

COMSAT

Technical Review

COMSAT Technical Review

Volume 13 Number 2, Fall 1983

COMSAT TECHNICAL REVIEW

Volume 13 Number 2, Fall 1983

Advisory Board Joseph V. Charyk
William W. Hagerty
John V. Harrington
John L. McLucas

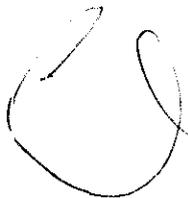
Editorial Board Pier L. Bargellini, Chairman
Ali E. Atia
S. J. Campanella
William L. Cook
H. W. Fliieger
Jorge C. Fuenzalida
William J. Getsinger
Geoffrey Hyde
R. W. Kreutel
Robert K. Kwan
Larry C. Palmer
James B. Potts
Akos G. Revesz
George R. Welti

Editorial Staff Daniel N. Crampton
MANAGING EDITOR
Margaret B. Jacocks
Pearl Coleman
Robert C. Heitman
Barbara J. Wassell
TECHNICAL EDITORS
Edgar Bolen
PRODUCTION
Shirley T. Cofield
CIRCULATION

- 221 CODED OCTAL PHASE SHIFT KEYING IN TDMA SATELLITE COMMUNICATIONS **S. A. Rhodes, R. J. Fang AND P. Y. Chang**
- 259 IMPACT OF SPEECH INTERPOLATION AND DATA ON FDM/FM TRANSMISSION **G. G. Szarvas AND S. J. Campanella**
- 275 WIDEBAND DIGITAL TRANSMISSION EXPERIMENTS IN THE INTELSAT V SYSTEM **D. Chakraborty, J. Ehrmann, R. Miner AND G. Rosch**
- 315 SERIAL IMPLEMENTATION OF VITERBI DECODERS **A. Shenoy AND P. Johnson**
- 331 ADAPTIVE SATELLITE POWER AMPLIFIER OPERATION FOR TDMA DOWN-LINKS **R. E. Eaves**
- 355 A COMPUTER PROGRAM FOR COMMUNICATIONS CHANNEL MODELING AND SIMULATION **A. Hamid, S. R. Baker AND W. Cook**
- 385 MULTIPLE SCATTERING IN COHERENT RADIO-WAVE PROPAGATION THROUGH RAIN **D. V. Rogers AND R. L. Olsen**
- 403 RAIN AND ICE DEPOLARIZATION MEASUREMENTS AT 4 GHz IN SITKA, ALASKA **S. J. Struharik**
- 437 ION IMPLANTATION OF BORON IN GAAS DEVICES **P. J. McNally**
- 451 TRANSLATIONS OF ABSTRACTS
FRENCH 451 SPANISH 457

COMSAT TECHNICAL REVIEW is published twice a year by Communications Satellite Corporation (COMSAT). Subscriptions, which include the two issues published within a calendar year, are: one year, \$10 U.S.; two years, \$15; three years, \$20; single copies, \$7; article reprints, \$1.50. Air mail delivery is available at an additional cost of \$10 per year. Make checks payable to COMSAT and address to Treasurer's Office, Communications Satellite Corporation, 950 L'Enfant Plaza, S.W., Washington, D.C. 20024, U.S.A.

© COMMUNICATIONS SATELLITE CORPORATION 1983
COMSAT IS A TRADE MARK AND SERVICE MARK
OF THE COMMUNICATIONS SATELLITE CORPORATION



Index: coding; error control; modulation, demodulation, and modems; synchronization; TDMA

Coded octal phase shift keying in TDMA satellite communications

S. A. RHODES, R. J. FANG, AND P. Y. CHANG

(Manuscript received November 19, 1982)

Abstract

Coded octal phase shift keying (coded OPSK) has been studied as a signaling technique for time-division multiple-access (TDMA) satellite communications. The performance of coded OPSK was investigated for forward error correction (FEC) codes of different complexities. Computer simulations were employed in the evaluation of probability of bit error, P_b , versus E_b/N_o for decoding of soft-detected OPSK symbols for different numbers of quantization levels. With the maximum likelihood (Viterbi) path decoder, coded OPSK provides a coding gain relative to the performance of uncoded quaternary phase shift keying (QPSK). Also, the use of a code rate of 2/3 allows two information bits to be conveyed by each OPSK symbol, thereby providing the same bandwidth efficiency as uncoded QPSK.

Significant coding gains of coded OPSK were obtained relative to uncoded QPSK for both an AWGN channel and a simulated INTELSAT V channel. Computer simulations showed that, for an FEC code of 64 states, coded OPSK transmitted over an AWGN channel yields a coding gain of about 3 dB when $P_b = 10^{-4}$. For $P_b = 10^{-6}$, the estimated coding gain for the AWGN case is approximately 4 dB. Corresponding coding gains for coded OPSK transmitted over the INTELSAT V channel were 3.2 dB at $P_b = 10^{-4}$ and 4.5 dB at $P_b = 10^{-6}$. It was also determined that coded OPSK performed somewhat more robustly than uncoded QPSK for large parabolic delay distortion in the input multiplexing filter of the simulated INTELSAT V satellite repeater.

Introduction

A study has been made of coded OPSK as a signaling technique for TDMA satellite communications. Performance evaluations are reported for Viterbi decoding [1], [2] of coded OPSK with convolutional FEC codes of rate $\rho = 2/3$ and of different complexities. These evaluations were made by computer simulations of coded OPSK for a channel with additive white Gaussian noise (AWGN channel) and for an INTELSAT V satellite channel. Investigations were also made to determine the decoding performance losses for different numbers of quantization levels.

In coded OPSK, the signaling scheme is coded phase modulation, in which the phase shifts are constrained so that only four possible phase changes are permitted during each signaling interval. Hence, only two information bits are conveyed by each octal modulation symbol. This signaling scheme inherently includes FEC coding of rate $\rho = 2/3$, since unconstrained octal signaling would yield three information bits per symbol. The coding constraint is conveniently introduced by a binary FEC code, and the output of the binary encoder is mapped into octal symbols for OPSK modulation.

Coded OPSK performance is expressed in terms of its coding gain, G (reduction in dB of the required E_b/N_o for a given probability of bit error, P_b), relative to the performance of uncoded QPSK. As E_b/N_o is increased, G increases and approaches the asymptotic coding gain G_a for vanishingly small P_b . Use of an FEC code rate of $2/3$ provides two information bits per OPSK symbol, the same as for uncoded QPSK. Thus, the coding gain is achieved without sacrificing bandwidth efficiency. A 120-Mbit/s information rate for coded OPSK can, therefore, be accommodated in a transponder channel allocation of 80 MHz (72-MHz "usable" transponder bandwidth).

A background for coded OPSK is given in the following section. Then, OPSK signaling is described, including the signaling constellation and the FEC codes of interest. Later, spectral results for coded OPSK are compared to those of uncoded QPSK at the output of the modem transmit filter and at the output of the high-power amplifier (HPA) of the transmitting station.

The Viterbi decoder is the most critical item with respect to high-speed implementation of coded OPSK. Decoder complexity is largely determined by the number of code states. Therefore, the performance of Viterbi decoding of coded OPSK was evaluated by computer simulations for codes of different complexities. Because of implementation considerations for high-speed applications, only codes with no more than 64 states were considered. Therefore, codes with 4, 16, and 64 states were investigated. Also, the number of quantization levels required for the soft-detected OPSK symbols contributes to decoder complexity. Decoding performance was investigated

for both 3- and 5-bit quantizations of the quadrature components of the received OPSK symbols.

Computer simulation results for the AWGN channel are presented for four different FEC codes. These codes are all of rate $\rho = 2/3$, with one code of 4 states, one code of 16 states, and two different codes of 64 states each. Performance sensitivity for coded OPSK to static phase errors in carrier synchronization was investigated for one of the codes of 64 states.

Results for the simulated INTELSAT V channel were first obtained for coded OPSK with an FEC code of 64 states. These results were obtained both for a single-channel simulation and for a four-channel simulation that included two adjacent channels and a co-channel as interferences for the desired channel.

Simulation results were also obtained for coded OPSK with three different FEC codes transmitted over the simulated INTELSAT V channel that included co-channel interference (CCI) and adjacent channel interference (ACI). The results of decoding probability of bit error P_b versus E_b/N_o for the INTELSAT V channel were obtained only with 32-level quantization of the detected quadrature components of the OPSK symbols.

Computer simulations were also performed to determine the decoding sensitivity of coded OPSK to delay distortion of the input multiplexing filter for the INTELSAT V satellite terminal. The sensitivity of uncoded QPSK performance was determined under the same conditions.

Background for coded OPSK

Most coherent digital communications presently use either binary or quaternary phase shift keying (BPSK or QPSK) modulation. When the information bit rate is limited by power considerations, FEC coding may be used to decrease the power requirement (E_b/N_o) necessary to achieve a desired bit-error probability, P_b . In such a case, modulation and FEC coding are treated as separate entities. The code redundancy that is necessary for error correction represents a bandwidth sacrifice in that the ratio of required bandwidth to information bit rate is increased by the use of FEC coding. Hence, bandwidth efficiency is traded for increased power efficiency.

Since modulation and FEC coding are closely related, bandwidth and power can be utilized more effectively by treating these two functions as a single entity. This combination of modulation and coding is referred to as coded trellis signaling (CTS). With CTS, the number of points in the signal constellation may be increased to some value M to provide redundancy for FEC coding without sacrificing bandwidth efficiency. Although CTS has M total modulation vectors, the code trellis may be constrained to only $\beta = 2^k$

branches, where $\beta < M$. The constraints of FEC coding allow only k information bits to be conveyed per M -ary symbol; hence, the effective code rate is $\rho = k/\log_2 M$.

Two main types of CTS have received considerable interest in recent years. The first type is digital FM or continuous-phase frequency shift keying (FSK) with specified modulation indices or frequency shifts. In this approach, the continuous-phase property results in a form of convolutional coding that increases the size of the modulation phase constellation at the detection sampling instants from 4 to M in a constrained manner which increases the minimum free Euclidean distance so that coding gain can be achieved. In the second type of CTS, a binary convolutional code of rate $\rho = k/n$ is used to produce code symbols with 2^n possible values from each input of k information bits, and the output code symbol is mapped into one of $M = 2^n$ possible modulation vectors for coded M -ary signaling. Any M -ary modulation can be employed, but this second type of CTS signaling is usually restricted to either quadrature-amplitude shift keying or phase shift keying. Memory of past bits in the encoder (usually consisting of shift registers) results in an increase in the minimum free distance of the CTS transmission.

One of the early papers on the distance properties of continuous-phase BPSK or binary FM by Pelchat, Davis, and Luntz [3] showed that certain modulation indices yielded minimum distances slightly greater than those for BPSK and QPSK. Minimum shift keying (MSK) [4] is a form of continuous-phase FSK or digital FM with a modulation index of 0.5 that provides the same minimum distance as BPSK and QPSK. DeBuda [5] developed a detection circuit and synchronizer configurations for differentially-encoded MSK, and named this signaling technique fast frequency shift keying (FFSK). In Reference 6, Osborne and Luntz determined the optimum detection configurations for continuous-phase FSK and presented a practical configuration for noncoherent detection of continuous-phase FSK. Schonhoff [7] investigated minimum distances versus frequency deviation for continuous-phase MFSK, with $M = 4$ and 8, and significant coding gains were shown to be possible relative to coherent QPSK. Also, Aulin and Sundberg [8] investigated the minimum distance properties of M -ary continuous-phase FSK for M values of 2, 4, 8, 16, and 32. Anderson and Taylor [9], [10] determined the minimum distances that could be obtained for binary FM with multiple choices of modulation index made in a cyclical manner. Later, this form of coding for continuous-phase FSK was named multi- h coding, and further work was reported by Anderson, Taylor, and Lereim [11], [12]. Aulin, Rydbeck, and Sundberg [13], [14] have reported on the bandwidth reductions and Euclidean distances that can be achieved from the use of partial-response signaling with continuous-phase FSK.

Considerable work on the second type of CTS has been done by Ungerboeck. In References 15, 16, and 17, he has determined Euclidean distances as a function of constraint length that can be obtained for several types of M -ary modulations when used in conjunction with FEC codes of rate $\rho = (n - 1)/n$, where $M = 2^n$. Hui and Fang [18]–[20] have obtained codes for coded CTS that have slightly larger minimum Euclidean distances than the Ungerboeck codes in some instances, and they have also developed coded CTS techniques that are suitable for severely bandlimited channels.

In CTS signaling, the size 2^n of the signaling alphabet must equal the size M of the modulation constellation. Thus, code rate of 1/2 would be suitable only for quaternary modulation, while codes of rate 3/4 would be for $M = 16$. Phase accuracies and resolution may be insufficient in the INTELSAT V system to support 16-ary PSK communications. Whereas rate 7/8 FEC coding can be employed to yield large minimum free Euclidean distances for coded QPSK signaling, the Viterbi decoder for maximum likelihood path decoding is too complicated for practical high-speed operation for codes of high rate that yield large coding gains.

Coded OPSK is a CTS method of the second type that is suitable for satellite communications. Some results from this application have been presented in conference papers [21], [22]. Also, there has been some recent work by Wilson [23] on coded 16-ary PSK with FEC coding of rate 3/4.

Coded QPSK with FEC coding of rate $\rho = 7/8$ is planned for selective use in the INTELSAT V TDMA implementation. A 60-Msymbol/s rate will be employed for the 72-MHz usable bandwidth of the INTELSAT V transponders, which have frequency allocations of 80 MHz each. With $\rho = 7/8$, the modem rate of 120 Mbit/s yields an information rate of only 105 Mbit/s.

It is desirable to improve both the bandwidth and power efficiencies for INTELSAT V TDMA communications. The purpose of this study is to evaluate the performance of a suitable signaling technique that can accomplish these two goals. Coded OPSK can be used for the INTELSAT V channel at a signaling rate of 60 Msymbol/s. Therefore, the FEC code rate of $\rho = 2/3$ for coded OPSK will provide two information bits per symbol for an information rate of 120 Mbit/s. Also, soft detection and Viterbi decoding for coded OPSK yield greater coding gain than the Bose-Chaudhuri-Hocquenghem (BCH) code of $\rho = 7/8$ used with hard decisions on the QPSK symbols. Consequently, coded OPSK can improve both bandwidth efficiency and power efficiency over that of the present coded QPSK for INTELSAT V TDMA communications. The potential performance of coded OPSK for INTELSAT V communications is evaluated in this study primarily by computer simulations.

Signal generation, FEC codes, and decoding

A description of the signal generation of coded OPSK will now be presented, and the four FEC codes used in the computer simulations will be described. The discussion will include soft detection and Viterbi decoding.

Decoding performance for CTS depends on the distinguishability of the different possible sequences of coded modulation symbols. The best measure of code sequence distinguishability is the minimum Euclidean distance squared, d_m^2 . For binary antipodal signaling, such as for BPSK and QPSK with Gray mapping of the binary code digits to the four phase positions, the minimum Hamming distance, h_m , for binary FEC coding is directly proportional to d_m^2 . However, FEC codes that provide large values of h_m do not necessarily yield large d_m^2 values for coded OPSK. Hence, codes for use with OPSK modulation must be selected on the basis of d_m^2 rather than h_m .

For any two code sequences, the Euclidean distance squared is equal to the sums of the Euclidean distances squared for the corresponding symbols of each sequence. Thus, for sequences A and B that merge at the i th code symbol,

$$d_{AB}^2 = \sum_{i=0}^{I-1} [(X_{Bi} - X_{Ai})^2 + (Y_{Bi} - Y_{Ai})^2].$$

The minimum value of d^2 for any two code sequences is d_m^2 .

Signal generation

Figure 1 is a block diagram for the generation of coded OPSK. Note that information bits are input to the encoder in pairs for a code with rate $\rho = 2/3$. During each encoding step, three binary code symbols are obtained from certain modulo-2 sums of the present input pair of bits and other information

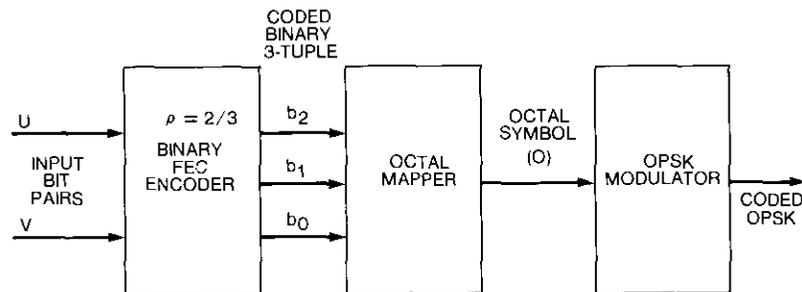


Figure 1. Block Diagram of Signal Generation for Coded OPSK

bits within the memory of the encoder. Each output binary 3-tuple must then be mapped into an octal symbol for OPSK signal generation. Any unique mapping of 3-tuples into octal symbols is permissible if encoder connections are used that maximize d_m^2 for the selected mapping function. However, if the encoder is specified as one that yields a large minimum Hamming distance, then a mapping such as Gray encoding [24] should be used so that a large Hamming distance also corresponds to a large Euclidean distance.

Figure 2 illustrates the OPSK signal constellation. The 3-tuple for each OPSK signal point is based on Gray encoding, for which adjacent signal

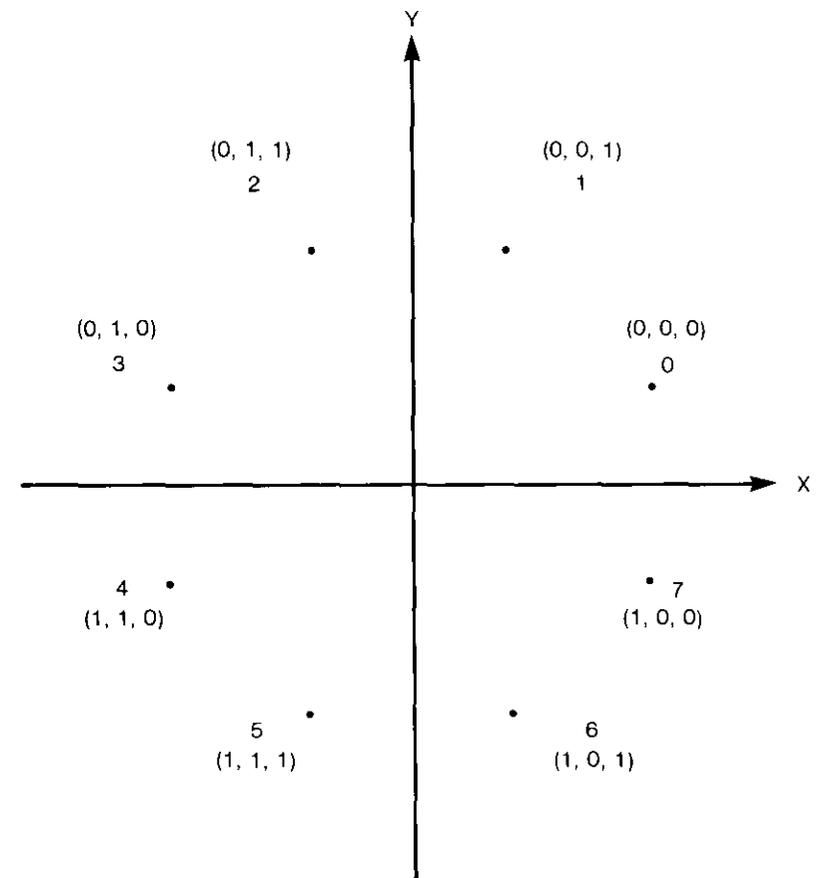


Figure 2. OPSK Signal Constellation With Gray Mapping of Coded Binary 3-Tuple

points differ in only one of the three components of the binary 3-tuple. In some cases, the FEC codes for coded OPSK will use a natural mapping of binary words into octal values.

In the OPSK signaling space of Figure 2, the eight possible symbol values are uniformly distributed around a circle of radius $\sqrt{E_s} = \sqrt{2E_b}$, where E_s is signal energy per OPSK symbol and E_b is the energy allocated to each of the two information bits conveyed by each coded OPSK symbol. Note that the quadrature components, X and Y , can take on values of only $\pm 0.383\sqrt{E_s}$ and $\pm 0.924\sqrt{E_s}$. Also, when the absolute value of one coordinate is $0.383\sqrt{E_s}$, the other component must have an absolute value of $0.924\sqrt{E_s}$.

There are five values of Euclidean distance between the various octal symbols. These distances may be expressed as d_{0j} for different j , the distance between the "0" octal symbol and the "j" octal symbol. In terms of the angle θ_{0j} between the j th symbol and the 0 symbol in the signal constellation of Figure 2,

$$d_{0j} = 2\sqrt{E_s} \sin\left(\frac{\theta_{0j}}{2}\right) \tag{1}$$

where $\theta_{0j} = 0.25j\pi$ radians. With $E_s = 2E_b$, the squared Euclidean distances for OPSK are as follows:

$$\begin{aligned} d_{00}^2 &= 0 \\ d_{01}^2 &= d_{07}^2 = 0.58 E_s = 1.16 E_b \\ d_{02}^2 &= d_{06}^2 = 2.00 E_s = 4.00 E_b \\ d_{03}^2 &= d_{05}^2 = 3.42 E_s = 6.84 E_b \\ d_{04}^2 &= 4.00 E_s = 8.00 E_b \end{aligned} \tag{2}$$

FEC codes suitable for coded OPSK

Four convolutional FEC codes of rate $\rho = 2/3$ were used in the present computer simulations of coded OPSK. These codes differed in constraint length, with $K = 4, 6,$ and 8 bits. Thus, $\gamma = K - 2 = 2, 4,$ and 6 bits are the memories of the four codes. Corresponding to code memory γ , a code will have $S = 2^\gamma$ code states. Therefore, the codes used in the simulations were for code complexities of $S = 4, 16,$ and 64 states. Two of the codes have $\gamma = 6$ bits memory, or $S = 64$ states. However, these two codes of equal complexity have different minimum distances.

The first of the four codes has $\gamma = 2$ bits memory, which results in only $S = 4$ code states. This code is optimum with respect to minimum free Euclidean distance for $\gamma = 2$ when used with natural binary-to-octal mapping

into coded OPSK. The square of the minimum free Euclidean distance is $d_m^2 = 8E_b$, where E_b is the energy per information bit. For uncoded QPSK, which has the same bandwidth efficiency as coded OPSK, the minimum Euclidean distance squared is $d_1^2 = 4E_b$. Thus, the asymptotic coding gain relative to uncoded QPSK is 3 dB for coded QPSK using this code. The code was obtained from results from Ungerboeck [15]–[17].

Code 2, with $\gamma = 4$ bits memory, is an optimum code with respect to minimum free Hamming distance, h_m . For this code, described by Paaske [25], $h_m = 5$. With Gray mapping to OPSK, the code is also optimum [18] for minimum free Euclidean distance, and $d_m^2 = 10.34 E_b$, which gives an asymptotic coding gain of $G_a = 4.1$ dB relative to uncoded QPSK.

The third code, with $\gamma = 6$ bits memory, was also an optimum code for minimum free Hamming distance as obtained by Paaske. For this code, $h_m = 7$. With Gray mapping to OPSK, the minimum free Euclidean distance is $d_m^2 = 11.52 E_b$, for an asymptotic coding gain of 4.6 dB relative to uncoded QPSK. However, this code is not optimum in the sense of maximizing the minimum Euclidean distance, as shown in References 18, 19, and 20. Hui [18], [19] has obtained several codes of rate $2/3$ and memory $\gamma = 6$ that maximize the minimum Euclidean distance and yield asymptotic coding gains of 5.0 dB for coded OPSK. Code 4 is one of these codes.

For coded OPSK, the encoder consists of two shift registers with tap connections to three modulo-2 adders. Figure 3 illustrates the encoder

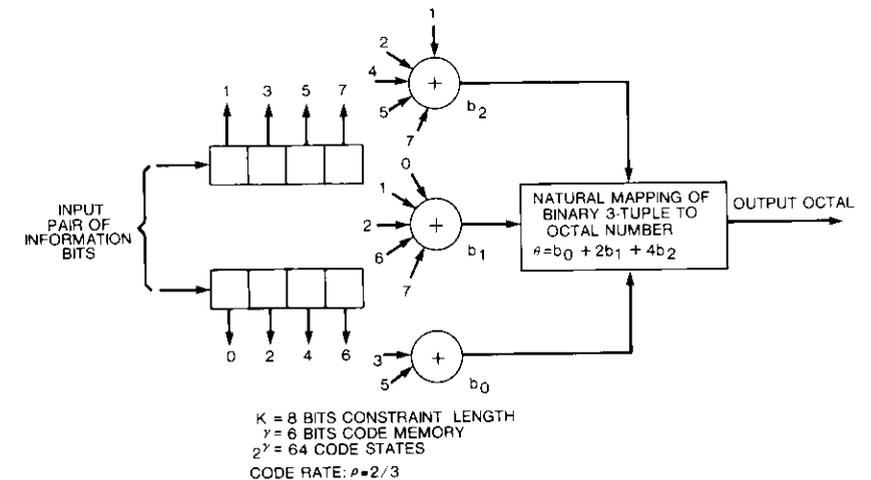


Figure 3. Code 4: Hui Code of Rate $\rho = 2/3$ and Memory of $\gamma = 6$ Bits With Natural Mapping to Coded OPSK

connections for one of the four convolutional codes used in the computer simulations of coded OPSK. Encoder connections for all four codes are given by their code polynomials in Table 1. An element a_{ijk} of a code polynomial denotes whether or not the j th tap of the i th shift register is connected to the k th adder. The important characteristics of the four codes are summarized in Table 1. "Natural" mapping (as opposed to Gray mapping) refers to the straightforward encoding of a binary word (b_2, b_1, b_0) into an octal number θ by

$$\theta = \sum_{i=0}^2 2^i b_i = b_0 + 2b_1 + 4b_2.$$

TABLE 1. COEFFICIENTS a_{ijk} FOR ENCODER POLYNOMIALS OF CONVOLUTIONAL FEC CODES USED IN COMPUTER SIMULATIONS OF CODED OPSK

CODE	CONSTRAINT LENGTH K	REGISTER NO. i	TAP NUMBER j											
			1			2			3			4		
			ADDER NO. k	ADDER NO. k	ADDER NO. k	ADDER NO. k	ADDER NO. k	ADDER NO. k	ADDER NO. k	ADDER NO. k	ADDER NO. k	ADDER NO. k	ADDER NO. k	
1	4	1	0	1	0	0	0	1	0	1	0	---	---	---
		2	1	0	0	---	---	---	---	---	---	---	---	---
2	6	1	1	0	1	1	0	0	1	1	0	---	---	---
		2	0	1	1	1	0	1	0	1	1	---	---	---
3	8	1	1	0	1	0	1	1	0	1	0	1	0	1
		2	0	1	1	1	1	1	1	0	1	0	1	1
4	8	1	1	1	0	0	0	1	1	0	1	1	1	0
		2	0	1	0	1	1	0	1	0	0	0	1	0

Soft detection and decoding

Figure 4 is a block diagram for the detection and decoding of coded OPSK transmissions. A quadrature receiver structure is illustrated. Carrier synchronization provides a coherent reference for demodulating the quadrature components of the OPSK symbols. Symbol synchronization then determines the appropriate sampling times for the filtered quadrature components. Samples \hat{X} and \hat{Y} of the quadrature components are individually quantized and sent to a Viterbi decoder. During each OPSK symbol interval, the Viterbi decoder, which performs maximum-likelihood path decoding, outputs two bit decisions, \hat{U} and \hat{V} .

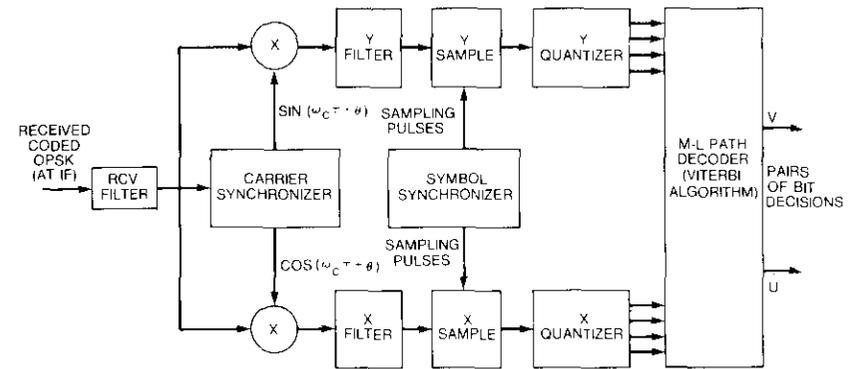


Figure 4. Block Diagram for Detection and Decoding of Coded OPSK

In the computer simulation program, the quadrature components of the received OPSK symbols are represented by quantized numbers. This soft quantization using b digits corresponds to sampling followed by analog-to-digital (A/D) conversion into 2^b levels. The computer simulation program allows selection of quantization from $b = 1$ bit, or hard binary decisions, up to $b = 5$ bits, or 32 quantization levels.

As an example of levels, 3-bit soft detection is represented by levels of $-7, -5, -3, -1, +1, +3, +5,$ and $+7$. The quantizer scaling is adjusted by selection of the Q setting, which represents the clipping level relative to the nominal signal level of $\sqrt{E_s}$, where E_s is the signal energy per OPSK symbol. In the designation of clipping level, selection of $Q = 1.2$ for 3-bit soft detection would result in the levels of -7 and $+7$ beginning at $-1.2\sqrt{E_s}$ and $+1.2\sqrt{E_s}$, respectively.

Decoder complexity is the primary consideration with respect to the feasibility of a high-speed implementation of coded OPSK signaling. The primary measure of decoder complexity is the number of code states, $S = 2^\gamma$. The required quantization also contributes to complexity in that the number of bits required for branch and path metrics depends on the number of quantization levels. Thus, it is important to determine the decoding performance (P_b versus E_b/N_0) as a function of the number of code states and the number of quantization levels.

Bandwidth requirements for coded OPSK

As with uncoded QPSK, coded OPSK conveys two information bits per symbol. Also, the power spectrum of coded OPSK has the same $\text{sin}^2(\pi f T_s)/$

$(\pi f T_s)^2$ function as uncoded QPSK, where $R_s = 1/T_s$ is the modulation symbol rate. Therefore, the spectral efficiency of coded OPSK should be about the same as for uncoded QPSK.

The effective bandwidth efficiencies can only be compared in terms of losses in performance of P_b versus E_b/N_o , when adjacent channel transmissions cause interference in addition to the intersymbol interference (ISI) caused by the bandwidth constraints of filtering. However, one important factor for bandwidth requirements in satellite communications can be studied by comparing power spectra alone. This aspect is the spectral regrowth for filtered signals when passed through a nonlinearity. Of particular interest is the spectrum of the transmitted signal from the ground terminal.

The modem transmit filter is used to restrict the spectrum of the transmitted signal. However, this filtered signal must then be amplified by the HPA, which introduces some spectral regrowth. Modem filtering for INTELSAT V TDMA communications is specified as Nyquist with an $\alpha = 0.4$ rolloff factor. For matched filtering under modem back-to-back operation, the Nyquist response is divided equally between the transmit and receive filters. Also, the transmit filter will contain an aperture compensation response of the form $(\pi f/R_s)/[\sin(\pi f/R_s)]$, where R_s is the QPSK symbol rate, if rectangular pulses of duration $T = 1/R_s$ are employed as inputs to the transmit filter. The filter output goes to the earth station HPA (coupled-cavity TWTA), which is assumed to have the nonlinear characteristics shown in Figure 5.

Computer simulations were used to compare spectral regrowths for coded OPSK and uncoded QPSK. Figure 6 illustrates the spectral results for various input backoffs (IBOs) of the HPA. Both modulations produce symmetrical spectra about the carrier frequency. For comparison, only one side of each power spectrum is shown, with uncoded QPSK results on the left and coded OPSK results on the right. The biggest contributor to spectral regrowth is the 180° phase transition. Such a transition has a probability of $1/8$ for coded OPSK and of $1/4$ for uncoded QPSK. Consequently, the spectral regrowth of sidelobes was slightly less for coded OPSK. Prior to filtering (point A) the spectra are those for rectangular pulses. After filtering (point B) the main spectral lobe is narrowed somewhat and the spectral sidelobes are suppressed. There is increased regrowth of sidelobes (point C) as the HPA IBO is reduced.

Coded OPSK performance for the AWGN channel

Computer simulations of coded OPSK transmitted over an AWGN channel were made for all four FEC codes of Table 2. The simulations for the first three codes were made for 5- and 3-bit quantizations of the quadrature

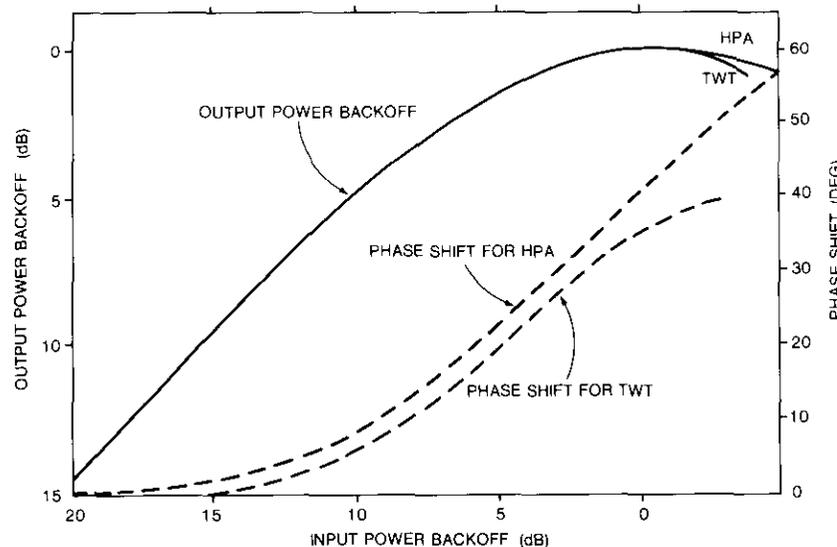


Figure 5. Nonlinear Response Curves for Earth Station HPA and Satellite TWTA of INTELSAT V

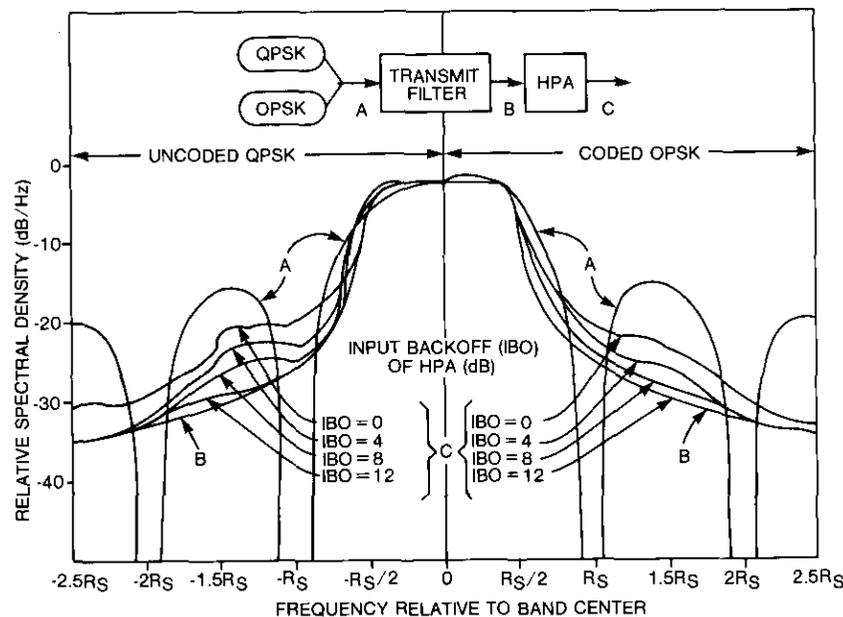


Figure 6. Comparison of Spectral Regrowth of Coded OPSK and Uncoded QPSK

TABLE 2. FEC CODES USED IN COMPUTER EVALUATIONS OF CODED OPSK

CODE	CODE RATE ρ	CONSTRAINT LENGTH K (Bits)	MAPPING TO OPSK	CODE MEMORY γ (Bits)	NUMBER OF CODES OF STATES S	ASYMPTOTIC CODE GAIN G_a (dB)
1	2/3	4	Natural	2	4	3.0
2	2/3	6	Gray	4	16	4.1
3	2/3	8	Gray	6	64	4.6
4	2/3	8	Natural	6	64	5.0

components of the received OPSK symbols. For both quantizations, the scaling factor Q was adjusted to minimize the bit-error probability of Viterbi decoding in the region of $P_b = 10^{-2}$. Because of the robustness of the decoding with respect to this scaling, it is not necessary to adjust Q differently at lower P_b values. Code 4 was simulated only for 5-bit quantization.

Figures 7, 8, and 9 give P_b versus E_b/N_o for Codes 1, 2, and 3 of Table 2. Results for all three codes indicate that the performance for 3-bit quantization is about 0.4-dB worse than that for 5-bit quantization in the region of $P_b = 10^{-4}$. Use of 4-bit quantization would approximately halve this degradation, which would result in a more acceptable loss. Hence, 4-bit soft detection is a more reasonable choice than 3-bit.

Figure 10 compares the performance of coded OPSK for the four codes when 5-bit quantization is employed. Code 1 has only four states and provides an asymptotic coding gain of only 3.0 dB. The actual coding gain at $P_b = 10^{-4}$ is about 2.0 dB. Use of Code 2, with 16 states, increases the asymptotic coding gain to 4.1 dB. Also, the gain at $P_b = 10^{-4}$ is increased considerably to 2.6 dB. However, use of Code 3 with 64 states does not yield as much improvement. For this code, the asymptotic gain is 4.6 dB and the coding gain is about 2.9 dB at $P_b = 10^{-4}$. Code 4 has the same complexity as Code 3, with 64 states, but its asymptotic coding gain is 5.0 dB. Consequently, the coding gain at $P_b = 10^{-4}$ is increased to 3.1 dB by the use of Code 4.

Because of the time required for the simulations on a Perkin Elmer 7/32 minicomputer to accumulate enough decision errors for statistical reliability in predicting decoding performance, it is not feasible for results to be obtained at $P_b < 10^{-4}$. Hence, the P_b curves must be extrapolated to estimate coding gain at $P_b = 10^{-6}$. These estimates for the four codes are 2.5, 3.3, 3.7, and 3.9 dB. The estimated coding gains for coded OPSK transmitted over the AWGN channel at various P_b values are summarized in Table 3 for the four codes.

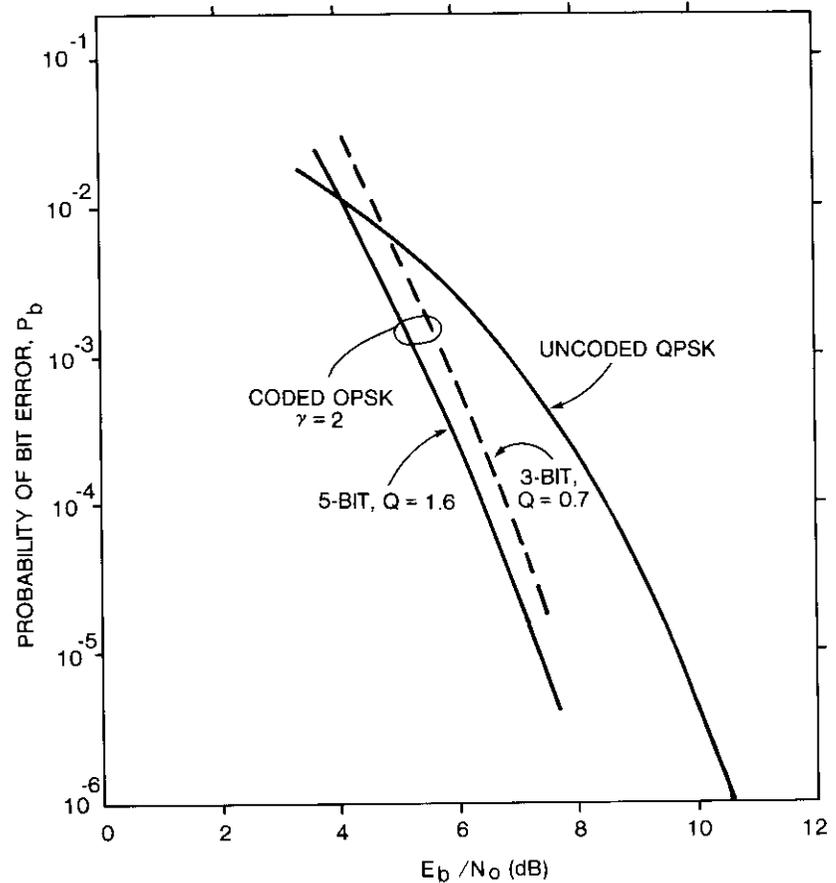


Figure 7. Performance of Coded OPSK for AWGN Channel and Code 1

TABLE 3. CODING GAINS WITH RESPECT TO UNCODED QPSK OF CODED OPSK OVER AWGN CHANNEL WITH 5-Bit SOFT DETECTION OF QUADRATURE COMPONENTS

P_b	CODING GAIN* (dB)			
	CODE 1	CODE 2	CODE 3	CODE 4
10^{-2}	0.2	0.2	0.2	0.2
10^{-3}	1.5	1.8	1.9	2.0
10^{-4}	2.0	2.6	2.9	3.1
10^{-5}	2.3	3.0	3.4	3.6
10^{-6}	2.5	3.3	3.7	3.9

* Coding gains estimated from curve extrapolations when $P_b < 10^{-4}$.

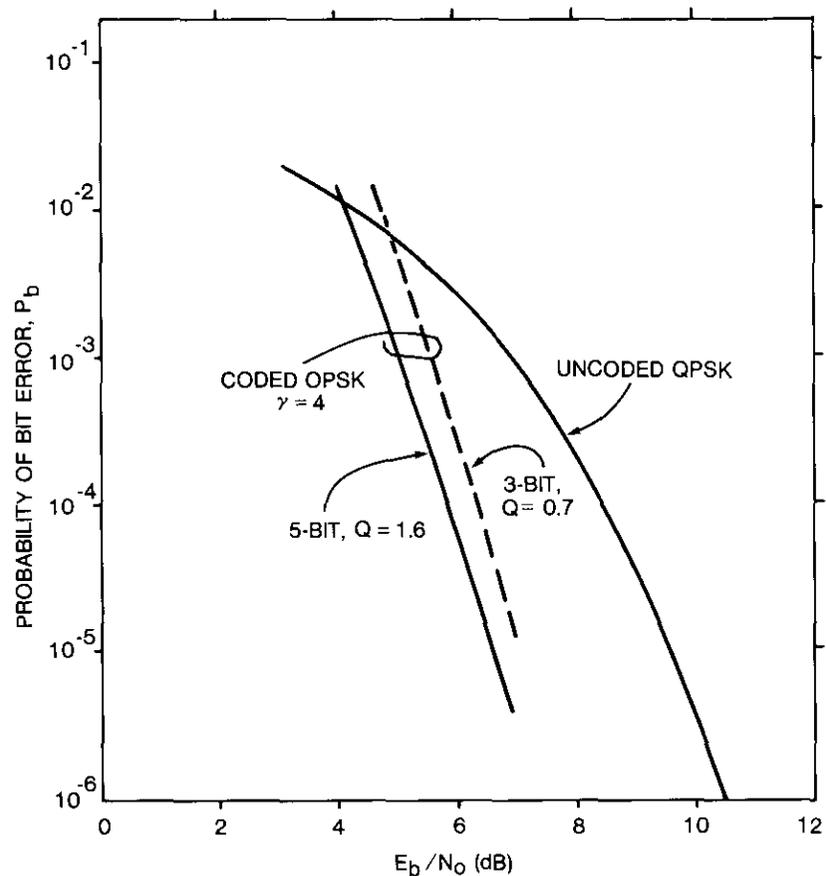


Figure 8. Performance of Coded OPSK for AWGN Channel and Code 2

As a further check on quantization, coded OPSK using Code 3 was simulated for 3-, 4-, and 5-bit quantizations. Figure 11 shows P_b versus E_b/N_0 for these simulations over the AWGN channel. The results for 5-bit soft detection are probably very close to the theoretical curve for this particular coded OPSK signaling with infinite quantization. Note that 4-bit soft detection gave a performance that was only about 0.1-dB worse than that for the 5-bit case. Also, 3-bit soft detection caused an additional loss of only about 0.2 to 0.3 dB. With 2-bit quantization, the performance was unsatisfactory. Although 3-bit quantization appears to be satisfactory for the two quadrature components of the octal symbols, 4 bits may be necessary when the channel contains

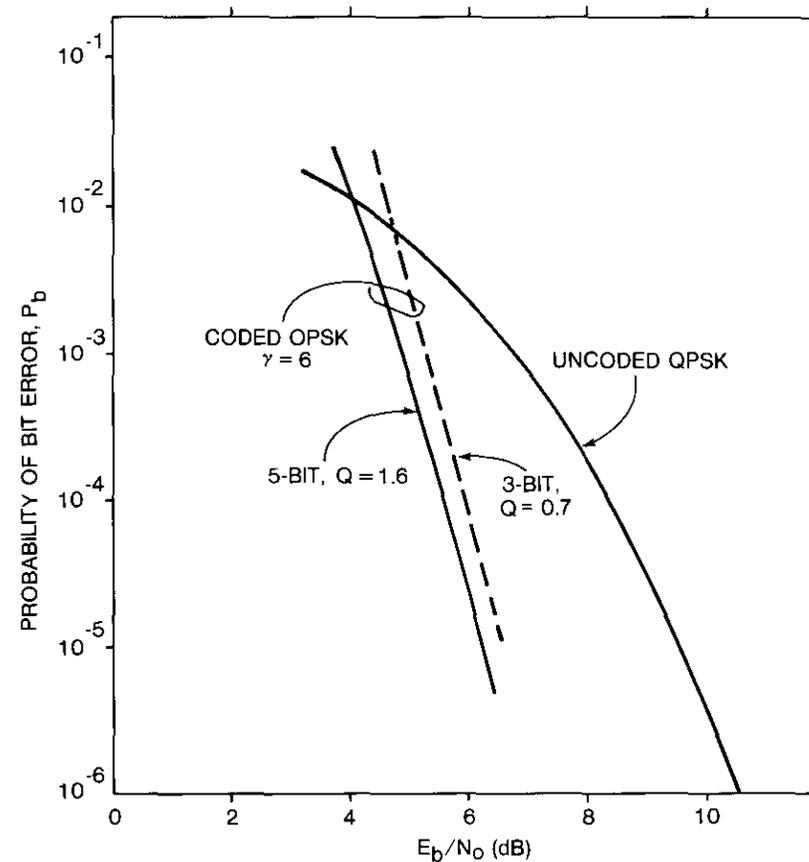


Figure 9. Performance of Coded OPSK for AWGN Channel and Code 3

filters and nonlinearities. A path memory or trellis storage of $L = 160$ bits, or path length $PL = 80$ OPSK symbol intervals, was used in the computer simulations for the results given in Figure 11.

In addition to coded OPSK results, Figure 11 also depicts the theoretical performances for uncoded coherent QPSK and uncoded coherent OPSK. Uncoded 8-phase simulation results for an AWGN channel were obtained to ensure that the computer program had the correct noise scaling for each E_b/N_0 , and these results were in agreement with theory.

The computer simulations of coded OPSK and uncoded QPSK were made with perfect synchronization of carrier phase and symbol timing. It is known that the close spacing of 45° between adjacent symbols causes uncoded OPSK

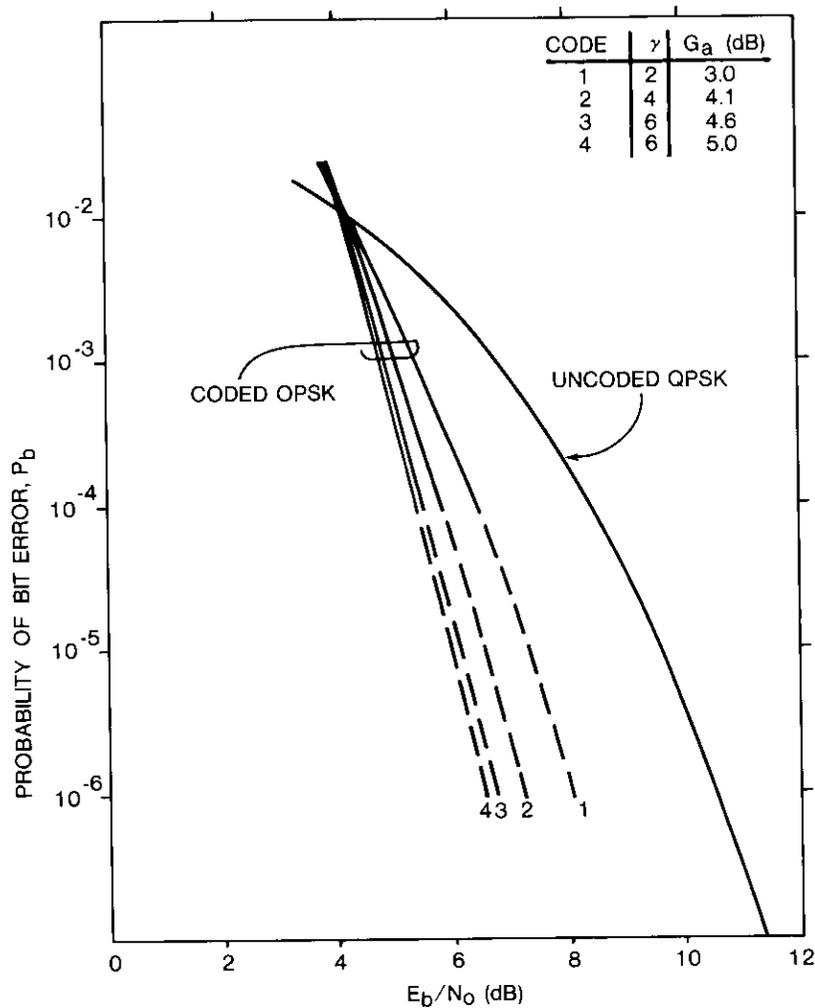


Figure 10. Coded OPSK Performance for AWGN Channel Parametric in Code Complexity, With Soft Detection of 32 Levels for OPSK Quadrature Components

performance to be extremely sensitive to errors in carrier phase. An evaluation was made to determine how much this sensitivity was reduced by the path decoding for coded OPSK.

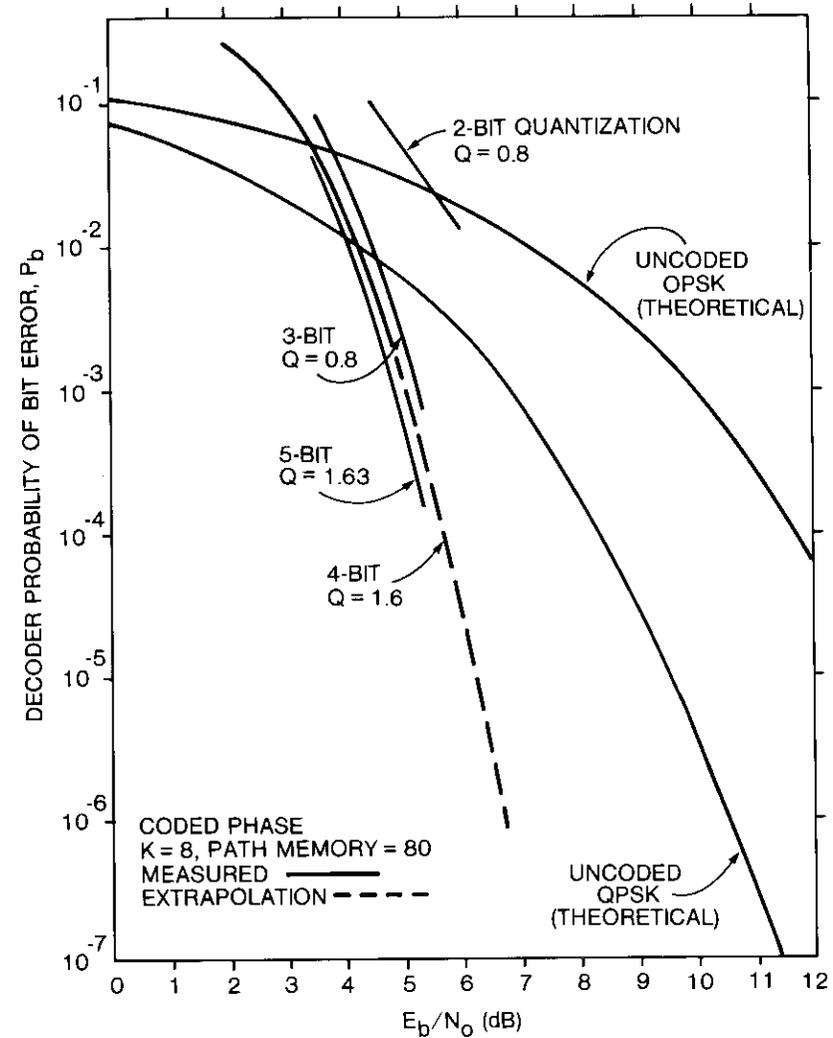


Figure 11. Sensitivity of Decoding Performance to the Number of Detection Quantizer Levels for Coded OPSK Transmitted Over an AWGN Channel Using Code 3

Computer simulations were performed to assess the decoding losses of coded OPSK associated with static phase errors, ϕ , of 5° and 10° when Code 4, with 64 states and an asymptotic coding gain of 5.0 dB, was employed.

These simulations of coded OPSK transmitted over an AWGN channel resulted in an almost constant loss of 0.4 dB when $\phi = 5^\circ$ and of 1.4 dB when $\phi = 10^\circ$ for a range of P_b from 10^{-2} to 10^{-4} . Table 4 compares the performance losses for coded OPSK, uncoded OPSK, and uncoded QPSK at $P_b = 10^{-4}$. Also, asymptotic losses at very low P_b are given as determined analytically. Coded OPSK performance is seen to have a moderate sensitivity to carrier phase error similar to that of uncoded QPSK, rather than the extreme sensitivity exhibited by uncoded OPSK.

TABLE 4. COMPARISON OF PERFORMANCE SENSITIVITIES TO STATIC ERRORS IN CARRIER PHASE SYNCHRONIZATION

SIGNALING SCHEME	CARRIER PHASE ERROR ϕ	PERFORMANCE LOSS WHEN $P_b = 10^{-4}$ (dB)	ASYMPTOTIC LOSS FOR VERY LOW P_b (dB)
Coded OPSK	5°	0.4	0.64
	10°	1.4	1.43
Uncoded OPSK	5°	1.6	2.10
	10°	4.5	4.95
Uncoded QPSK	5°	0.5	0.83
	10°	1.4	1.82

Coded OPSK performance for the simulated INTELSAT V channel

Computer simulation results for coded OPSK transmitted over the simulated INTELSAT V channel were first obtained for the single-channel case. Only Code 3 of Table 2 was employed for these initial results. A co-channel transmission and two adjacent channels were then introduced to determine the effect of CCI and ACI on coded OPSK using Code 3.

After the Code 3 tests, the four-channel case (the desired channel plus a co-channel and two adjacent channels) was also simulated for Codes 3 and 4. No coded OPSK simulations for the INTELSAT V channel were made for Code 1, which is considered to be of insufficient complexity to yield significant coding gains.

Channel description

Modem filtering provides a Nyquist response [26] with $\alpha = 0.4$ for the rolloff factor. The modem filters include both a square-root Nyquist response

and an aperture compensation response for the transmit filter and a square-root Nyquist response for the receive filter. The aperture response is an inverse $\sin(\pi f/R_s)/(\pi f/R_s)$ that compensates for the use of rectangular pulse inputs, instead of impulses, to the transmit filter.

Figure 5 gives the nonlinear responses of the HPA for the transmit terminal and the traveling wave tube amplifier (TWTA) of the satellite transponder. Both of these devices are nonlinear in output amplitude and phase versus input amplitude. The nonlinear responses are based on the use of a coupled-cavity amplifier for the transmit HPA and a helix-type amplifier for the satellite TWTA. Also, both amplifiers are modeled as memoryless. In the simulations, either an 8.5- or 12-dB IBO was used for the HPA, and the decoding results were almost identical for both operating points. The IBO for the TWTA was 2 dB in all cases.

Figures 12 and 13 give the responses of amplitude and delay vs frequency for the input and output multiplexer filters of the satellite transponder. Both the amplitude and delay functions are for the nominal responses of these filters. The delay distortions are small and cause insignificant performance losses.

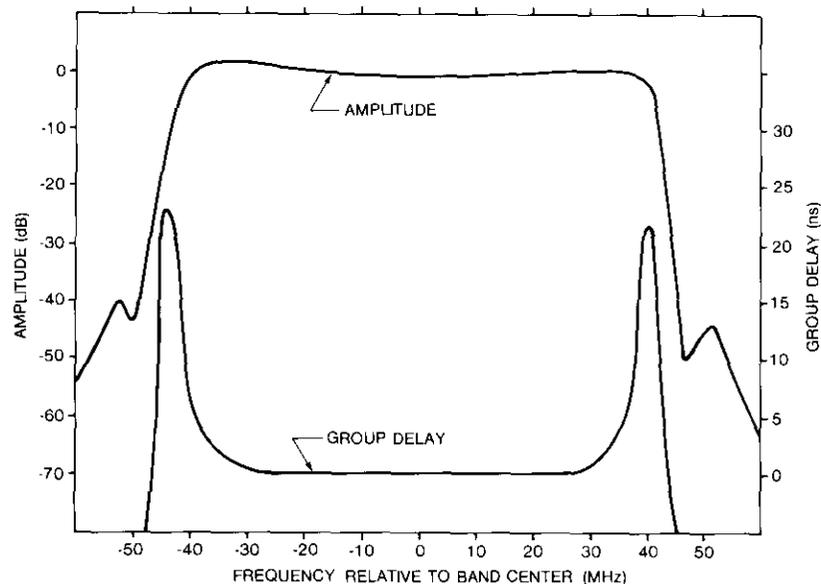


Figure 12. Nominal Amplitude and Delay Responses for the Satellite Input Multiplexing Filter

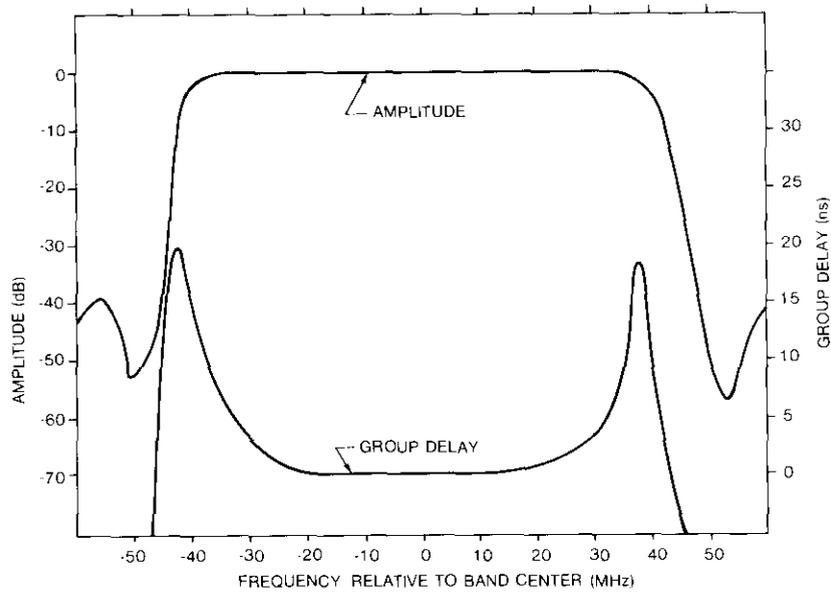


Figure 13. Nominal Amplitude and Delay Responses for the Satellite Output Multiplexing Filter

Exact carrier synchronization and symbol timing are employed in the computer simulations of coded OPSK. Hence, the decoder results do not include any losses caused by imperfect synchronization and are thus for ideal, coherent, synchronous detection.

Figure 14 illustrates the channel elements for the computer simulation of the four-channel case, which includes CCI and ACI. The signaling rate is 60 Msymbol/s for an information rate of 120 Mbit/s for either uncoded QPSK or coded OPSK. Responses of the satellite multiplexing filters yield a usable bandwidth of 72 MHz per transponder. The allocated bandwidth or channel spacing is 80 MHz per transponder. All CCI is lumped into the up-link at an interference power level 18.5 dB below that of the desired transmission. Multipath transmission of the desired signal through adjacent transponders is included in the simulation by means of the overlapping frequency responses of the transponder multiplexing filters. Regrowth of spectral sidelobes because of the HPA nonlinearity compounds the multipath problem.

Simulation results for coded OPSK using Code 3

Code 3 was simulated first for coded OPSK over an INTELSAT V channel both with and without CCI and ACI. As in all of the simulations that include

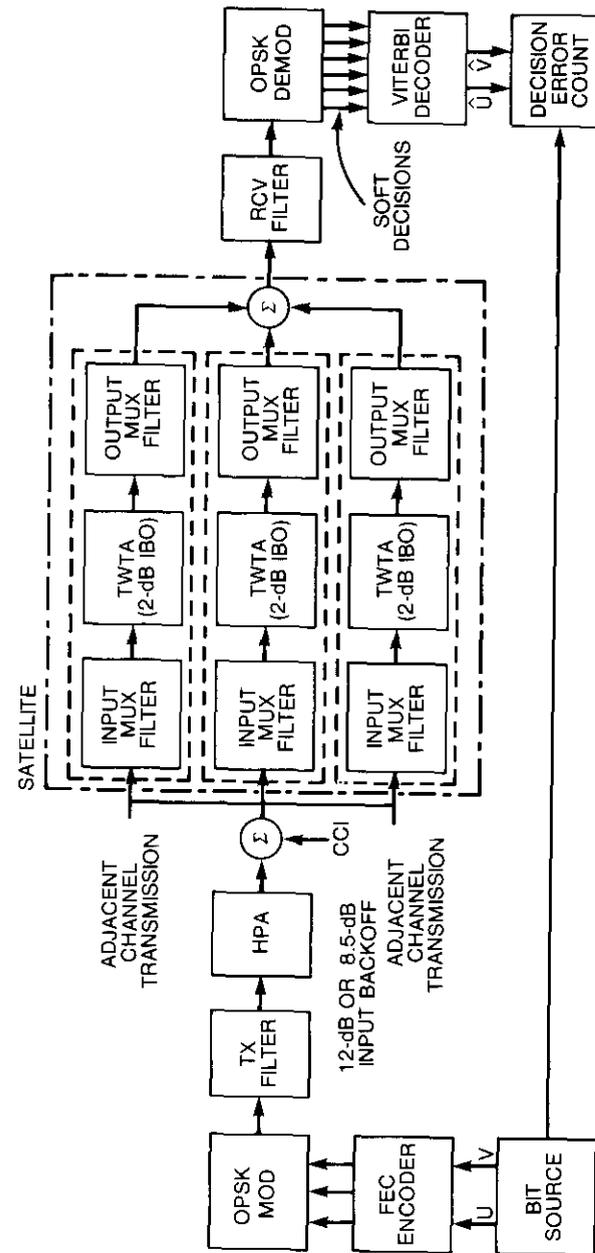


Figure 14. Block Diagram of Circuits for Coded OPSK Simulation Over the INTELSAT V Satellite Channel

the simulated INTELSAT V channel, 5-bit or 32-level quantization was employed. Equal spacing was used between soft-detection thresholds, and the highest threshold corresponded on an absolute basis to about twice the maximum voltage level in the absence of noise. With this quantization of the quadrature components of the detected OPSK symbols, the Viterbi algorithm yields nearly ideal maximum-likelihood path decoding.

Perturbations of the signal levels by ISI or other interferences cause a reduction in distance between some OPSK symbols, and performance is considerably worse than for the purely AWGN channel. In a nonlinear band-limited channel, the saturation power levels of nonlinear amplifiers prevent constructive ISI from increasing the signal level as much as it is lowered by destructive ISI. Although the signal scatter from ISI is actually asymmetrical for this nonlinear case, the ISI can be considered to result in approximately symmetrical scattering about a reduced mean signal level. The reduced mean causes a common reduction in Euclidean distance for both coded and uncoded signaling. However, the scattering about the reduced mean in any code sequence contains some symbols with constructive ISI and some with destructive ISI. These constructive and destructive contributions tend to be averaged in path decoding so that the Euclidean distances between code paths are not reduced as much as the distances for uncoded communications. Consequently, the Viterbi decoder does not usually suffer as much performance loss from interferences as does uncoded QPSK or any other communications method that does not employ path decoding. Therefore, the coding gain of coded OPSK is somewhat greater for the nonlinear, band-limited satellite channel than for the AWGN case.

Figure 15 compares performances for coded OPSK and uncoded QPSK. Code 3, with a memory of $\gamma = 6$ bits and $S = 64$ states, has an asymptotic coding gain of 4.6 dB for the AWGN channel. Coding gain is about 3.1 dB for the simulated INTELSAT V channel at $P_b = 10^{-4}$, as opposed to 2.9 dB for Code 3 when used for the AWGN channel. Additional losses caused by ACI and CCI were approximately the same for coded OPSK and uncoded QPSK. Thus, the coding gains are similar for the single-channel and four-channel simulations.

The results presented in Figure 15 were for an HPA IBO of 12 dB. Computer simulations were also made in which the HPA IBO was 8.5 dB. For these simulations, the P_b versus E_b/N_o results were virtually identical. Thus, performance was almost independent of HPA backoff in the quasi-linear region of operation that includes 8.5- and 12-dB IBOs. In Figure 15, coded OPSK results for $P_b < 10^{-4}$ are obtained by curve extrapolation.

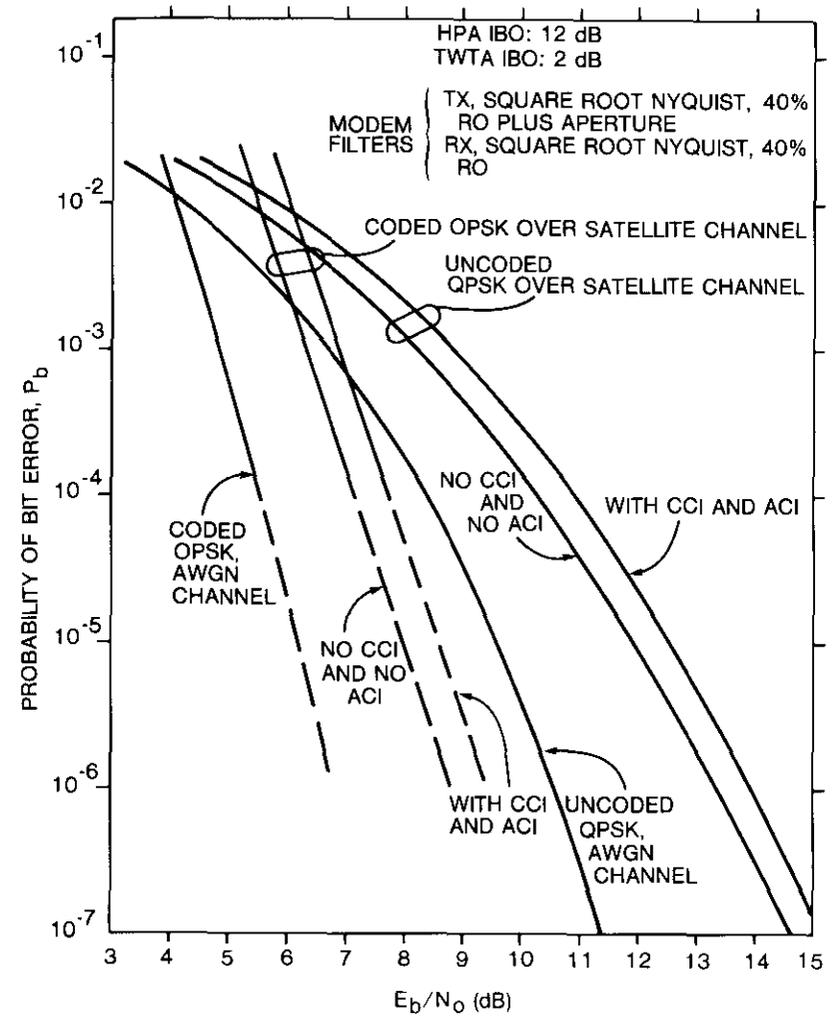


Figure 15. Coded OPSK Performance for the INTELSAT V Channel Using Code 3

Coded OPSK results for INTELSAT V using Codes 2, 3, and 4

Computer simulations were performed for coded OPSK using Codes 2, 3, and 4. The simulations were for the four-channel case (Figure 14) that

included CCI and ACI. Results were obtained as P_b versus E_b/N_o and were compared to the performance of uncoded QPSK.

The results for the present simulations of coded OPSK are for a TWTA IBO of 2 dB and an HPA IBO of 8.5 dB. The simulation included both adjacent channels and a co-channel interferer. The adjacent channels were of the same level as the desired transmission to simulate clear-sky propagation conditions. All CCI was represented by a single up-link co-channel transmission at a level 18.5 dB below that of the desired transmission. Multipath transmission of the desired signal through adjacent transponders is also simulated by the overlapping responses of the transponder multiplexing filters.

Coded OPSK can provide considerably more coding gain by using codes that have at least 16 states. Therefore, the shorter code (Code 1 of Table 2) was not employed in the coded OPSK computer simulations for the INTELSAT V channel. The simulation results for Codes 2, 3, and 4 are plotted in Figures 16, 17, and 18, respectively, with 5-bit quantization used on the quadrature components of the received OPSK symbols. These figures also give the AWGN results for coded OPSK, and uncoded QPSK performances are included for comparison.

The bandwidth limitations imposed by filtering and the nonlinearities of the satellite channel cause a performance loss (required increase in E_b/N_o to achieve a given P_b) relative to the AWGN case. The CCI and ACI cause further losses. At low E_b/N_o , the closer spacing of the eight-phase signal space causes greater losses for coded OPSK than for uncoded QPSK. Hence, coding gain at $P_b > 10^{-3}$ is less than for the AWGN case. At higher E_b/N_o values, however, the maximum likelihood path decoding of coded OPSK reduces these losses somewhat by averaging the interference perturbations over an interval greater than a constraint length of the FEC code. Thus, coded OPSK has smaller losses than uncoded QPSK at $P_b = 10^{-4}$ or lower; consequently, coded OPSK has a slightly higher coding gain for the satellite channel than for the AWGN channel when $P_b \leq 10^{-4}$.

In cases of interference-limited communications performance, the curve of P_b can exhibit flaring at high E_b/N_o because of an irreducible bit-error probability that results from interference alone. Such flaring was not observed in any of the computer simulation results for either uncoded QPSK or coded OPSK.

In laboratory hardware tests by Jones [27] of a 120-Mbit/s TDMA/QPSK modem, flaring of P_b was evidenced for the INTELSAT V channel with adjacent transponder interference (ATI) and CCI. The flaring was especially evident with CCI introduced on both the up- and down-links for three CCIs. The computer simulation program lumps all CCI into one equivalent up-link composite interferer. Also, the simulations were made with perfect synchro-

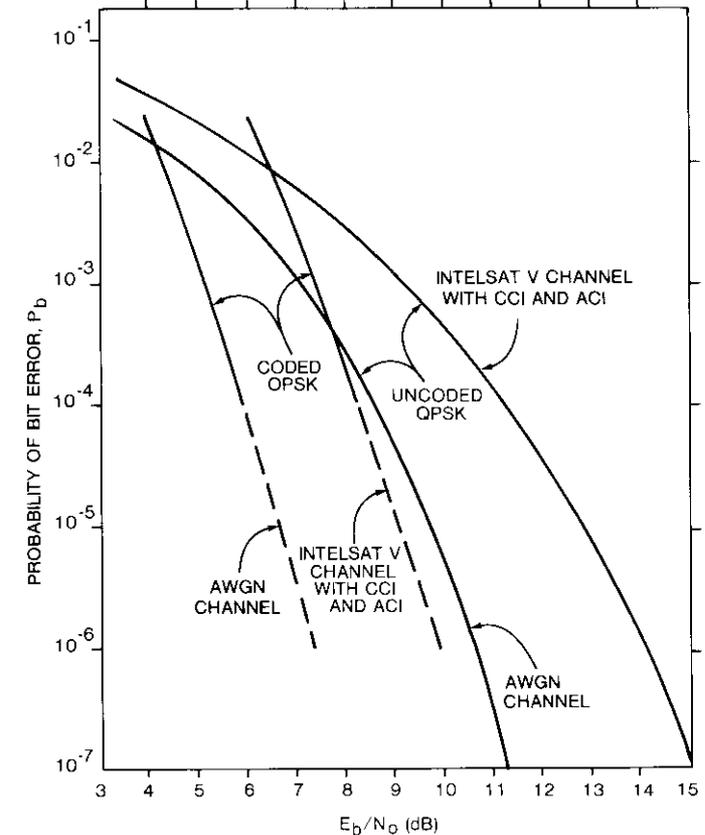


Figure 16. Coded OPSK Performance for Code 2 With $\gamma = 4$ Bits Memory and Asymptotic Gain of 4.1 dB

nization. Large numbers of interferences are necessary to produce levels of interference that are high enough to exceed the magnitude of the desired signal and thereby cause an irreducible error probability and P_b curve flaring. For the small number of low-level interferences in the computer simulation, the total interference can only cause a moderate reduction in the received power level of the desired transmission.

If flaring should indeed be present in the actual INTELSAT V system, the hardware simulations indicate an irreducible error probability for QPSK of about 10^{-6} . Down to $P_b = 10^{-4}$ there is no evidence of curve flaring. The effect of FEC decoding is invaluable whenever such flaring occurs, because

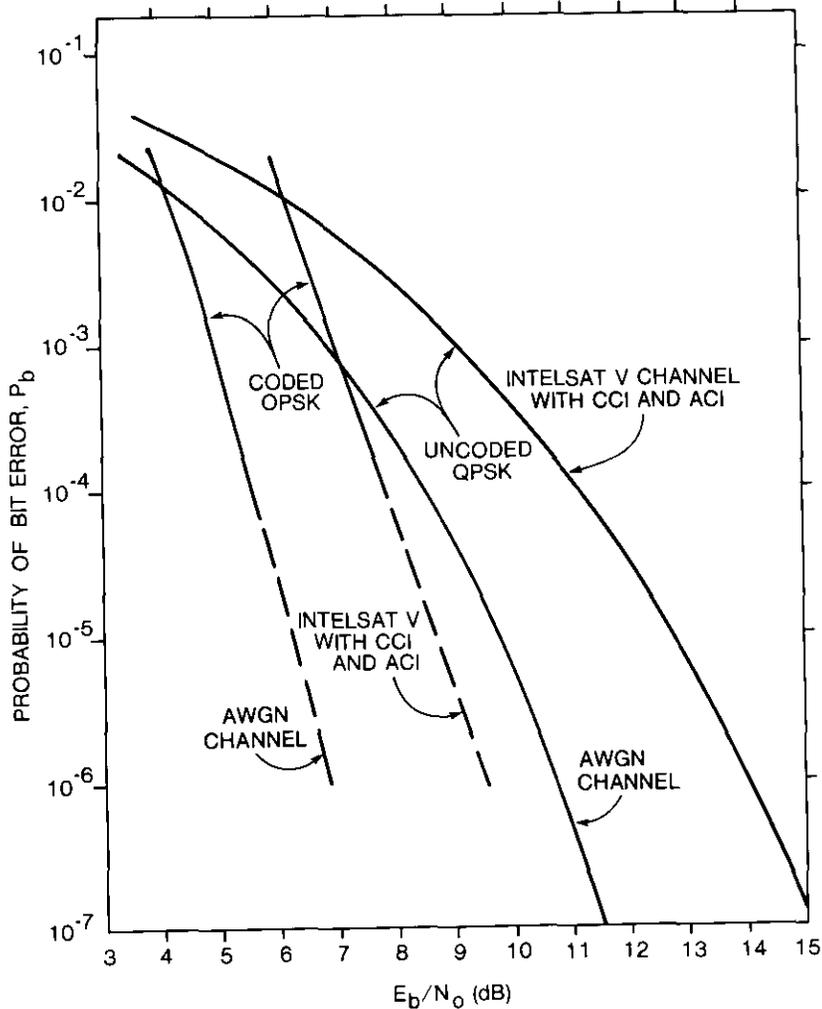


Figure 17. Coded OPSK Performance for Code 3 With $\gamma = 6$ Bits Memory and Asymptotic Gain of 4.6 dB

the FEC decoder can provide low P_b values at low E_b/N_0 , where noise rather than interference is the dominant source of signal disturbance. With Viterbi path decoding of coded OPSK, the irreducible value of P_b is expected to be very low.

For the INTELSAT V TDMA channel, $E_b = C/R_b$ is the definition used for

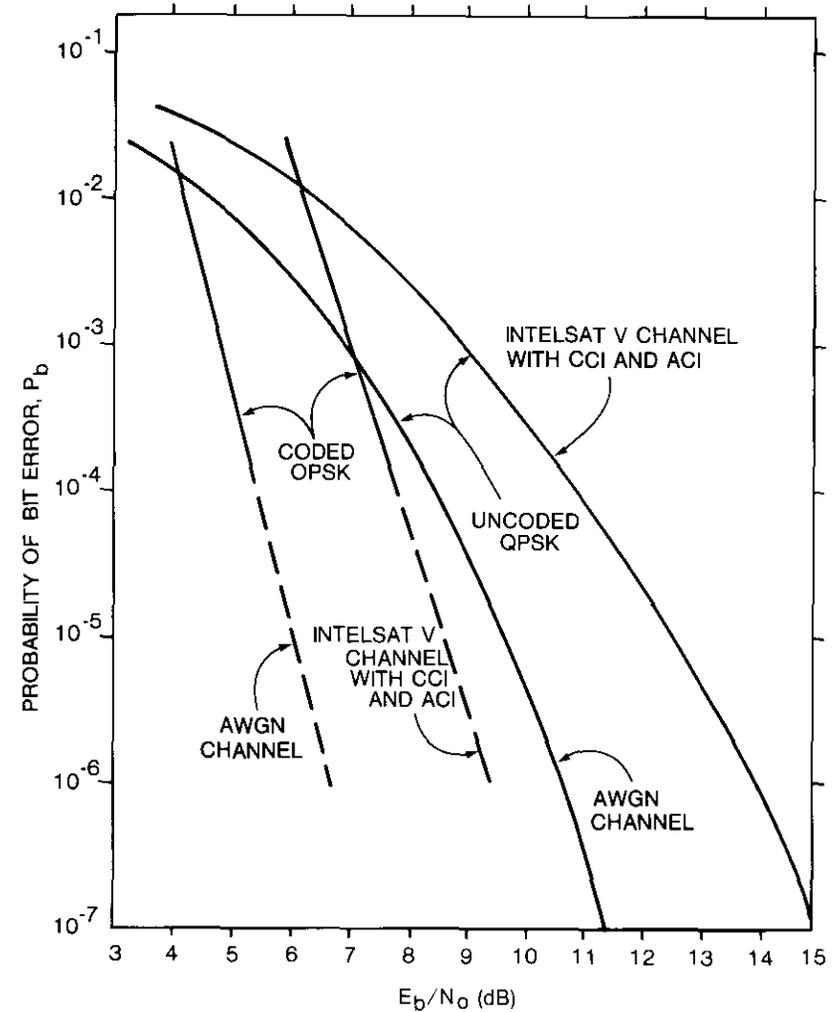


Figure 18. Coded OPSK Performance for Code 4 With $\gamma = 6$ Bits Memory and Asymptotic Gain of 5.0 dB

energy per information bit, where R_b is the information bit rate and C is the received level of unmodulated power. This definition is suitable for operation of the TWTA in the nonlinear region, where power is peak-limited. If the communications system were linear, E_b would be based on an average received power level, P , for the modulated signal. In general, peak power

is relevant to TDMA and other single-carrier-per-transponder communications, while average power is relevant to FDMA and other cases of multiple transmissions per transponder.

Table 5 summarizes the results for coded OPSK transmitted via the simulated INTELSAT V channel. Note the changes in coding gains relative to the AWGN cases. For instance, Codes 2, 3, and 4 yielded respective coding gains of 1.8, 1.9, and 2.0 dB at $P_b = 10^{-3}$ for the AWGN channel. For the simulated INTELSAT V satellite channel with CCI and ACI, these coding gains are changed only slightly to 1.7, 2.1, and 2.0 dB. At $P_b = 10^{-4}$, the respective coding gains of the three codes were 2.6, 2.9, and 3.1 dB for the AWGN case. However, these gains were increased to 2.9, 3.2, and 3.2 dB for the INTELSAT V case. At $P_b = 10^{-6}$, the estimated coding gains are much higher for the simulated INTELSAT V channel than for the AWGN case. However, these results are obtained by curve extrapolations and must be treated as rough approximations of decoding performances.

TABLE 5. CODED OPSK RESULTS FOR THE SIMULATED INTELSAT V AND AWGN CHANNELS

P_b	CODING GAIN* (dB)					
	AWGN CHANNEL			INTELSAT V CHANNEL		
	CODE 2	CODE 3	CODE 4	CODE 2	CODE 3	CODE 4
10^{-2}	0.2	0.2	0.2	-0.3	-0.2	-0.1
10^{-3}	1.8	1.9	2.0	1.7	2.1	2.0
10^{-4}	2.6	2.9	3.1	2.9	3.2	3.2
10^{-5}	3.0	3.4	3.6	3.6	3.9	4.0
10^{-6}	3.3	3.7	3.9	4.2	4.5	4.7

* Coding gains estimated from curve extrapolation when $P_b < 10^{-4}$.

As stated previously, Code 2 with $\gamma = 4$ bits memory is optimum for minimum free Euclidean distance. This code yields an asymptotic coding gain for coded OPSK of 4.1 dB relative to uncoded QPSK performance. It is clear that coded OPSK with only 16 code states can yield excellent coding gain.

Code 3 with memory $\gamma = 6$, or $S = 64$ code states, yields an asymptotic coding gain of 4.6 dB. As seen from Table 4, this code yielded better results than Code 2, which had $\gamma = 4$ and $S = 16$. Code 4 has $\gamma = 6$ bits memory and rate $\rho = 2/3$ and yields an asymptotic code gain of 5.0 dB. Results for coded OPSK with this code are shown in Table 5 to be slightly better than for Code 3.

Sensitivity to delay distortion

Computer simulations of coded OPSK signals via the INTELSAT V channel were made to determine performance sensitivity to delay distortion of the satellite multiplexing filters. Code 4 of Table 2 was employed in these simulations. This code, with $S = 64$ states, has an asymptotic coding gain of 5 dB for the AWGN channel. The sensitivity of uncoded QPSK performance was also determined under the same conditions.

At present, the nominal delay response for the input multiplexing filter does not result in significant performance losses. However, the INTELSAT V specification bounds on delay distortion will allow filter responses that could degrade the detection performance. Any excessive delay distortion of the output multiplexing filter can be compensated for at the receiving earth station. Also, the bit-error probability was found by computer simulation to be more sensitive to distortion of the input multiplexing filter. Consequently, the study of performance sensitivity to delay distortion was made only for the input filter.

Delay distortion increases the amount of ISI, thereby degrading detection performance. There are two opposing characteristics for coded OPSK with respect to this interference. First, because the eight signaling points for OPSK are more closely spaced than the four phase positions of QPSK, coded OPSK could be greatly affected by such an increase in ISI. Second, however, the path decoding for coded OPSK tends to average the ISI over an interval of at least one constraint length before a path decision is made. For uncoded communications, detection decisions must be made on a symbol-by-symbol basis. Thus, the performance loss caused by ISI can be smaller for coded OPSK than for uncoded QPSK.

Figure 12 illustrates the nominal amplitude characteristic of the input multiplexer filter for INTELSAT V. This amplitude response was employed along with different delay responses. The five delay responses used in simulating this filter are illustrated in Figure 19. These include the nominal delay response and four levels of parabolic delay distortion. Of the four curves (A, B, C, and D) of parabolic distortion, curve A depicts a zero parabolic slope rate, which corresponds to flat delay or no delay distortion. Thus, this ideal delay response is better than the nominal response. Curves B, C, and D represent parabolic delay distortions of 4, 8, and 12 ns, respectively, at 30 MHz from band center. Curve D, with the largest distortion, represents approximately the maximum parabolic delay distortion that would fall within the specification mask for the filter.

For uncoded QPSK, the effect of delay distortion may be readily tabulated.

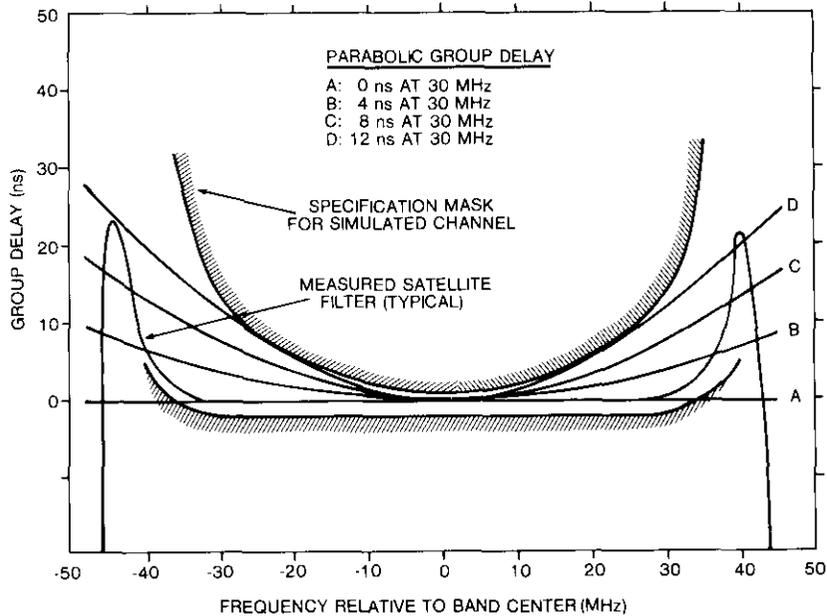


Figure 19. Group Delay for INTELSAT V Input Multiplexing Filter

At P_b values from 10^{-3} down to 10^{-7} , performance losses caused by delay distortion are given in Table 6 for the simulated INTELSAT V channel. Performance loss is measured in decibels and is defined as the required

TABLE 6. PERFORMANCE LOSS OF UNCODED QPSK CAUSED BY DELAY DISTORTION OF THE INPUT SATELLITE MULTIPLEXING FILTER

P_b	LOSS FOR DIFFERENT CASES OF DELAY DISTORTION* (dB)				
	A	N	B	C	D
10^{-3}	0.0	0.1	0.1	0.3	0.7
10^{-4}	0.0	0.18	0.15	0.4	1.1
10^{-5}	0.0	0.25	0.2	0.5	1.5
10^{-6}	0.0	0.3	0.25	0.6	1.9
10^{-7}	0.0	0.35	0.3	0.7	2.2

* Delay distortion:

- A—Ideal (zero)
- N—Nominal
- B—Parabolic with 4 ns at 30 MHz
- C—Parabolic with 8 ns at 30 MHz
- D—Parabolic with 12 ns at 30 MHz

increase in E_b/N_o to maintain the same P_b value. It is seen that the loss increases considerably for lower P_b values. Also, the loss for the case D of greatest parabolic distortion is much larger than for case C, the next greatest distortion.

In Figure 20, the performance of coded QPSK and uncoded QPSK are compared for cases A, N, and D. One peculiar characteristic may be deduced

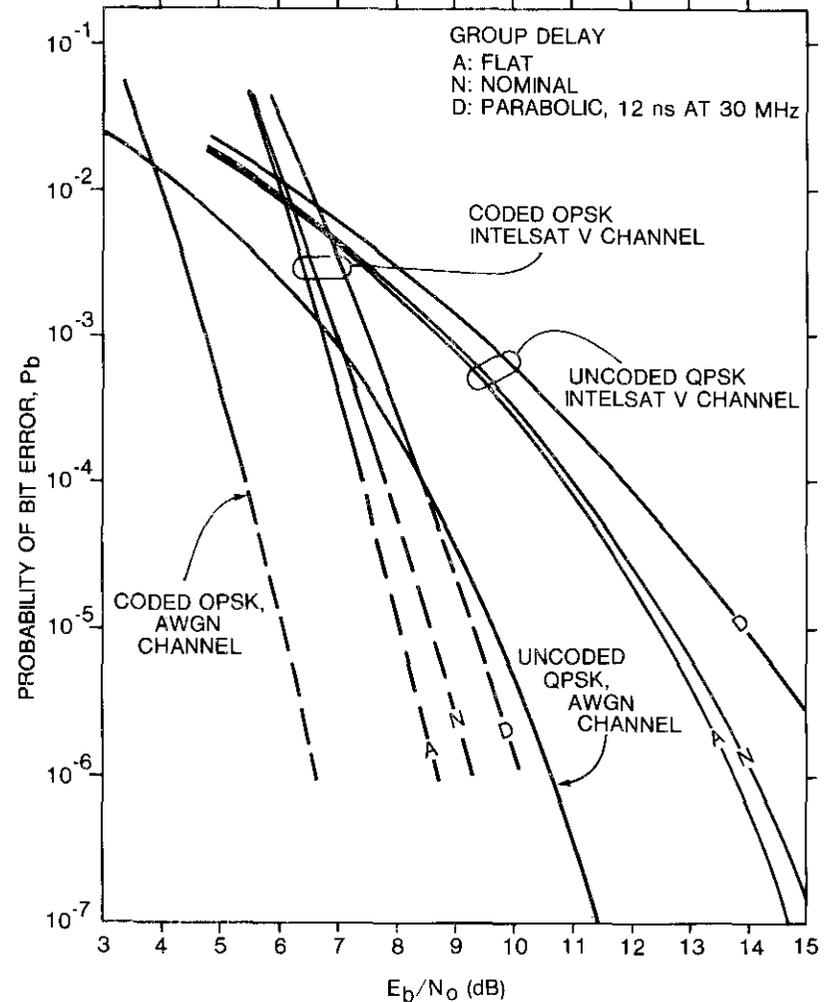


Figure 20. Sensitivity of Uncoded QPSK and Coded QPSK Performance to Parabolic Group Delay of the Input Multiplexer Filter

from these curves. The nominal delay distortion of the satellite input multiplexing filter causes more loss for coded OPSK than for uncoded QPSK. However, the additional loss caused by the large parabolic delay distortion of case *D* is less for coded OPSK than for uncoded QPSK. This latter comparison is relevant because all previous results have been for case *N*, the nominal group delay. Therefore, the performance loss relative to the nominal case indicates how a large parabolic delay distortion would degrade performance with respect to previously obtained results. In this comparison, coded OPSK performance is robust relative to that of uncoded QPSK.

Table 7 compares the losses for coded OPSK and uncoded QPSK as a result of the large parabolic distortion of Case *D*. These losses are given relative to the case *N* for nominal delay. Also, the improvement in performance that could be obtained with an ideal delay response is listed as a negative loss. This table indicates clearly the robust performance of coded OPSK for large delay distortion relative to the nominal performance. Also, the opposing characteristic is shown for coded OPSK in the nominal Case *N* relative to an ideal flat delay or zero distortion for Case *A*. The overall loss of Case *D* relative to Case *A* is only slightly less for coded OPSK than for uncoded QPSK.

TABLE 7. COMPARISONS OF PERFORMANCE LOSSES CAUSED BY DELAY DISTORTION FOR CODED OPSK AND UNCODED QPSK

P_b	PERFORMANCE LOSS RELATIVE TO NOMINAL DELAY CASE <i>N</i> *			
	CODED OPSK**		UNCODED QPSK	
	CASE <i>A</i>	CASE <i>D</i>	CASE <i>A</i>	CASE <i>D</i>
10^{-3}	-0.3	0.5	-0.1	0.6
10^{-4}	-0.4	0.6	-0.18	0.9
10^{-5}	-0.5	0.7	-0.25	1.25
10^{-6}	-0.6	0.8	-0.3	1.6

* Delay distortion:

N—Nominal

A—Ideal (zero)

D—Parabolic with 12-ns at 30 MHz

** Results for coded OPSK obtained by curve extrapolation when $P_b < 10^{-4}$.

Conclusions

Coding gain for coded OPSK relative to uncoded QPSK is quite impressive, while preserving the same bandwidth efficiency. Because maximum likelihood path decoding (Viterbi decoding) averages signal perturbations resulting from

various channel impairments, coded OPSK suffers smaller losses than uncoded communications over a satellite channel. Therefore, the coding gain for coded OPSK is somewhat larger for the simulated INTELSAT V case than for the AWGN case.

An FEC code memory of $\gamma = 4$ information bits, yielding $S = 16$ code states, was shown to provide a coding gain of about 2.6 dB for the AWGN case and 2.9 dB for the INTELSAT V channel at $P_b = 10^{-4}$. Use of a code with $\gamma = 6$, for $S = 64$ states, allows coded OPSK to obtain a 3.1-dB gain at $P_b = 10^{-4}$ for the AWGN channel and a 3.2-dB gain for the INTELSAT V channel. At $P_b = 10^{-6}$, the estimated coding gains for the AWGN case are 3.3 dB when $S = 16$ and 3.9 dB when $S = 64$. The corresponding gains for the INTELSAT V channel at $P_b = 10^{-6}$ are 4.2 dB for $S = 16$ and 4.7 dB for $S = 64$.

The advantage of using coded OPSK in a band-limited nonlinear satellite channel has been shown by computer simulation. At this juncture, it is desirable to demonstrate the feasibility of coded OPSK by actual hardware development at 120 Mbit/s for consideration in INTELSAT V system applications.

The recent CCITT Recommendation G.821 [28] on the Integrated Satellite Digital Network would tighten the requirements on the allowable bit-error rate (BER). In particular, the new recommendation of a BER of 10^{-6} or less for 90 percent of the time for end-to-end network performance would result in a requirement of $P_b \leq 10^{-6}$ for 99.2 percent of the time for the satellite link alone. Therefore, the large coding gain provided by coded OPSK at $P_b = 10^{-6}$ can be important in allowing INTELSAT TDMA communications to meet the new P_b recommendations, even with additional frequency reuse and the consequent increase in CCI levels for future satellite communications.

References

- [1] A. J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," *IEEE Transactions on Information Theory*, IT-13, April 1967, pp. 260-269.
- [2] J. A. Heller and I. M. Jacobs, "Viterbi: Decoding for Satellite and Space Communications," *IEEE Transactions on Communications*, COM-19, No. 5, Part II, October 1971, pp. 835-848.
- [3] M. G. Pelchat, R. C. Davis, and M. D. Luntz, "Coherent Demodulation of Continuous Phase Binary FSK Signals," International Telemetry Conference, Washington, D.C., 1971, *Proc.*, pp. 181-190.
- [4] S. Gronemeyer and A. McBride, "MSK and Offset QPSK Modulation," *IEEE Transactions on Communications*, COM-24, No. 8, 1976, pp. 809-820.

- [5] R. DeBuda, "Coherent Demodulation on Frequency-Shift Keying With Low Deviation Ratio," *IEEE Transactions on Communications*, COM-20, June 1972, pp. 429-435.
- [6] W. P. Osborne and M. B. Luntz, "Coherent and Noncoherent Detection of CPFSK," *IEEE Transactions on Communications*, COM-22, No. 8, August 1974, pp. 1023-1036.
- [7] T. A. Schonhoff, "Bandwidth vs Performance Considerations for CPFSK," IEEE National Telecommunications Conference, 1975, *Proc.*, pp. 38-1-38-5.
- [8] T. Aulin and C. E. Sundberg, "M-ary CPFSK Type of Signaling With Input Data Symbol Pulse Shaping—Minimum Distance and Spectrum," IEEE International Conference on Communications, Boston, June 10-14, 1979, *Proc.*, pp. 42.3.1-42.3.6.
- [9] J. B. Anderson and D. P. Taylor, "Bandwidth Efficient Modulation and Coding," Communications Research Laboratory of McMaster University, Hamilton, Ontario, Report No. CRL-55, 1978.
- [10] J. B. Anderson and D. P. Taylor, "A Bandwidth-Efficient Class of Signal Space Codes," *IEEE Transactions on Information Theory*, IT-24, No. 6, November 1978, pp. 703-712.
- [11] A. T. Lereim, "Spectral Properties of Multi-h Phase Codes," Communications Research Laboratory of McMaster University, Hamilton, Ontario, Report No. CRL-57, July 1978.
- [12] J. B. Anderson, D. P. Taylor, and A. Lereim, "A Class of Trellis Phase Modulation Code for Coding Without Bandwidth Expansion," IEEE International Conference on Communications, Toronto, Canada, June 4-7, 1978, *Conference Proceedings*, pp. 50.3.1-50.3.5.
- [13] T. Aulin, N. Rydbeck, and C. E. Sundberg, "Transmitter and Receiver Structures for M-ary Partial-Response FM," International Zurich Seminar on Digital Communications, March 4-6, 1980, *Proc.*, pp. A2.1-12.6.
- [14] T. Aulin, N. Rydbeck, and C. E. Sundberg, "Continuous Phase Modulation—Part II: Partial Response Signalling," *IEEE Transactions on Communications*, COM-29, No. 3, March 1981, pp. 210-225.
- [15] G. Ungerboeck and I. Csajka, "On Improving Data Link Performance by Increasing the Channel Alphabet and Introducing Sequence Coding," IEEE International Symposium on Information Theory, Ronneby, Sweden, June 21-24, 1976.
- [16] G. Ungerboeck, "Trellis Coding with Expanded Channel-Signal Sets," IEEE International Symposium on Information Theory, Cornell University, Ithaca, New York, October 10-14, 1977.
- [17] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," *IEEE Transactions on Information Theory*, IT-28, No. 1, January 1982, pp. 55-67.
- [18] J. Hui, "Joint Coding and Modulation Designs for Bandlimited Satellite Channels," Master's Degree Thesis, Massachusetts Institute of Technology, June 1981.
- [19] J. Hui and R. J. Fang, "Convolutional Code and Signal Waveform Design for Band-Limited Satellite Channels," IEEE International Conference on Communications, 1981.

- [20] R. J. Fang, "A Bandwidth and Power Efficient Modulation System," *Innovations in Telecommunications, Part A*, New York: Academic Press, 1982, pp. 29-57.
- [21] S. A. Rhodes and S. H. Lebowitz, "Performance of Coded QPSK for TDMA Satellite Communications," Fifth International Conference on Digital Satellite Communications, Genoa, Italy, March 23-26, 1981.
- [22] S. H. Lebowitz and S. A. Rhodes, "Performance of Coded 8-PSK Signaling for Satellite Communications," IEEE International Conference on Communications, Denver, Colorado, June 14-18, 1981.
- [23] S. G. Wilson, P. J. Schottler, and H. A. Sleeper, "Rate 3/4 16-PSK Phase Codes," IEEE International Conference on Communications, Philadelphia, June 13-17, 1982, *Conference Record*, Vol. 3, pp. 6F.1.1-6F.1.5.
- [24] Technical Staff of Bell Telephone Laboratories, *Transmissions Systems for Communications*, Fourth Edition, pp. 589-591.
- [25] E. Paaske, "Short Convolutional Codes with Maximal Free Distance for Rates 2/3 and 3/4," *IEEE Transactions on Information Theory*, IT-20, September 1974, pp. 683-687.
- [26] R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communications*, New York: McGraw-Hill Book Co., 1968, pp. 45-51.
- [27] M. Jones, "Laboratory Hardware Simulation Measurements of 120-Mbit/s QPSK/TDMA Transmission Performance in the INTELSAT V System," Sixth International Conference on Digital Satellite Communications, Phoenix, Arizona, September 1983.
- [28] International Telegraph and Telephone Consultative Committee (CCITT), "Report of the Meeting of Working Party XVIII/3: Network Performance Objectives," Study Group XVII, Contribution COM 18, No. R5, Geneva, July 1981.

Smith A. Rhodes received a B.S.E.E. from Virginia Polytechnic Institute in 1959 and an M.S.E.E. from North Carolina State University in 1962. From 1959 to 1965, he was a Member of the Technical Staff at Bell Laboratories, where he was engaged in various communications system studies. From 1965 to 1968, he was involved in digital communications studies for the R&D directorate of Page Communications. At Computer Sciences Corporation, he participated in digital satellite communications studies for military applications from 1968 to 1974. Since joining COMSAT in 1974, Mr.



Rhodes has been engaged in system studies for satellite communications. He is currently a Staff Scientist in the Communications Studies Department of the Transmission Systems Laboratory. His primary interests are in the digital communications areas of modulation, synchronization, and error control. Mr. Rhodes is a member of the IEEE groups on Communications and Information Theory.



Russell J. F. Fang received a B.S.E.E. in 1962 from National Taiwan University, Taipei, China, and an M.S.E.E. and Ph.D. from Stanford University in 1964 and 1968. From 1964 to 1966, he was a research engineer at Stanford Research Institute. Since joining COMSAT Laboratories in 1968, Dr. Fang has been a Member of the Technical Staff, Staff Scientist, and Department Manager, and is currently Director of the Transmission Systems Laboratory. He has authored or co-authored more than 20 papers in professional journals, not including conference papers. He is a Senior Member of IEEE.

Peter Y. Chang received a B.S.E.E. in 1966 from Cheng Kung University, Taiwan, China, and an M.S.E.E., M.Ph., and Ph.D. from Columbia University in 1969, 1972, and 1975, respectively. Prior to joining COMSAT, he was with the Satellite Communications Division of Computer Sciences Corporation, Falls Church, Virginia. He joined COMSAT Laboratories in 1976 and is currently a Staff Scientist in the Communications Studies Department of the Transmission Systems Laboratory. His work at COMSAT has included TDMA and FDMA modulation, digital signal transmission, intermodulation analysis involving the nonlinear effect of transponders, direct-broadcast (TV video and audio) satellite transmission systems, intersatellite links, simulation of satellite marine or mobile communications systems with fading channels, and the simulation of digital signal performance in satellite channels involving various coding schemes. Dr. Chang is a member of IEEE, Eta Kappa Nu, and Sigma Xi.



Index: telephone transmission, multiplexing, modulation, speech processing, frequency allocations

Impact of speech interpolation and data on FDM/FM transmission*

G. G. SZARVAS AND S. J. CAMPANELLA

(Manuscript received December 17, 1982)

Abstract

Speech interpolation decreases the bandwidth required for a given number of voice channels in an FDM/FM system. The bandwidth decrease is a function of the interpolation gain, g , which is the ratio of the number of channels available without interpolation to that with interpolation. This advantage is partially cancelled by increased multi-channel load and the corresponding increase in FM deviation required to restore the proper noise floor. For a given g and for a fixed FM bandwidth, the net result is an FDM/FM carrier voice frequency channel capacity enhancement between $g^{2/3}$ and $g^{5/6}$, depending on the number of channels. If the FM deviation is not modified, and the FM bandwidth is reduced by the same amount by which the FDM bandwidth is reduced, *i.e.*, by the factor g , channel performance degrades. Channels with data load decrease the interpolation efficiency and the obtainable carrier capacity enhancement. Assigning data-loaded channels, which have a lighter load than the speech-interpolated channels, to high-frequency FDM baseband positions minimizes the multichannel load deviating the FM carrier.

Introduction

In voice communications over four-wire circuits, channels are active less than 50 percent of the time on the average because of listening periods,

* This paper is an extended version of a paper presented at the Seventh Colloquium on Microwave Communications, Budapest, Hungary, September 1982.

inactivity between calls, and hesitations and silences punctuating speech. Speech interpolation systems use this low speech activity to reduce the transmission circuits required. Speech bursts on n_t voice channels (trunks) are interpolated and transmitted over n ($n < n_t$) time-shared channels interconnecting speech interpolation terminals. Interpolation gain is defined as

$$g = \frac{n_t}{n} \quad (1)$$

(The interpolation gain takes data and speech signals into account and is generally less, as will be shown, than the speech interpolation gain, g_s , which takes only speech signals into account.) Thus, interpolation leads to an increased average interpolated trunk load as compared to a noninterpolated trunk load.

FDM/FM carriers are usually designed for transmitting n telephone channels carrying noninterpolated speech. The increased load of interpolated speech would overdeviate the carrier, as was first pointed out by Campanella in 1975 [1] for $n > 240$ channels. This paper examines how the excess load may be accommodated for a given power and bandwidth of a regular FDM/FM carrier, either by degrading channel performance or by decreasing the number of transmitted channels, to sacrifice some of the speech interpolation gain but to retain the channel performance. The impact of data signals on the FDM/FM transmission is also taken into account. Finally, it is shown that for a mixed load of interpolated and noninterpolated speech, the allocation of channels with interpolated speech or generally with a heavier load to the low-frequency positions in the FDM baseband spectrum minimizes the multichannel load [2].

Carrier with n channels of interpolated speech

The well-known Carson's Rule for the FM bandwidth is

$$B = 2[f_r \log^{-1}(P_{eq}/20) + f_m] \quad (2)$$

where f_r = rms deviation of the 0-dBm0 test tone

P_{eq} = peak load of n channels (dBm0)

f_m = maximum modulating frequency ($f_m \simeq 4.2n$ kHz)

The peak load of n channels with noninterpolated speech load, P_{eq0} , is

given by the following equations based on the classical work of Holbrook and Dixon [3] and References 4 and 5:

$$P_{eq0} = P_{av0} + 10; \quad n \geq 240 \quad (3a)$$

$$P_{eq0} = P_{av0} + (24 - 6 \log n); \quad n < 240 \quad (3b)$$

with

$$P_{av0} = -15 + 10 \log n$$

The quantity P_{av0} is the long-term mean power (-15 dBm0/channel). All values are in decibels referenced to 1 mW (dBm0). Equation (3a) corresponds to the 10-dB crest factor (the level exceeded with a probability less than 0.2 percent) for a Gaussian noise model, which is considered valid for $n \geq 240$. Equation (3b) expresses an increased crest factor for the departure from the Gaussian noise model that occurs with fewer channels.

Substituting the value of P_{av0} into equations (3a) and (3b) yields

$$P_{eq0} = -5 + 10 \log n; \quad n \geq 240 \quad (4a)$$

$$P_{eq0} = 9 + 4 \log n; \quad n < 240 \quad (4b)$$

The curves representing equations (4a) and (4b) intersect at $n = 215$, as shown in Figure 1. However, by tradition, the value of n that determines which approximate expression to use is $n = 240$, which corresponds to the grouping of four supergroups. Figure 1 shows that doubling the channel numbers for $n < 215$ results in a peak load increase of 1.2 dB ($4 \log 2$) as compared to the 3.0-dB ($10 \log 2$) increase when $n > 215$. This difference is due to the differing multichannel signal crest factors previously described and results in different expressions for the effect of loading on FM bandwidth as a function of n .

Tradeoffs

The impact of speech interpolation on FDM/FM transmission is best shown by comparing extreme cases for the same FM bandwidth and carrier power, *i.e.*, all regularly loaded channels (reference case) vs all interpolated-speech-loaded channels.

Tradeoff between excess load and channel performance

When all n channels are loaded with interpolated speech with interpolation

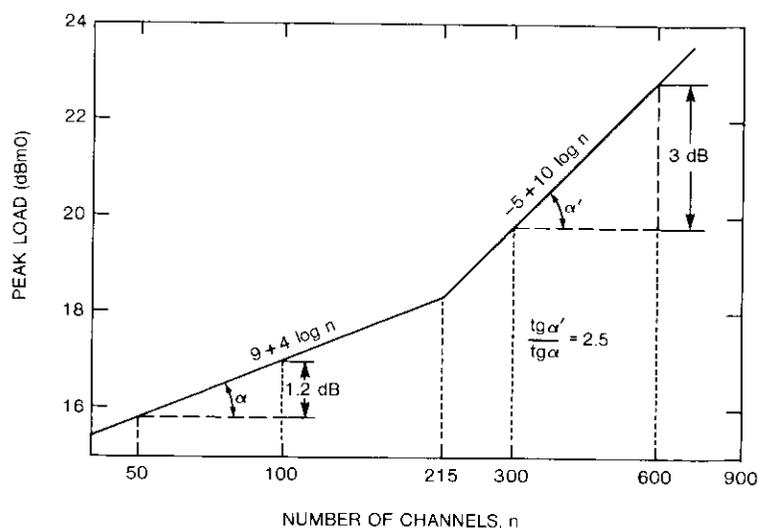


Figure 1. Multichannel Signal Peak Load

gain g_s , the speech signals of ng_s channels are transmitted on n channels. Using equations (4a) and (4b) with n for the noninterpolated speech load and ng_s for the interpolated-speech load yields the excess load due to interpolation, $\Delta P_{eq} = P_{eq} - P_{eq0}$, for various ranges of n and ng_s :

$$\Delta P_{eq} = 10 \log g_s \text{ (dB)} ; \quad n, ng_s \geq 240 \quad (5a)$$

$$\Delta P_{eq} = -14 + 6 \log n + 10 \log g_s \text{ (dB)} ; \quad n < 240, ng_s \geq 240 \quad (5b)$$

$$\Delta P_{eq} = 4 \log g_s \text{ (dB)} ; \quad n, ng_s < 240 \quad (5c)$$

Overdeviation is avoided by inserting an attenuator with a numerical value of ΔP_{eq} before the FM modulator. This step decreases the signal-to-noise performance of each channel by the same amount. The degradation at $g_s = 2$ corresponds to the excess load caused by doubling the number of channels, an example previously given. Table 1 shows the channel signal-to-noise degradation for three different values of n and several values of g_s . Figure 2 shows the channel performance degradation as a function of channel numbers with g_s as a parameter.

Tradeoff between excess channel load and number of channels

Channel performance in an FDM/FM system is usually expressed in terms

TABLE 1. CHANNEL SIGNAL-TO-NOISE DEGRADATION, ΔP_{eq} (dB) VS NUMBER OF CHANNELS (n) AND INTERPOLATION GAIN (g_s)

NUMBER OF CHANNELS n	INTERPOLATION GAIN, g_s			
	1.5	2.0	2.5	3.0
48	0.7	1.2	1.6	1.9
96	0.7	1.2	1.9	2.7
312	1.8	3.0	4.0	4.8

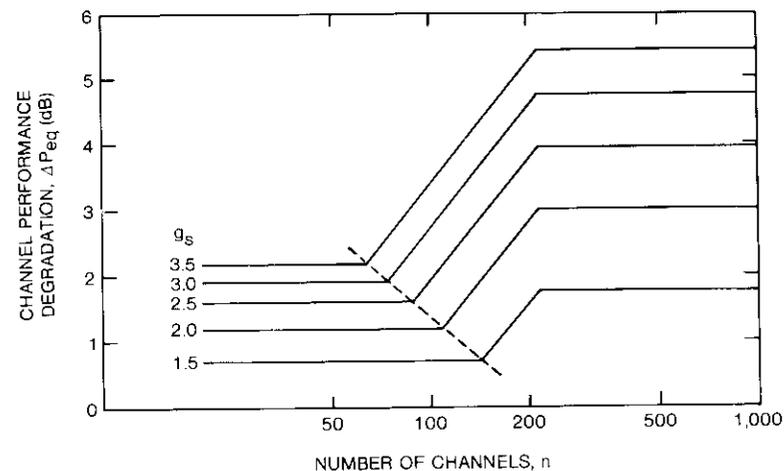


Figure 2. Channel Performance Degradation vs Channel Number

of the signal-to-noise ratio in the top channel in the baseband. The FM system signal-to-noise ratio for this channel is proportional to the square of the ratio of test tone deviation to maximum baseband frequency. For a given carrier power and bandwidth, the channel performance with and without interpolation is the same when

$$(f_r/f_{m0})^2 = (f_r/f_m)^2 \quad (6a)$$

or

$$f_r = (n/n_0) f_{r0} \quad (6b)$$

where the zero subscript refers to the noninterpolated system. The value of n , which must be adjusted between n_0 and n_0g_s to result in equal performance, can be calculated by substituting equation (6) into equation (2). The following expressions for n are obtained by neglecting the maximum modulating frequency term, f_m , in equation (2) because it is small compared to the maximum deviation:

$$n = n_0g_s^{-1/6}; \quad n_0, ng_s < 240 \quad (7a)$$

$$n = n_0g_s^{-1/3}; \quad n_0, ng_s \geq 240 \quad (7b)$$

$$n = 2.93n_0^{0.8}g_s^{-1/3}; \quad n_0 < 240, ng_s \geq 240 \quad (7c)$$

Each of the n channels carries ng_s voice frequency (VF) channels. Introducing $g' = ng_s/n_0$ as the capacity enhancement due to speech interpolation yields the following expressions:

$$g' = g_s^{5/6}; \quad n_0, ng_s \leq 240 \quad (8a)$$

$$g' = g_s^{2/3}; \quad n_0, ng_s \geq 240 \quad (8b)$$

$$g' = 2.93n_0^{-0.2}g_s^{2/3}; \quad n_0 < 240, ng_s \geq 240 \quad (8c)$$

Figure 3 shows g' as a function of n_0 , with g_s as a parameter. Equations (8a) to (8c), which agree respectively with D. S. Arnstein's result [6] and with Campanella's original result [1], show that a substantial part of the interpolation gain remains available when interpolation is applied to FDM/FM transmission. For example, for the conditions of equation (8b) with $g_s = 2$, 80 percent of the interpolation gain remains. This proportion increases to almost 90 percent for the conditions of equation (8a).

Influence of data load

Because continuous data signals cannot be interpolated, the interpolation gain decreases. The decrease of interpolation gain combined with the data level of -13 dBm0 [7], which differs from the average speech level of -15 dBm0, influences the FDM/FM transmission parameters.

Multichannel load

If a fraction β of the n noninterpolated speech channels carries data, the peak load is

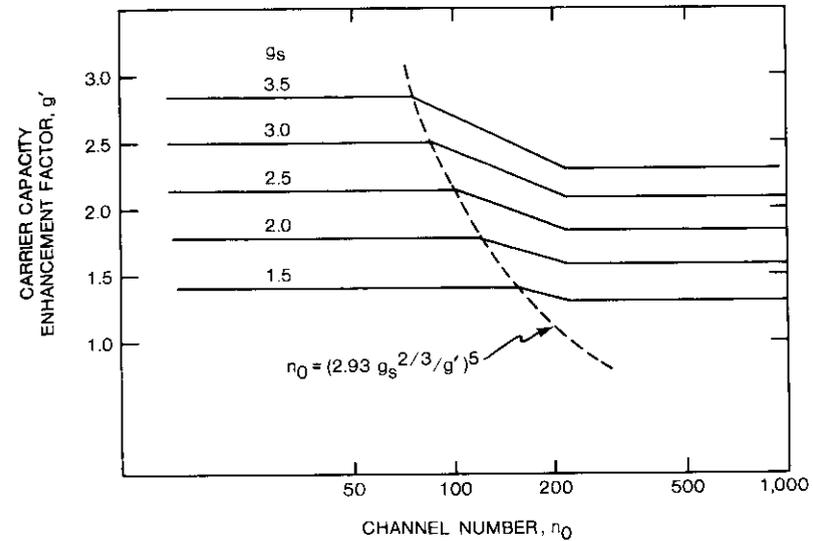


Figure 3. Carrier Capacity Enhancement vs Channel Number, With Interpolation Gain as a Parameter

$$P_{eq0} = 10 \log n [10^{-1.3\beta} + 10^{-1.5(1-\beta)}] + 10 \text{ (dBm0)}; \quad n \geq n_{pi} = f(\beta), 0 \leq \beta \leq 1 \quad (9a)$$

$$P_{eq0} = 9 + 4 \log n \text{ (dBm0)}; \quad n < n_{pi} = f(\beta), 0 \leq \beta \leq 1 \quad (9b)$$

where n_{pi} is the value of n where the curve representing equation (9a) intersects that representing equation (9b).

Figure 4 shows the curves for P_{eq0} in the ranges of n corresponding to equations (9a) and (9b) and also for speech load only [equations (4a) and (4b)].

Similarly to equation (4a), equation (9a) is based on the Gaussian distribution of multichannel signals in terms of the central limit theorem of statistics, which requires independent but not necessarily homogeneous individual channel signals. The average value of the multichannel signal increases from 0 to 2 dB as β increases from 0 to 1. For equation (9b), which is identical to equation (4b), it is assumed that the less prominent peaks of data signals, compared to those of speech signals, compensate for the increased average value. The channel number n_{pi} that determines the

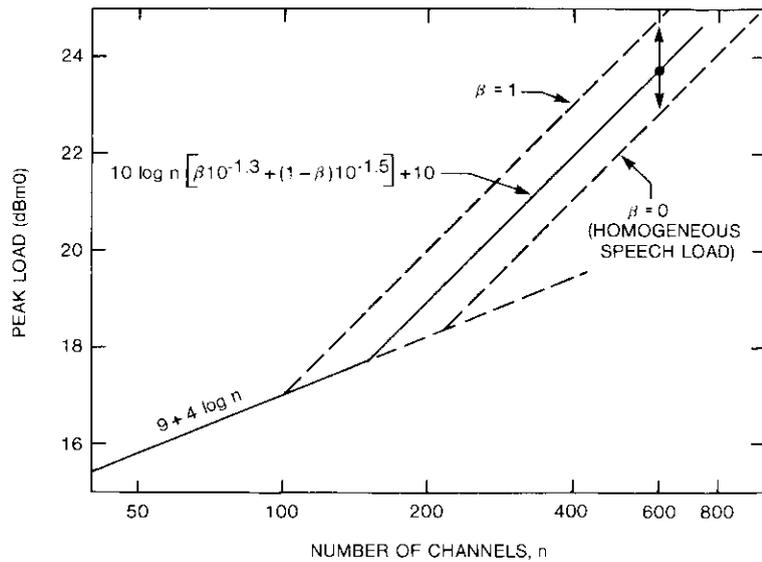


Figure 4. Multichannel Signal Peak Load vs Channel Number, With β as a Parameter

point of intersection is obtained by equating equations (9a) and (9b). Figure 5 shows this variable as a function of β . For homogeneous speech channels, equations (4a) and (4b) assume a speech activity of 0.25. Assuming continuous talkers, this means that for $240 \times 0.25 = 60$ channels the multichannel load has a Gaussian distribution. Figure 5 shows that when all channels carry data ($\beta = 1$), the Gaussian distribution starts at only $n = 100$. Therefore, the approximations used for the peak load of the mixed load of data and speech are conservative ones.

Decrease of circuit gain

The βgn trunks carrying data seize the same number of interpolated channels. Therefore, the speech signals of $(1 - \beta)gn$ trunks must be transmitted on the remaining $(n - n\beta g)$ channels, as shown by

$$(1 - \beta)gn = g_s(n - n\beta g) \tag{10}$$

which when solved for g yields

$$g = g_s / [1 + \beta(g_s - 1)] \tag{11}$$

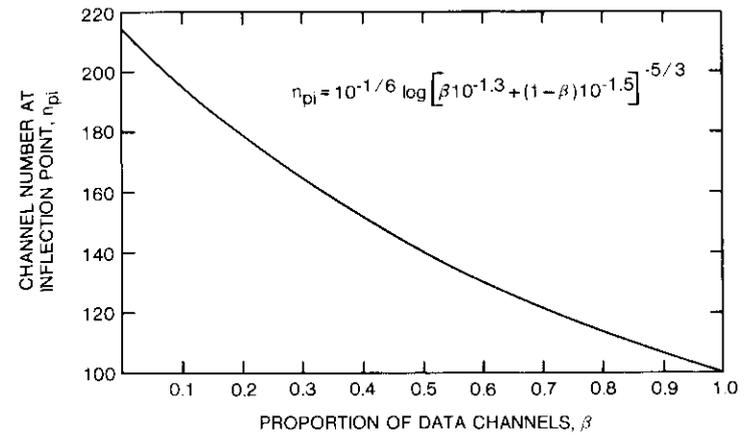


Figure 5. Channel Number at Point of Intersection

This expression shows that when β increases from 0 to 1, the interpolation gain g decreases from the speech interpolation gain of g_s to 1.

Tradeoffs for the mixed load

The tradeoffs for the mixed load are similar to those for speech load only. The mixed FDM/FM carrier load of $ng\beta$ data channels and $n(1 - g\beta)$ channels with interpolated speech load are referenced to the FDM/FM carrier mixed load of $n\beta$ data channels and $n(1 - \beta)$ channels with noninterpolated speech load.

Tradeoff between excess load and channel performance

Solving equations (9a) and (9b) for n channels and for gn channels and taking the difference results in equations (12a) and (12b) for the excess load tradeoff:

$$\Delta P_{eq} = 10 \log g \text{ (dB)} ; \quad n, gn \geq n_{pi}, n_{pi} = f(\beta) \tag{12a}$$

$$\Delta P_{eq} = 10 \log g [10^{-1.3}\beta + 10^{-1.5}(1 - \beta)] + 6 \log n + 1 \text{ (dB)} ; \quad n < n_{pi}, gn \geq n_{pi} \tag{12b}$$

$$\Delta P_{eq} = 4 \log g \text{ (dB)} ; \quad n, gn < n_{pi} \tag{12c}$$

Figure 6 shows the channel performance degradation at speech interpolation

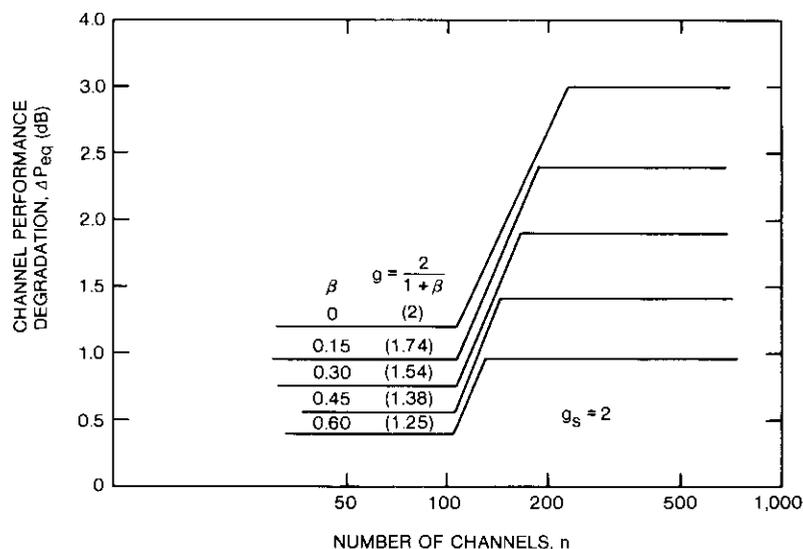


Figure 6. Channel Performance vs Channel Number at $g_s = 2$, With β as a Parameter

gain $g_s = 2$ as a function of n , with β as a parameter. The performance degradation is inversely proportional to β if $g_s > 1.6$, corresponding to the 2-dB difference between the data load of -13 dBm0 and the noninterpolated average speech load of -15 dBm0. If this condition is met, the performance for the mixed data and interpolated speech load is degraded less than for the homogeneous interpolated speech load.

Tradeoff between excess load and capacity enhancement

The carrier capacity enhancement for the mixed load is derived similarly as for the homogeneous load, with g given by equation (11), to obtain

$$g' = g^{5/6}; \quad n_0, ng < n_{pi} = f(\beta) \quad (13a)$$

$$g' = g^{2/3}; \quad n_0, ng \geq n_{pi} = f(\beta) \quad (13b)$$

$$g' = 0.93 n_0^{-0.2} g^{2/3} [\beta 10^{-1.3} + (1 - \beta) 10^{-1.5}]^{-1/3} \\ n_0 < n_{pi}; \quad ng > n_{pi} \quad (13c)$$

Table 2 shows the carrier capacity enhancement g' for a large and small number of channels [equations (13a) and (13b)] and the relative circuit gain g/g_s [equation (11)] as the function of β at $g_s = 2$. It can be seen that, when

β is increased, the carrier capacity enhancement decreases less than g , the interpolation gain. This relation holds because the data channels are less heavily loaded than the channels carrying interpolated speech.

TABLE 2. CARRIER CAPACITY ENHANCEMENT AND RELATIVE CIRCUIT GAIN VS β AT $g_s = 2$

β	0	0.15	0.3	0.45	0.6
g' (large n)	1	0.91	0.84	0.78	0.73
g' (small n)	1	0.89	0.80	0.73	0.68
g/g_s	1	0.87	0.77	0.69	0.63

Allocation of channels with different loads in the baseband spectrum

The multichannel peak load, P_{eq} , at the FM modulator input (after preemphasis) equals the baseband peak load, P'_{eq} , for a homogeneous channel load or for the case in which channels of different loads are uniformly distributed in the baseband spectrum, a tacit assumption for the mixed load in the preceding section, "Influence of Data Load." Generally, if the channels dedicated to different loads are not uniformly distributed, but are grouped, P_{eq} differs from P'_{eq} . This relation will be shown for the mixed load of channels with interpolated speech and noninterpolated speech. When interpolated speech is carried on n_i of the channels, and noninterpolated speech on n_r channels, the n mixed channels ($n = n_i + n_r$) carry the speech load of n' noninterpolated channels ($n' = n_r + g_s n_i$) referenced to the interpolator input. In the baseband, the peak load of n mixed channels is

$$P'_{eq} = P'_{av} + 10; \quad n' = n_r + g_s n_i \geq 240 \quad (14a)$$

$$P'_{eq} = P'_{av} + (24 - 6 \log n'); \quad n' < 240 \quad (14b)$$

where $P'_{av} = -15 + 10 \log n'$. The baseband long-term power P'_{av} consists of the power of n_r noninterpolated channels and the excess power of n_i interpolated channels. The two power components are shown in normalized form as areas A_1 and A_2 in the idealized baseband spectral density diagram of Figure 7a. The frequency is normalized by $x = f/f_m$; the spectral density is equal to unity for the n_r channels and equal to g for the n_i channels. The n_i channels are shown in the low-frequency part of the baseband occupying the normalized frequency interval $x < n_i/n$. Figure 7b shows the spectral density at the input of the frequency modulator. The shape of the curve is

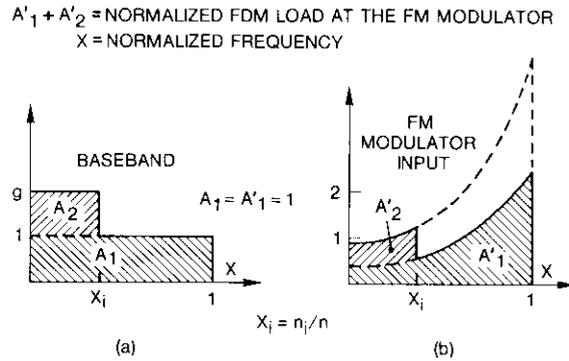


Figure 7. Spectral Densities

given by the regular FM preemphasis network transfer characteristic [8], which emphasizes channel signals occupying high-frequency positions in the baseband, where the noise is high because of the parabolic noise spectrum of the FM receiver. For regular FDM channel loading, the average preemphasis gain is adjusted to be unity over the frequency interval $0 \leq x \leq 1$; therefore, the areas are

$$A'_1 = A_1 = 1$$

It may be observed that A'_2 and the total normalized power at the frequency modulator, $A'_1 + A'_2$, has a minimum (maximum), by placing the n_i channels in the low (high) end of the baseband spectrum.

The ratio of the long-term power of the mixed channels after the preemphasis network to that before the preemphasis network is

$$\delta = \frac{A'_1 + A'_2}{A_1 + A_2} = \frac{1 + A'_2}{1 + (n_i/n)(g - 1)} \quad (15)$$

and in terms of δ , the peak load of n mixed channels is

$$P_{eq} = P'_{eq} + 10 \log \delta \text{ (dBm0)} \quad (16)$$

The minimum and maximum values of A'_2 , corresponding to grouping the interpolated channels at the low or high ends of the FDM spectrum, respectively, may be obtained by integrating a second-order polynomial approximating the normal preemphasis curve over the appropriate interval of x . Hence,

$$A'_{2\min} = (g_s - 1) \left[a_0 x_i + \frac{1}{2} a_1 x_i^2 + \frac{1}{3} a_2 x_i^3 \right] \quad (17a)$$

$$A'_{2\max} = (g_s - 1) \left\{ 1 - [a_0(1 - x_i) + \frac{1}{2} a_1(1 - x_i)^2 + \frac{1}{3} a_2(1 - x_i)^3] \right\} \quad (17b)$$

where $x_i = (n_i/n)$, $a_0 = 0.45$, $a_1 = -0.762$, and $a_2 = 2.802$. Substituting these minimum and maximum values of A'_2 into equation (15) yields corresponding minimum and maximum values of δ .

Figure 8 shows δ_{\max} and δ_{\min} in decibels as a function of the interpolated channels, n_i/n . For example, it can be seen that when 50 percent of channels loaded with interpolated speech of $g_s = 3$ are placed in the low-frequency baseband positions, the peak load at the FM modulator is about 1.3 dB less than in the baseband.

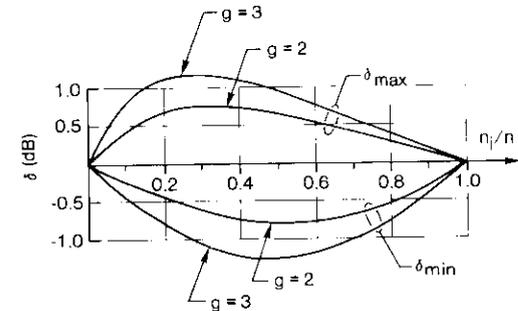


Figure 8. The Ratio δ vs Proportion of Channels Loaded With Interpolated Speech

Similarly, it is generally advantageous to assign the channels with interpolated speech, with data, and with noninterpolated speech in the given order to the baseband frequency spectrum.

Discussion

In regard to the use of TASI on FDMA/FM circuits, it is expected that in the long range, digital speech interpolation and digital transmission with TDM

multiplexing will replace analog speech interpolation and FDM/FM transmission. The transition is, however, a slow process, and there will be occasions when both transmission schemes are applied. For example, the speech interpolation system TASI-E [9] is designed for digital and analog transmission facilities by using analog-to-digital and digital-to-analog converters in the latter case.

Concerning the loading assumptions for voice and data used in this paper, it should be noted that CCITT Rec. H.51 [10] allows subsequent studies to specify "a reasonable proportion P " of channels to be loaded with data signals at -13 dBm0. This allowance may be explained by the measured decrease of the average speech level from the nominal -15 -dBm0 level, the base of multichannel load calculation. A decrease of the -13 -dBm0 data load to -15 dBm0 is also discussed, taking into account the development in data transmission. In this fluid state, it was reasonable for comparison and calculation to assume that the speech level is -15 dBm0 and that the data level is -13 dBm0 and to account for this difference in multichannel load calculation.

The use of preemphasis network characteristic for optimization has been used for other purposes as well. L. Schiff [11] suggested it for achieving equal signal-to-noise ratio for receiving stations with different down-link performance in multidestination FDM/FM satellite telephony. K. Jonnalagadda [12] used it for calculation of a mixed load consisting of companded channels and uncompanded channels.

Conclusion

For the transmission of interpolated VF channels by an FDM/FM carrier designed for transmitting noninterpolated channels, compromises in terms of channel performance and/or channel capacity are necessary. The presence of voice band data diminishes the circuit gain of interpolation and the carrier capacity enhancement, but the latter to a lesser degree. For the mixed-load case, the efficiency of the FDM/FM transmission is increased by placing the channels carrying the lesser load in increasing frequency positions in the baseband spectrum.

Acknowledgments

The authors wish to thank D. S. Arnstein for discussions and A. F. Ghais for valuable comments and critique. P. L. Bargellini provided extensive help in clarifying the text.

References

- [1] S. J. Campanella, Private communication.
- [2] G. G. Szarvas, "Impact of Speech Interpolation on FDM/FM Transmission," Seventh Colloquium on Microwave Communications, Budapest, September 6-10, 1982, *Proc.*, pp. 95-99.
- [3] B. D. Holbrook and J. T. Dixon, "Load Rating Theory for Multichannel Amplifiers," *Bell System Technical Journal*, Vol. 18, No. 4, October 1939, pp. 624-644.
- [4] CCITT, *Yellow Book*, Vol. III-2, Rec. G.223, Geneva, 1980.
- [5] CCIR, *Green Book*, Vol. IV, Report 708, Kyoto, 1978.
- [6] D. S. Arnstein, Private communication.
- [7] CCITT, *Yellow Book*, Vol. VIII-1, Rec. V.2, Geneva, 1980.
- [8] CCIR, *Green Book*, Vol. IV, Rec. 464, Kyoto, 1978.
- [9] R. L. Easton et al., "TASI-E Communications System," International Conference on Communications, Denver, Colorado, June 14-18, 1981, *ICC Conference Record*, pp. 49.3.1-49.3.5.
- [10] CCITT, *Yellow Book*, Vol. III-4, Rec. H.51, Geneva, 1980.
- [11] L. Schiff, "Compensatory Pre-emphasis," *RCA Review*, Vol. 37, December 1976, pp. 515-527.
- [12] K. Jonnalagadda, "Syllabic Companding and Voice Capacity of a Transponder," *RCA Review*, Vol. 41, September 1980, pp. 275-295.

George G. Szarvas received a Dipl. Eng. Degree in electrical and mechanical engineering from the Technical University in Budapest, Hungary. He completed post-graduate courses at the Royal Institute for Technology in Stockholm, Sweden. Before joining COMSAT in 1969, Mr. Szarvas worked for the Telecommunications Research Institute in Hungary, for L. M. Ericsson and Facit Electronics in Sweden, and for the Swedish Teleadministration. He is currently a Member of the Technical Staff in the Voice Band Processing Department of the Signal Processing Laboratory. He is a member of the IEE.





S. J. Campanella received a B.S.E.E. from the Catholic University of America in 1950, an M.S.E.E. from the University of Maryland in 1956, and a Ph.D. in electrical engineering from the Catholic University of America in 1965. As the Executive Director of the Communications Technology Division at COMSAT Laboratories, Dr. Campanella contributed significantly to techniques for TDMA network control and satellite-switched TDMA. He was Technical Director of the INTELSAT TDMA field trials in the Atlantic Ocean Region. Recently, he was named Chief Scientist at COMSAT Laboratories.

Dr. Campanella is a Member of AIAA, Sigma Xi, Phi Eta Sigma, and a Fellow of the IEEE and AAAS. He holds numerous patents in digital processing techniques.

Index: Intelsat, TDMA

Wideband digital transmission experiments in the Intelsat V system

D. CHAKRABORTY, J. EHRMANN, R. MINER, AND G. ROSCH

(Manuscript received June 6, 1983)

Abstract

A systematic experimental study was performed at the Andover, Maine, Earth Station to evaluate the performance of 120-Mbit/s QPSK transmission through a 77-MHz bandwidth INTELSAT V F2 channel. The test objectives included an assessment of amplitude and group delay equalizers for time division multiple access (TDMA). Tests were also conducted for electrical path delay, transversal equalizer sensitivity, and the effects of QPSK spectrum regrowth on FM carrier performance.

Equalization tests were successfully completed with both fixed and transversal equalizers. Transversal equalizers were used to optimize the amplitude and group delay characteristics of the station and satellite link for optimum BER performance. No difficulties were experienced in equalizing the terminal up- and down-chains to meet specifications. The optimally equalized channel yielded a BER of 10^{-6} at $E_b/N_0 \cong 12.9$ dB. Equalization was easily achieved with the transversal equalizer with different high-power amplifiers (HPAs). A 1-week test demonstrated stable link performance.

Extensive tests on the up- and down-path sensitivities to amplitude and group delay distortions revealed that the paths were sensitive to parabolic delay perturbations, particularly on the down-link. Test procedures for measuring electrical path length with a time interval counter were successfully demonstrated to within an estimated accuracy of 1 ns. Spectrum regrowth attributable to satellite transponder nonlinearity affected the adjacent FM channel by degrading the noise-power ratio (NPR) 0.5 to 1 dB. These tests provided valuable information about the practical constraints associated with wideband digital signaling.

Introduction

The INTELSAT TDMA/DSI specifications (BG-42-65) apply to network connectivity and operation during the INTELSAT V and post-INTELSAT V eras, with a 120-Mbit/s QPSK modulation format and transponders that have 72- to 77-MHz usable bandwidths with hemispheric or zonal beam coverage at 6/4 GHz.

Because sufficient data were not available to quantify 120-Mbit/s QPSK nonlinear satellite channel performance in an actual earth station environment, a test program was undertaken [1] to evaluate the transmission characteristics of 120-Mbit/s QPSK signals through a 77-MHz transponder (channel 1-2). The goals of this program were as follows:

- a. To determine an equalization strategy to characterize the earth station up- and down-chain and the transponder amplitude and group delay responses for optimum channel bit-error rate (BER) performance.
- b. To verify the stability of equalizer adjustments and link characteristics in terms of channel BER performance.
- c. To determine a technique and its accuracy for electrical path length measurement needed for TDMA burst time management.
- d. To conduct any other relevant tests that may be beneficial to the day-to-day operation of TDMA service through the INTELSAT V system.

To accomplish these goals, INTELSAT approved the use of INTELSAT V F2 (channel 1-2) transponder 41, a west-zone beam right-hand circularly polarized (RHCP) up and west-zone beam left-hand circularly polarized (LHCP) down configuration. The test program was carried out at the Andover earth station with the Andover II (AN-II) antenna facility.

For operational reasons, INTELSAT V channel 1-2 has been allocated for TDMA signaling (77-MHz transponder bandwidth and 5,970-MHz up-link carrier frequency). Because the center of the transponder band is at 5,967.5 MHz, this frequency difference may introduce asymmetric transmission characteristics, and hence distortions. However, since the TDMA transponder (channel 1-2) bandwidth (77-MHz) is sufficiently large compared with the Nyquist bandwidth of the transmitted signal (60 MHz), the asymmetry should not unduly affect the 120-Mbit/s QPSK signal.

Amplitude and group delay equalizations play an important role in wideband digital signaling. Fixed equalizers [2] in the form of lumped LC networks are normally used in the IF band to compensate for linear and parabolic distortion. The equalizers can be designed for either positive or negative amplitude slopes. The basic approach to equalizing delay distortion is to separate the linear and parabolic components and to design equalizers for each component.

Thus, several parabolic or linear sections can be cascaded to achieve large delay compensation values.

When several equalizing sections are cascaded, amplitude and delay ripples appear inside the transmission band. If a ripple is sinusoidal, the resulting data system performance can be estimated by the classical echo theory [3]. In general, ripples in both amplitude and phase or ripples with several frequencies result in interactive terms, which complicate the solution of the problem.

In practice, adjustable equalizers in the form of variable tapped delay lines [3] are used. The equalizer consists of a delay line tapped at intervals of t_0 seconds, where t_0 is the symbol period. Each tap is connected through a variable gain element ($|G| \leq 1$) to a summing bus, as indicated in Figure 1. The signal is available at the center tap unaltered in amplitude, while appropriate adjustments of the other tap settings introduce echoes of the signal. Thus, in the frequency domain, the tapped delay line is a finite Fourier series synthesizer that compensates for the deviation of the overall channel characteristic from the desired characteristic. The equalizer is quite effective in correcting all types of amplitude and phase deviations, particularly ripples which result in echoes that fall within $n/2$ symbol intervals of the main pulse, where n is the number of equalizing sections in the equalizer. Fixed and 3-tap transversal equalizers were used in the link under investigation.

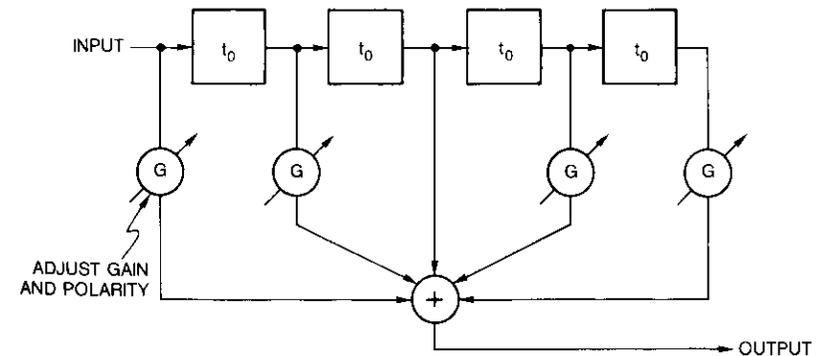


Figure 1. Tapped Delay Line Transversal Equalizer

Satellite link equalization strategy

Figure 2 is a simplified block diagram of a satellite link equalization scheme. Assuming that the HPA is operating in a linear region, the amplitude

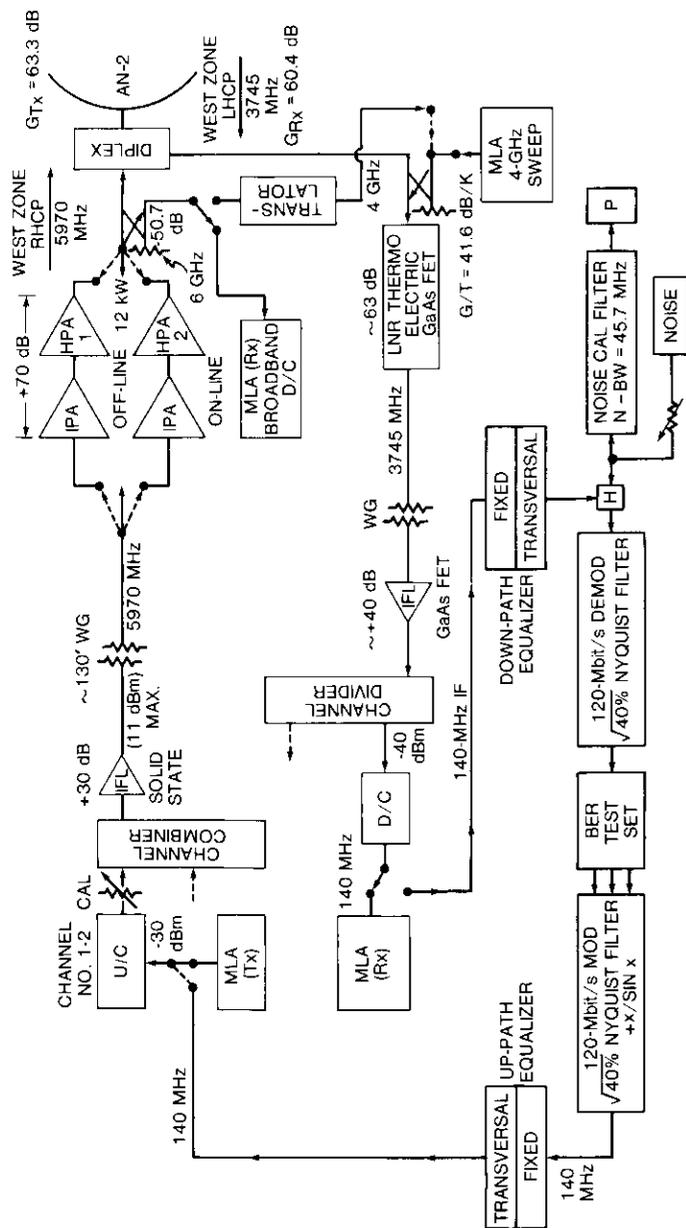
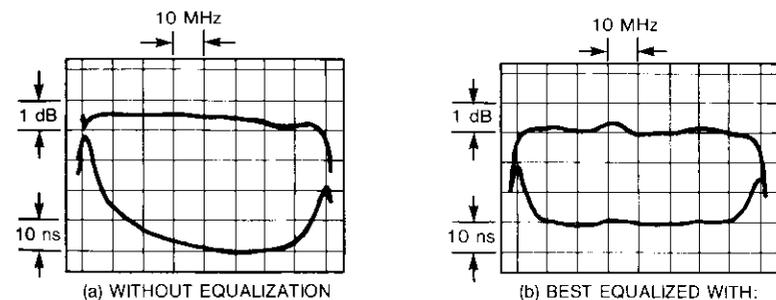


Figure 3. Simplified Block Diagram of Measurement Setup With the Andover II System



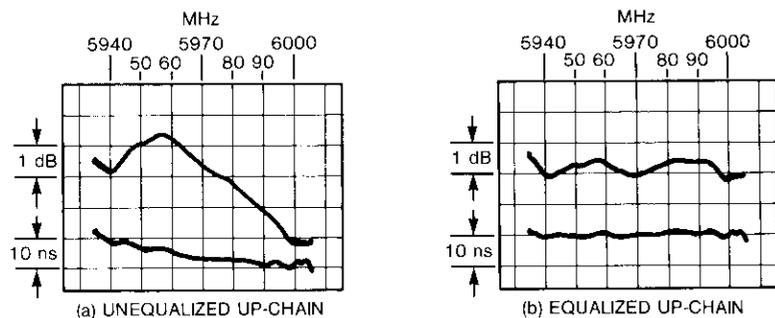
\cap 9-ns/ \pm 36-MHz PARABOLIC CONVEX OR CAP
 \cap +2.4-dB LINEAR AMPLITUDE/ \pm 36 MHz
 \cap 5-ns LINEAR GROUP DELAY OVER \pm 36 MHz
 \cap AND 3-TAP TRANSVERSAL EQUALIZER

Figure 4. AN-II Down-Chain Amplitude and Group Delay Characteristics

The up-path amplitude and group delay characteristics, including the 12-kW on-line HPA, were examined after the MLA signal was down-converted via the wideband down-converter. Figure 5 shows measured responses. Figure 5a represents the unequalized responses and Figure 5b the equalized responses. Again, fixed and transversal equalizers were used at IF and are identified on the right-hand side of Figure 5b. The HPA is the major contributor to the amplitude distortion in the up-chain. Figure 6 shows the INTELSAT TDMA/DSI

(Figure 3). Thus, the complete up-chain amplitude and group delay responses can be evaluated. In addition, a wideband 4-GHz microwave swept signal can be injected at the test port of the LNR (see Figure 3), and the down-chain (including the narrowband operational down-converter) amplitude and group delay responses can be examined by using the conventional MLA receiver. It should be noted that the wideband MLA down-converter with the receiver head is a specially designed unit and is transparent for amplitude and group delay variations.

The above procedure was used to measure the down-chain (from LNR input to down-converter output) amplitude and group delay characteristics, as shown in Figure 4. Figure 4a shows that the down-chain is subjected to asymmetric parabolic delay distortion, while the amplitude is substantially flat over the Nyquist bandwidth. The best equalized down-chain responses are shown in Figure 4b, where fixed equalizers comprising a 9-ns cap (convex parabolic), a 2.4-dB amplitude, and a 5-ns linear delay over \pm 36 MHz were used. A 3-tap transversal equalizer was also employed.



EQUALIZERS USED:

- + 3.5-dB / ± 36-MHz AMPLITUDE
- + 5-ns / ± 36-MHz LINEAR DELAY
- 6-ns \cap / ± 36-MHz PARABOLIC CONVEX OR CAP AND 3-TAP TRANSVERSAL EQUALIZER

Figure 5. Up-Chain Amplitude and Group Delay Characteristics With On-Line HPA 2

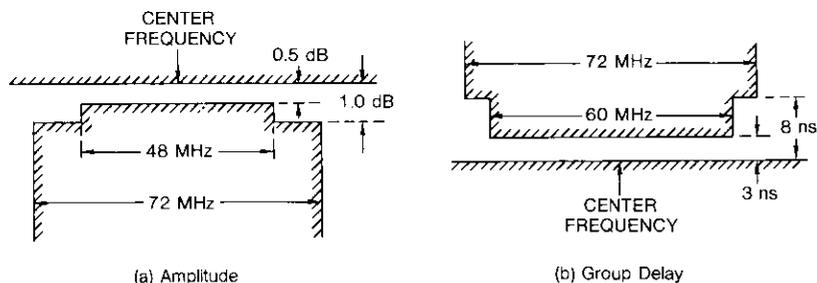


Figure 6. INTELSAT Specification (BG-42-65) for Up- or Down-Chain Amplitude and Group Delay Variation

specification masks (BG-42-65) for amplitude and group delay variation for the up- or down-chain. Figures 4b and 5b show that the INTELSAT specifications for up- and down-chain amplitude and group delay responses are in compliance. Thus, fixed and transversal equalizers can provide adequate flexibility in equalizing up- and down-chains in a typical earth station assigned for 120-Mbit/s QPSK TDMA operation.

Satellite loop measurements

SYSTEM CALIBRATION

Measurements through the satellite were performed by using the 12-kW

on-line HPA first and then switching the off-line 12-kW HPA into the system. Later, the on-line and off-line 3-kW HPA high-power multiplexer (HPM) were used to access the satellite.

The 12-kW on-line HPA was tuned to the manufacturer's specifications, and the earth station e.i.r.p. was calibrated in terms of the calibrated attenuator setting (see Figure 3). The HPA operation was linear for the maximum e.i.r.p. required for 120-Mbit/s QPSK TDMA transmission.

The transponder power transfer curve was generated based on the calibrated attenuator settings, and the measured curve is shown in Figure 7. The earth station e.i.r.p. was measured at the antenna feed point for transponder saturation, which was found to be 89.2 dBW.

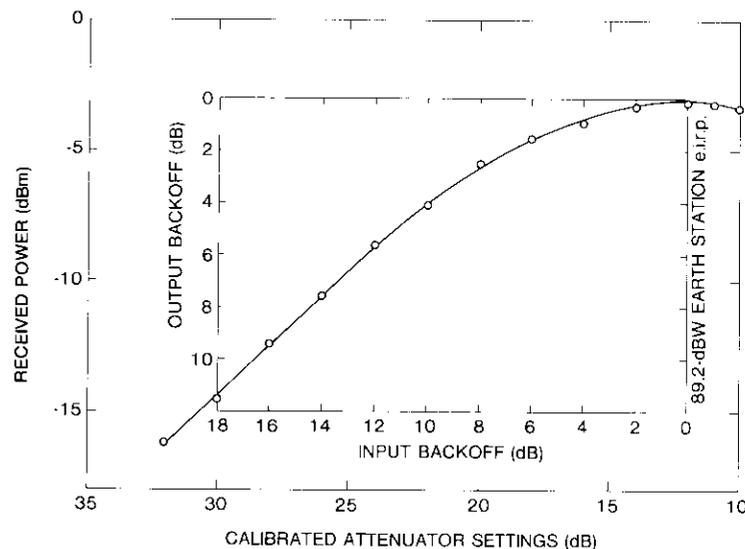


Figure 7. Transponder Power Transfer Curve

The measured cw carrier power and the noise power through a narrowband measuring filter were translated into the system C/N over the demodulator noise bandwidth of 66 MHz (receive noise bandwidth of $\sqrt{40\%}$ Nyquist filter). The corresponding E_b/N_o was $C/N - 10 \log(120/66)$. The measured C/N and E_b/N_o curves for the system under test are shown in Figure 8. The computed values of C/N are derived from the following system parameters:

Up-path loss at 5,970 MHz	200.0 dB
Up-path aspect (beam axis) correction	1.8 dB

Satellite beam axis G/T	-4.3 dB/K
Earth station e.i.r.p. at transponder saturation (measured)	89.2 dBW
Down-path loss at 3,745 MHz	195.9 dB
Down-path aspect (beam axis) correction	0.3 dB
Transponder e.i.r.p. at saturation	33.9 dBW
Receive G/T at Andover (LHCP receiver at 3,745 MHz)	41.6 dB/K

The computed C/N curve without up-path noise compression is shown in Figure 8. At or near saturation, the AM/AM characteristic of the TWTA compresses the input noise. From measured up-path noise compression characteristics of the transponder, the system C/N was recomputed, considering the up-path noise compression; the recomputed curve is shown in the figure. An examination of the three curves reveals that the measured C/N

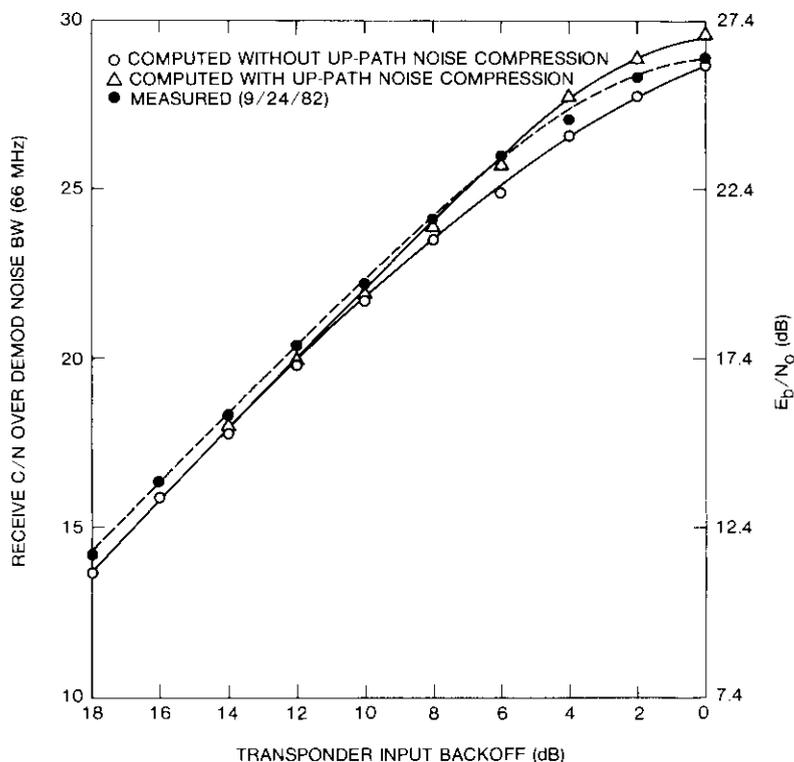


Figure 8. System C/N and E_b/N_0 vs Transponder Input Backoff

curve falls between the computed curves with and without noise compression in the operating range of the transponder input backoff (IBO); INTELSAT recommends 2-dB IBO.

LINK THROUGH 12-kw HPA

The satellite-loop amplitude and group delay characteristics are shown in Figure 9. Figure 9a shows the characteristics with the up- and down-chains equalized (refer to Figures 4b and 5b) and without transponder equalization. Channel 1-2 transponder equalizer intrinsic characteristics, shown in Figure 9b, are derived from satellite prelaunch data. These characteristics should essentially compensate for the unequalized output multiplexing filter responses. The transponder input multiplexing filter is well equalized and needs no additional equalization in the up-path.

With the transponder equalizer used in the system, further equalization in the up- and down-chains was necessary to achieve the best amplitude and

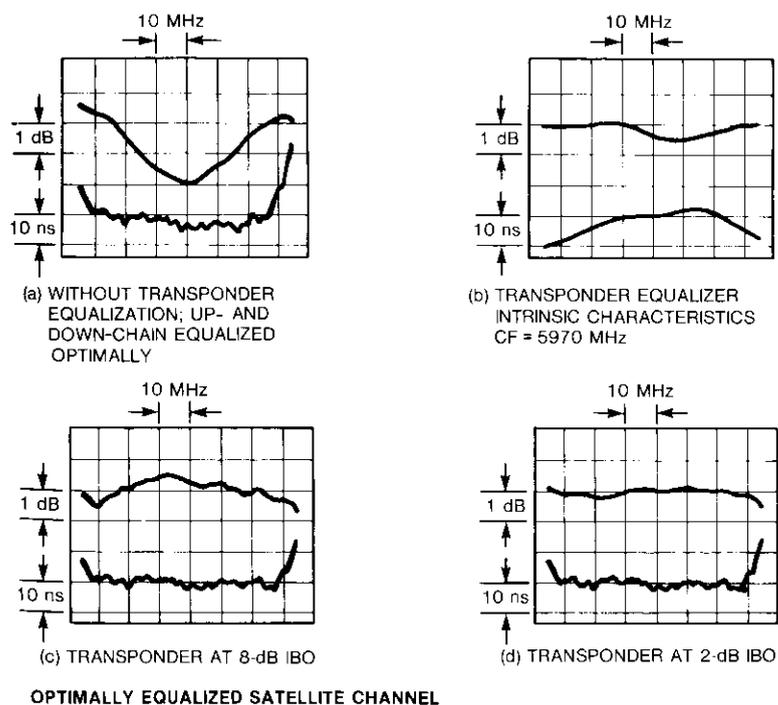


Figure 9. Overall System Amplitude and Delay Characteristics

group delay responses of the satellite loop. This was accomplished by adjusting the up-path transversal equalizer while keeping the fixed delay equalizers the same as for up-path equalization (compare with Figure 5b). The down-path equalization was readjusted. The best equalized satellite loop responses at 8-dB transponder IBO are shown in Figure 9c. These responses were repeated at 2-dB IBO of the transponder, as shown in Figure 9d. A comparison of parts c and d of Figure 9 shows the effect of amplitude compression. The satellite loop BER vs E_b/N_o (with noise added in the down-path) performance is shown in Figure 10, which compares loop performances with and without the transponder equalization to modem IF back-to-back

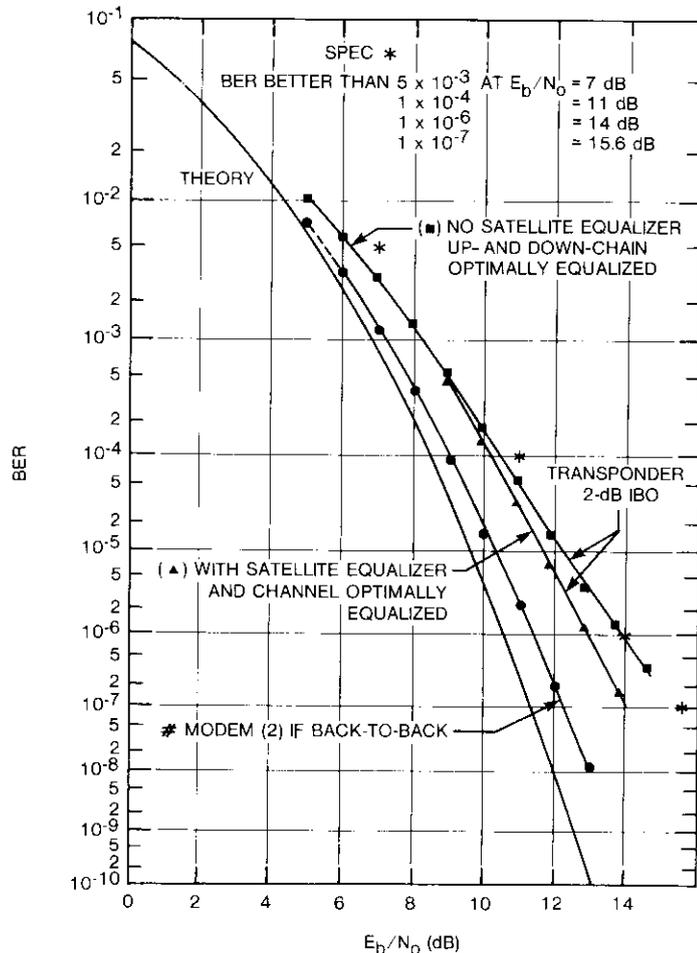


Figure 10. Satellite Loop Performance

performance. With transponder equalization, satellite loop performance improved by about 1 dB at BER 10^{-6} . The equalized loop performance is degraded only 1.5 dB from the IF back-to-back performance and 2.4 dB from theory at a BER of 10^{-6} and transponder IBO of 2 dB. This performance exceeds the INTELSAT BG-42-65 nonlinear channel modem BER vs E_b/N_o specification.

System C/N measurements taken under different weather conditions, using different HPAs, at different dates are shown in Figure 11. The satellite link BER was checked between experiments. Figure 12 gives two separate BER measurements performed over a period of 72 hr.

Next, the off-line 12-kW HPA was switched in and the link amplitude and group delay responses were measured at transponder IBOS of 8 and 2 dB with

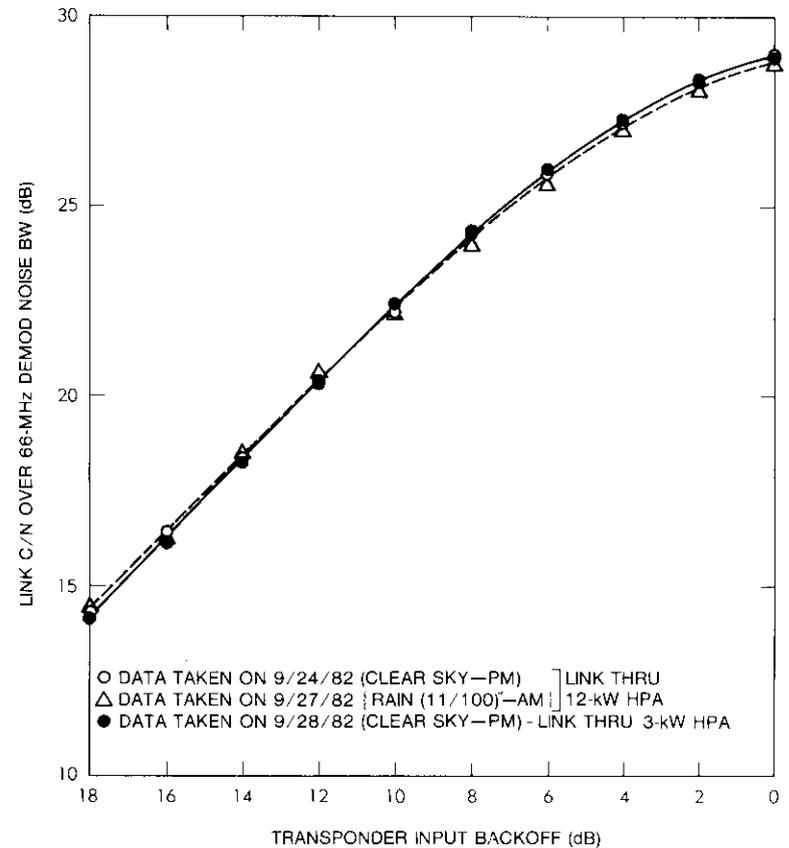


Figure 11. Link Stability Performance: C/N vs IBO

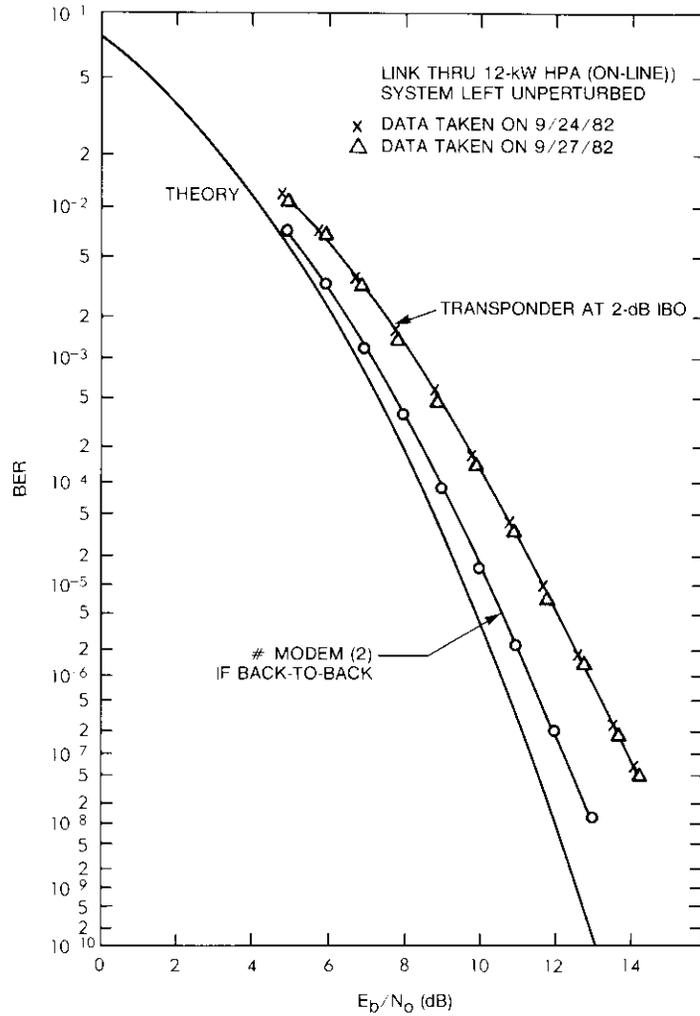


Figure 12. Link Stability Performance: BER vs E_b/N_0

the link equalization the same as in the on-line 12-kW system (Figure 9, parts c and d). The responses are shown in Figure 13, parts a and b, respectively. The effect of amplitude compression is obvious at 2-dB IBO. The significant difference in on-line and off-line HPA tube characteristics can

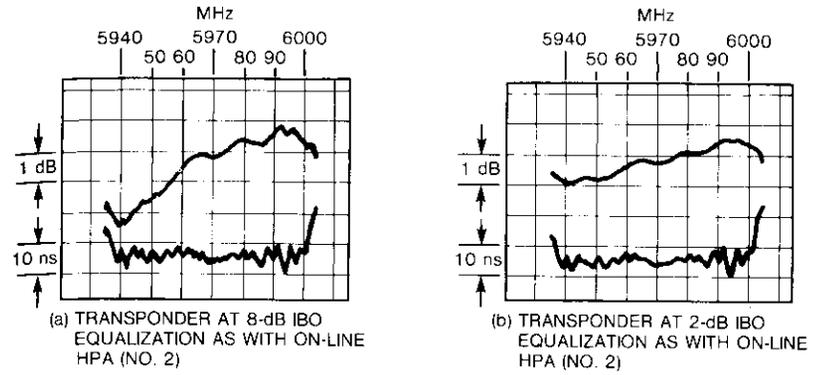


Figure 13. Overall Link Amplitude and Group Delay Characteristics With Off-Line 12-kW HPA 1

be observed by comparing the gain-frequency* responses, as shown in Figures 9a and 13a. The link characteristics with the 12-kW off-line HPA system were reequalized by adjusting the up-path transversal equalizer. The reequalized link responses taken at 8- and 2-dB transponder IBO are shown in Figure 14, parts a and b, respectively. The off-line 12-kW HPA link BER was

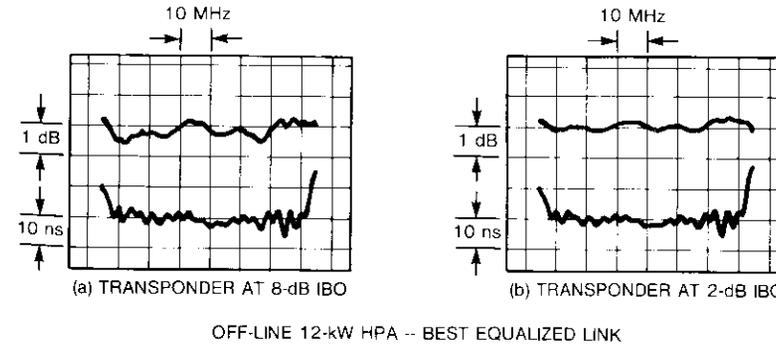


Figure 14. Off-Line HPA and the Satellite Link After Re-equalization With Transversal Equalizers, Fixed Equalizers as in On-Line HPA System

* The gain-frequency response of 12-kW coupled-cavity type TWTAs is somewhat sensitive to operating beam voltage and beam current. At peak efficiency, a typical gain variation over 500 MHz is about 6 dB.

measured first, with the same equalization as in the on-line HPA system. The measured data at 2-dB transponder IBO under two different equalization conditions are shown in Figure 15. A significant difference in BER was observed through the off-line HPA link with on-line HPA system equalization.

With the off-line 12-kW HPA link reequalized, the BER performance at 2-dB transponder IBO improved considerably and is comparable with the on-

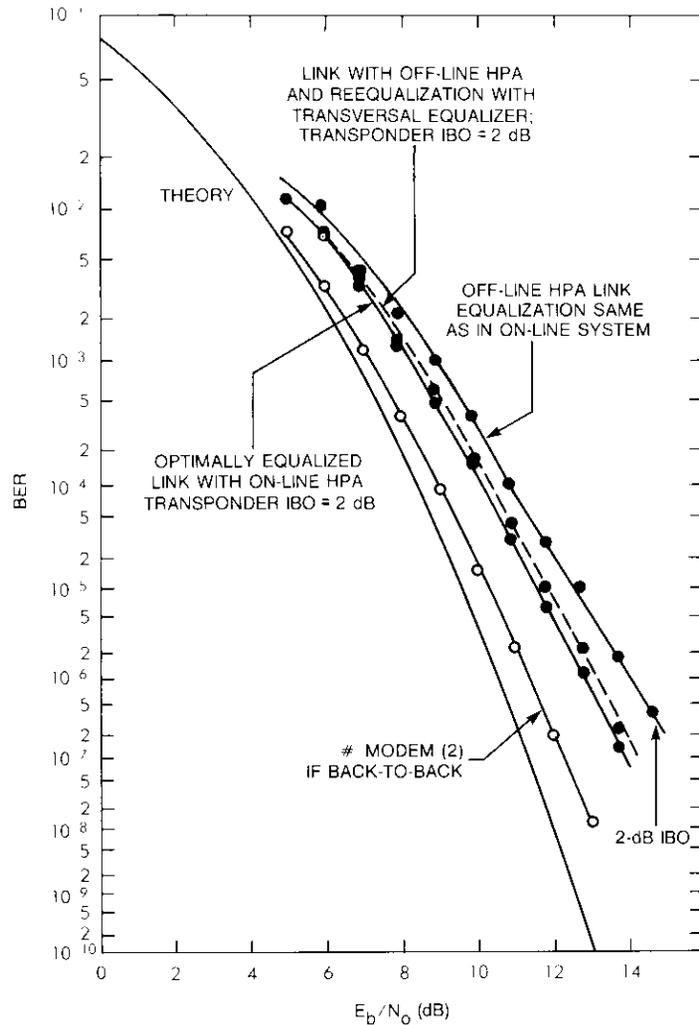


Figure 15. Link Performance With 12-kW Off-Line HPA Switched On Line

line 12-kW HPA link, as shown by the two adjacent curves of Figure 15.

To examine link performance repeatability, the on-line 12-kW HPA system was switched in, and after reequalization of the link with the transversal equalizers, the BER was measured and compared with the previous data. This comparison is shown in Figure 16.

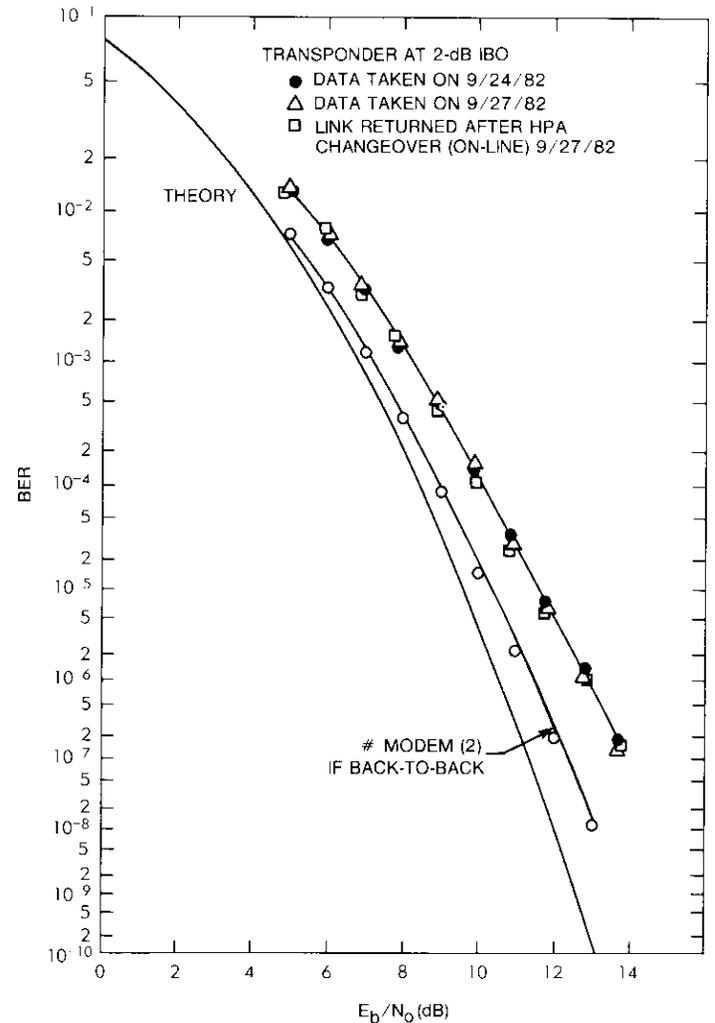


Figure 16. Link Performance Repeatability: AN-II On-Line 12-kW HPA Reequalized With Transversal Equalizers (Fixed equalizers in situ)

adjacent (ch 3) FDM/FM carrier top-channel NPR ranges from 0.5 to 1.0 dB because of the spectral regrowth of the PSK carrier (ch 1-2) when operated at a nominal IBO of 2 dB. This degradation can be easily corrected by a minor adjustment in the FDM/FM carrier e.i.r.p., if required.

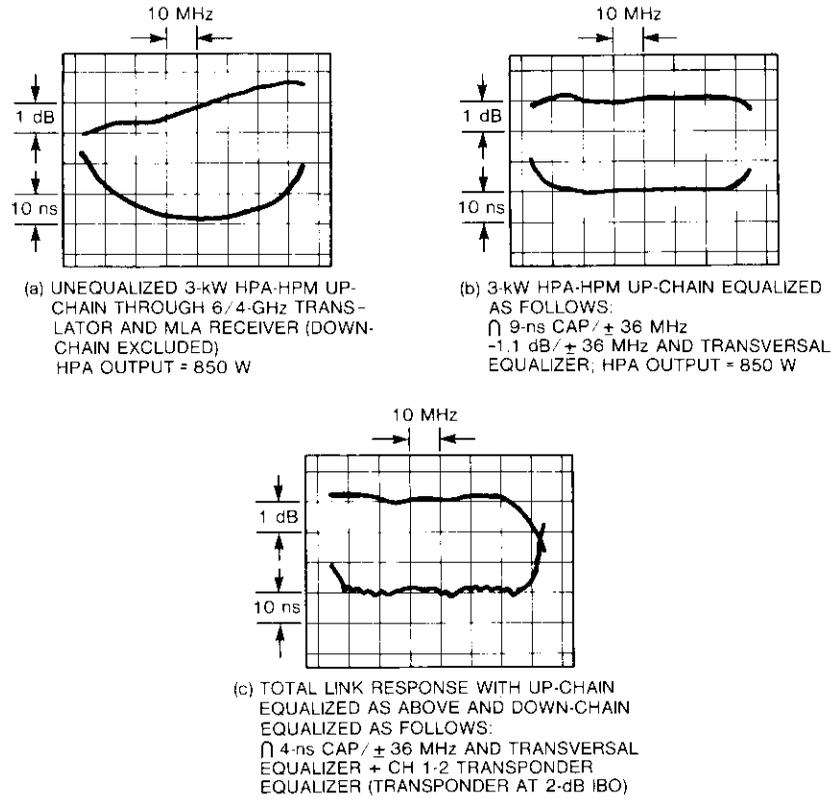


Figure 18. Link Amplitude and Delay Characteristics With 3-kW On-Line HPA and HPM

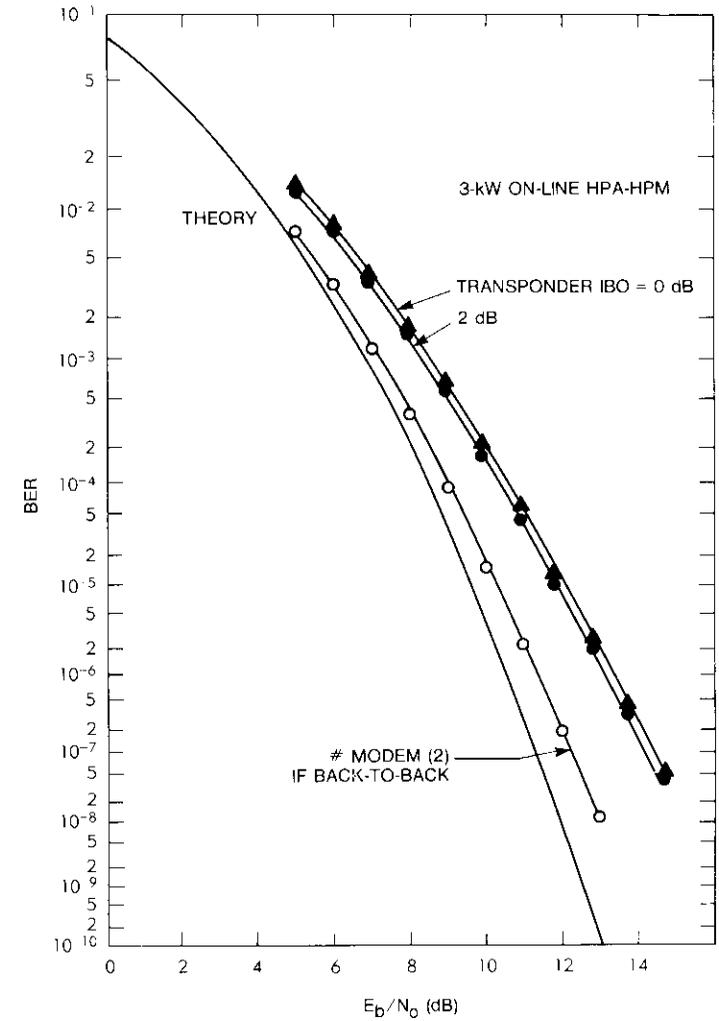


Figure 19. Link Performance With 3-kW On-Line HPA-HPM

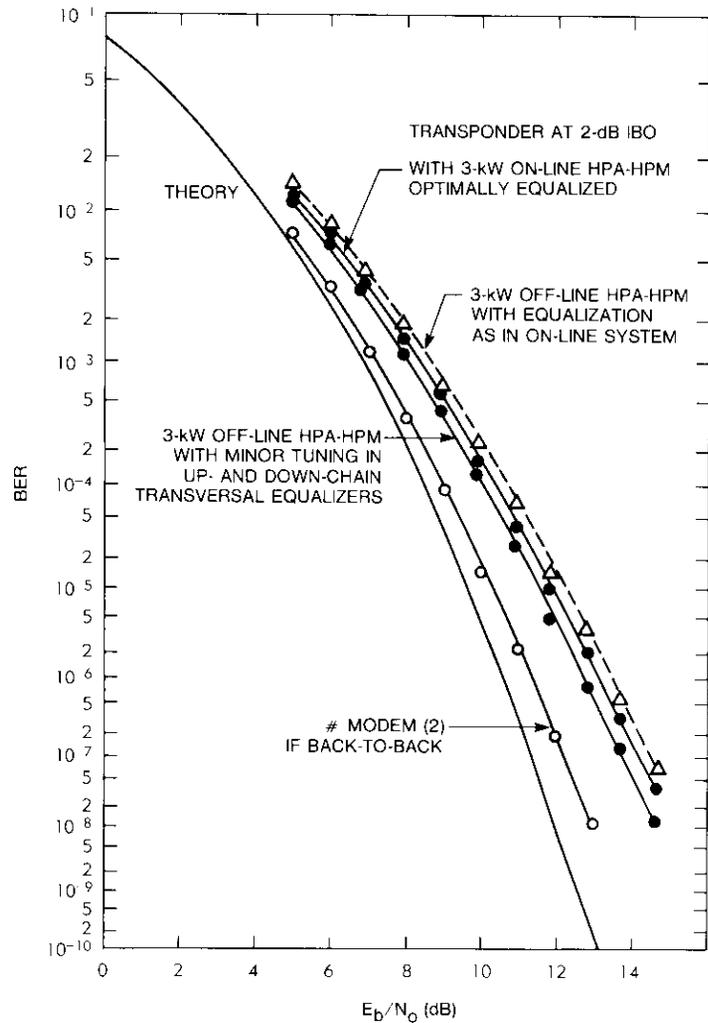


Figure 20. Link Performance With 3-kW Off-Line HPA-HPM

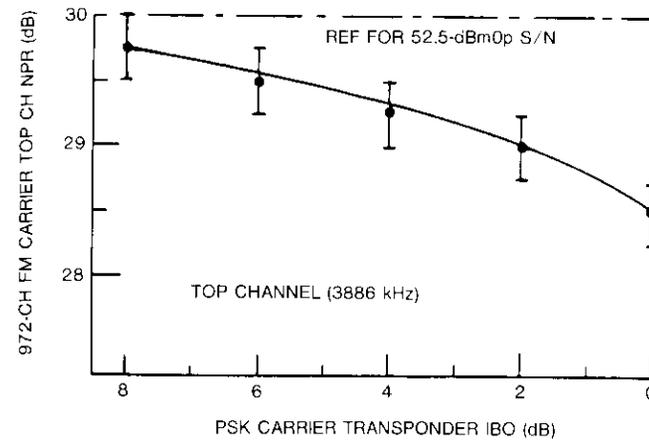


Figure 21. Adjacent Channel Interference Due to PSK Carrier Spectral Regrowth

TABLE 1. TRANSMISSION PARAMETERS

PARAMETERS	120-Mbit/s QPSK CARRIER	972-Ch FDM/FM CARRIER
Transponder Channel	1-2	3
Center Frequency (MHz)	5,970 Tx 3,745 Rx	6,030 Tx 3,805 Rx
Bandwidth (MHz)	60 (Nyquist)	36 (IFBW)
Top Baseband (kHz)	—	4028
Test Tone Frequency (kHz)	—	2,449.02
Test Tone rms Deviation (kHz)	—	802

CHANNEL EQUALIZATION SENSITIVITY STUDY

So far, the channel equalization has been studied by visual observation of amplitude and group delay responses of the looped-back link with the MLA. In an N -access TDMA configuration, $N - 1$ up-chains will not necessarily be equalized to match the N th down-chain equalization. This will result in an equalization mismatch in the overall TDMA network. Reciprocally, $N - 1$ down-chain equalization may not exactly match the N th up-chain equalization. This situation was examined by adding a small variation in up-path and down-path (IF) amplitude and delay characteristics of the best equalized

channel (by visual observation of MLA responses) with the on-line 3-kW HPA-HPM system and the transponder operating at a fixed IBO of 2 dB.

The BER performance variations for an amplitude slope perturbation of ± 1.2 dB over ± 36 MHz in up-path and down-path are shown in Figures 22 and 23, respectively. Positive amplitude slope ($+ 1.2$ dB/ ± 36 MHz)

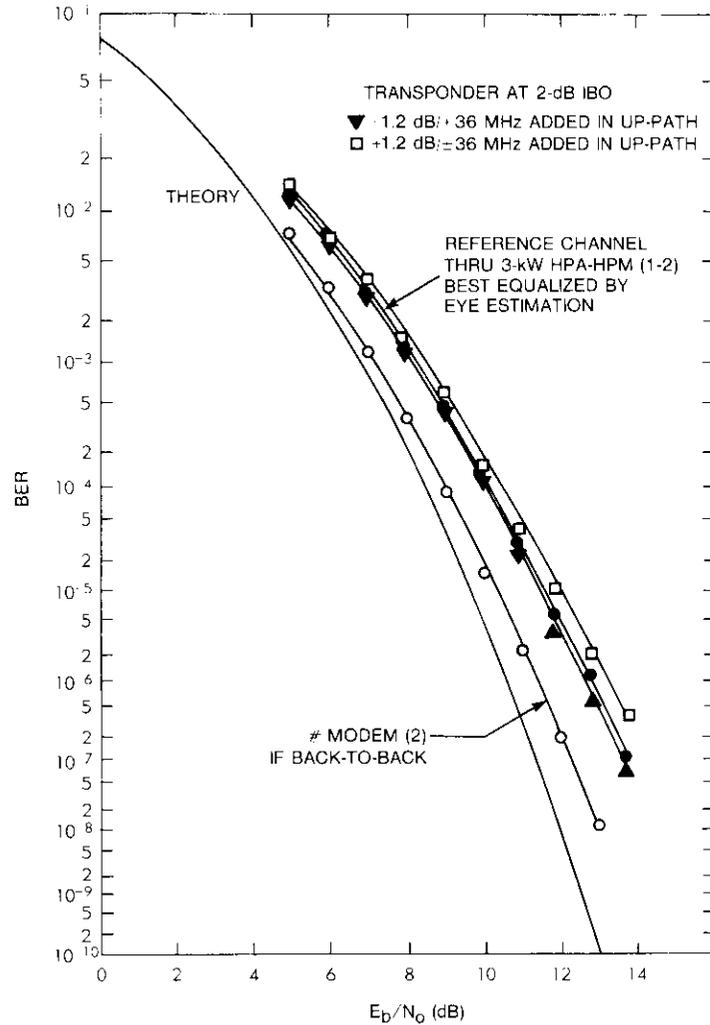


Figure 22. Channel Sensitivity—Amplitude Distortion in Up-Path as Parameter

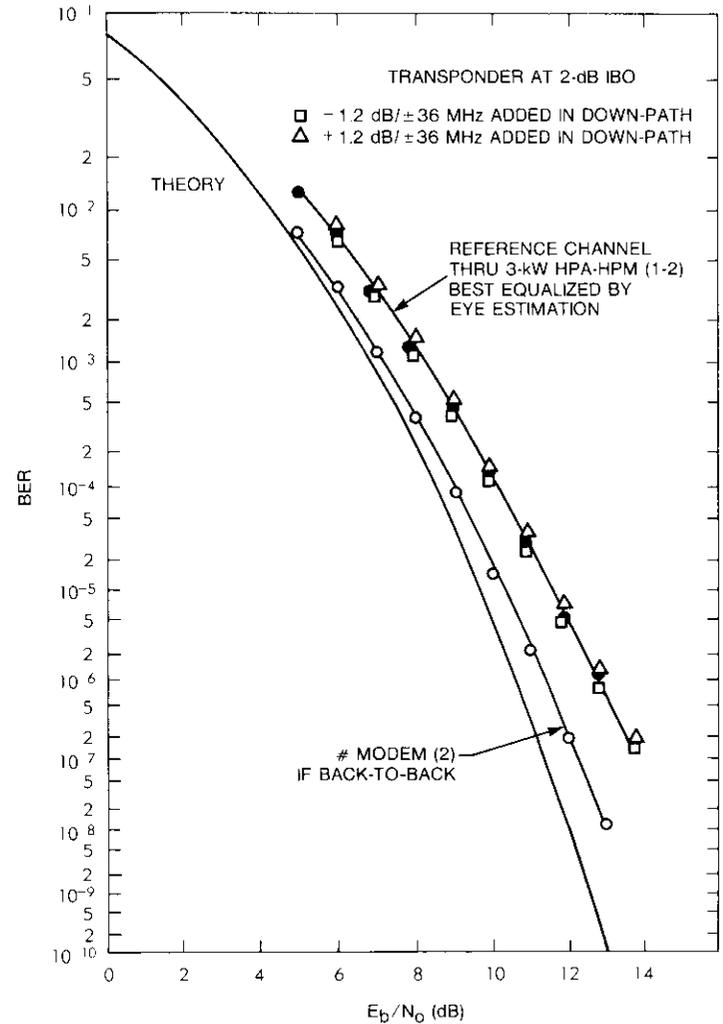


Figure 23. Channel Sensitivity—Amplitude Distortion in Down-Path as Parameter

in the up-path slightly degraded the reference channel, while the negative slope (-1.2 dB/ ± 36 MHz) marginally (within measuring error) improved the reference channel performance (see Figure 22). However, perturbation in down-path characteristics did not produce any measurable change in the reference channel BER.

Similar BER performances resulting from linear group delay perturbations (± 5 ns/ ± 36 MHz) in the up- and down-path responses of the reference channel are shown in Figures 24 and 25, respectively. Both positive and negative group delay (linear) perturbations in the up-path marginally degraded the reference channel. Negative perturbation (-5 ns/ ± 36 MHz) in down-

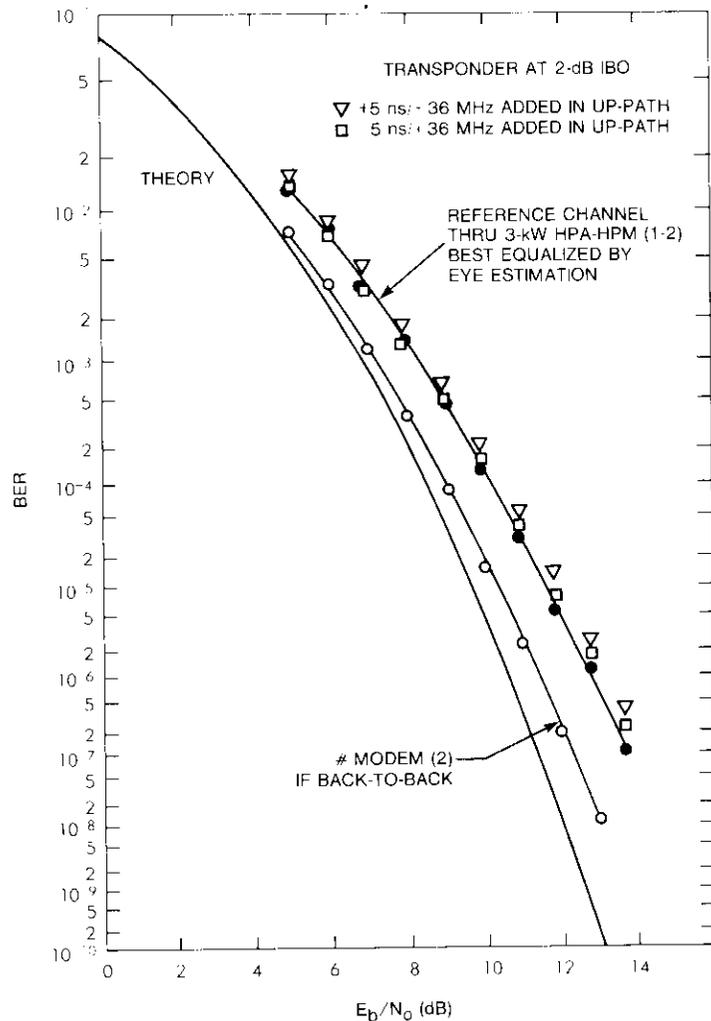


Figure 24. Channel Sensitivity—Linear Group Delay Distortion in Up-Path as Parameter

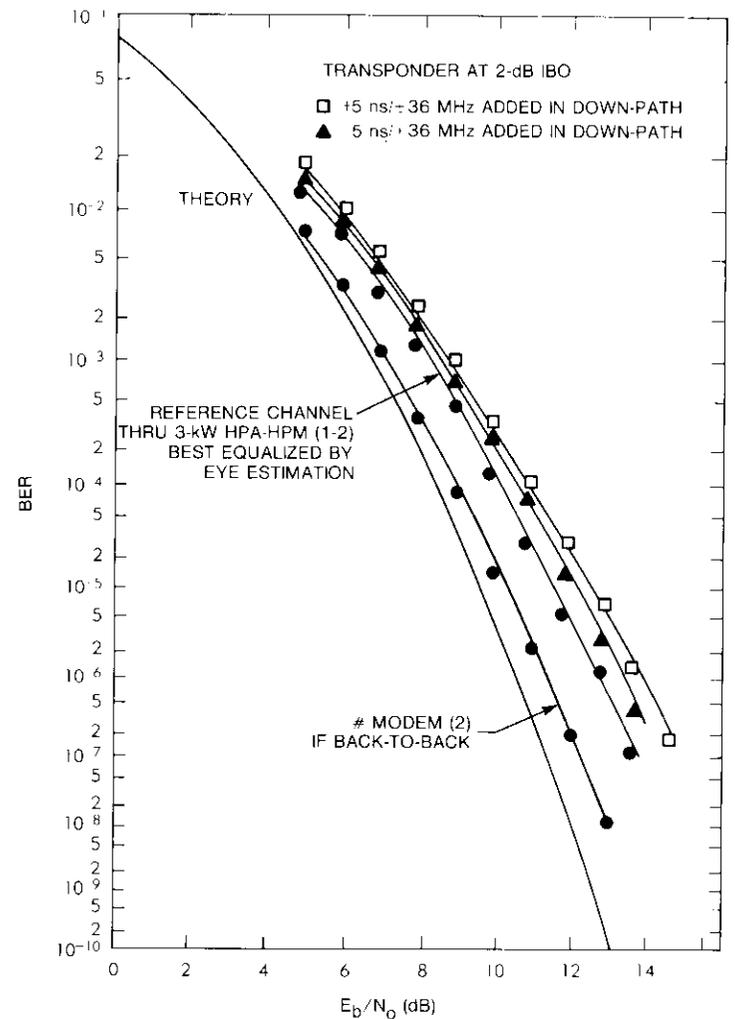


Figure 25. Channel Sensitivity—Linear Group Delay Distortion in Down-Path as a Parameter

path affected the reference channel BER, and positive perturbation ($+5$ ns/ ± 36 MHz) significantly degraded reference channel performance.

The effects of parabolic group delay perturbation (6 ns/ ± 36 MHz, convex or cap and concave or inverted cap) on BER are shown in Figures 26 and 27 for the up-path and down-path perturbation cases, respectively. Slight

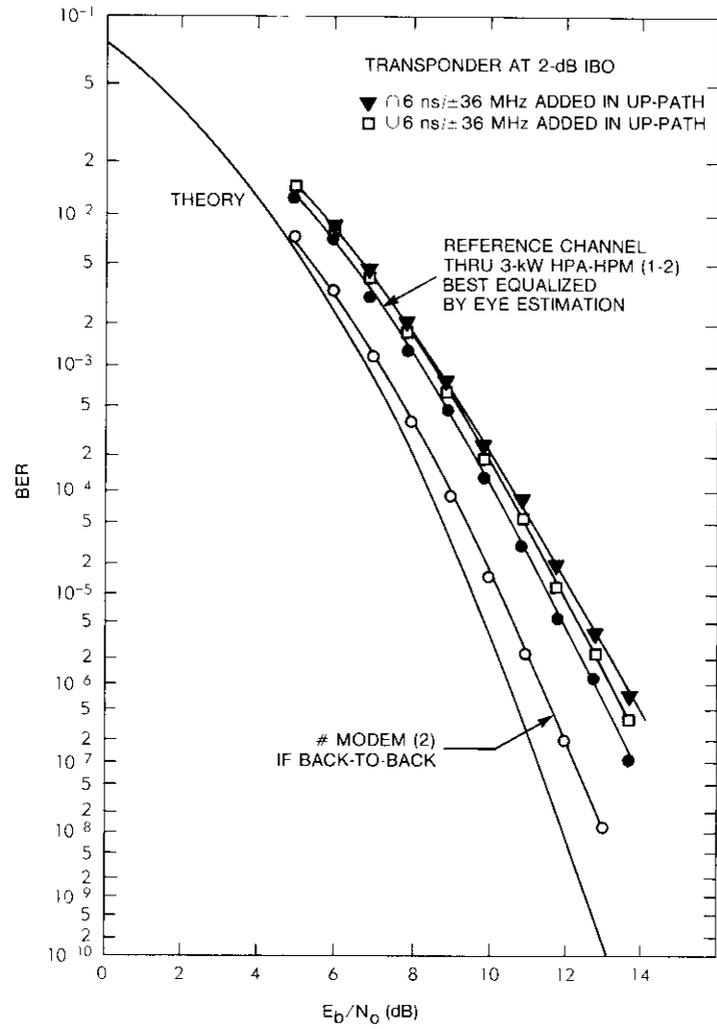


Figure 26. Channel Sensitivity—Parabolic Delay Distortion in Up-Path as Parameter

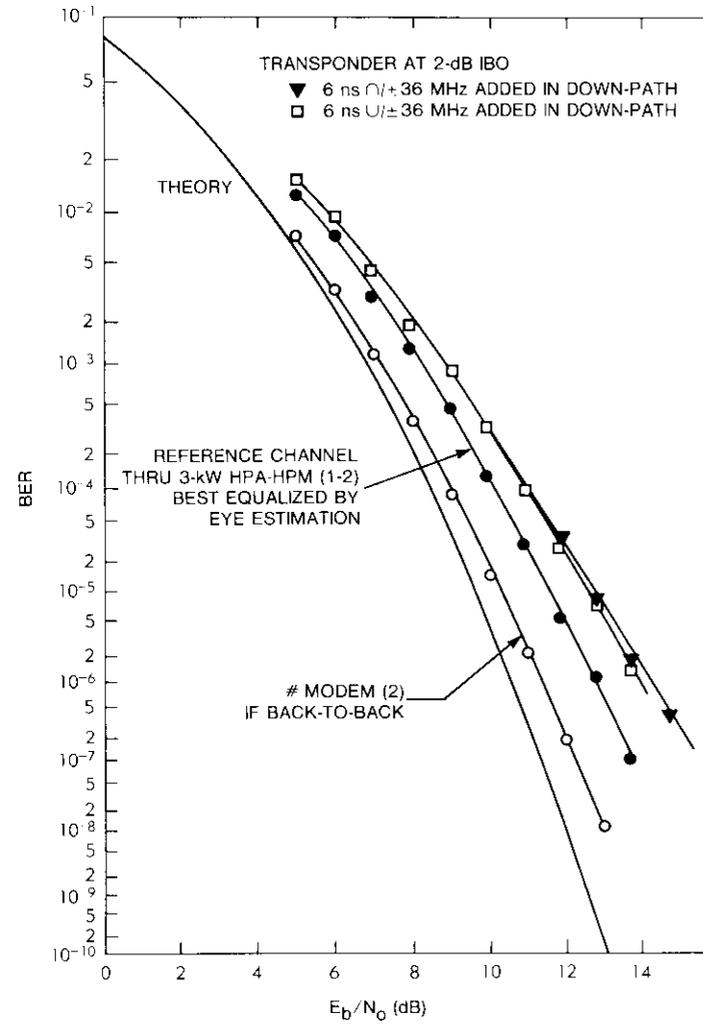


Figure 27. Channel Sensitivity—Parabolic Delay Distortion in Down-Path as Parameter

degradations in the reference channel BER were observed when perturbations were added in the up-path. However, considerable degradation was measured when parabolic (both cap and inverted cap) group delay perturbations were added in the down-path.

Figures 22 through 27 reveal that an optimally or near-optimally equalized channel can tolerate small perturbations in amplitude and group delay responses, whether added in the up-path or in down-path, without drastically altering the BER performance of the link. However, the link BER performance is more sensitive to parabolic delay perturbations when the perturbations are added in the down-path.

Electrical path length (time delay) measurement

To manage the burst time plan in the INTELSAT V TDMA system, the difference between any two electrical lengths of all transmit signal paths within the earth station must be maintained to within 16 ns (one symbol). In addition, the absolute RF/IF electrical path length of the TDMA reference and maintenance station (TRMS) facility must be known to an accuracy of 16 ns or less, including the signal paths from the LNA input coupler to the remote terminal equipment (RTE) interface, and from the RTE interface to the transmit antenna feed port couplers. The TRMS facility will be capable of maintaining the variations in absolute electrical path length to within ± 16 ns.

The basic principle of electrical path length measurement is outlined in Figure 28 [5]. Essentially, the time interval counter measures the time between the beginning of the transmit reference pulse and the delayed pulse through the system under test. A high-speed dual trace oscilloscope was also used as a check in conjunction with the counter to display the start pulse (reference pulse) in one trace and the stop pulse (delayed pulse through the system) in the second trace. From the calibration of the scope horizontal time scale, the delay between the start and stop pulses can be directly evaluated, and the voltage level observed on the oscilloscope can be used to determine the trigger level setting for the counter.

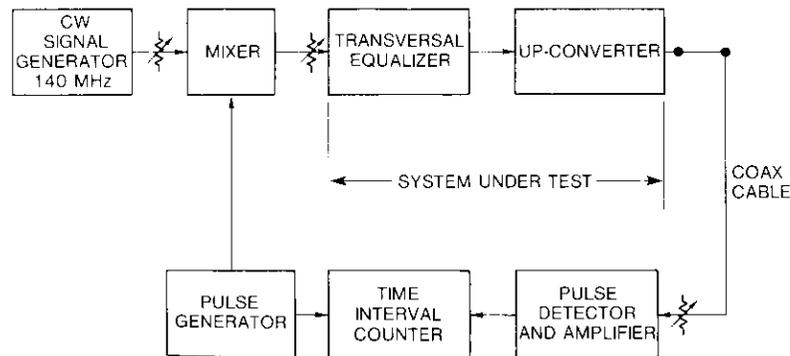


Figure 28. Path Length Measurement Principle

Up-path time delay measurements were made through the on-line and off-line 12-kW HPA systems, as well as the 3-kW HPA systems scheduled for TDMA use. Down-path time delay measurements were also performed. Table 2 gives the up-path delays (between up-converter input and transmit feed point) and the down-path delays (from LNR input to down-converter output).

TABLE 2. EARTH STATION UP- AND DOWN-CHAIN ABSOLUTE TIME DELAY

CONFIGURATION	ABSOLUTE DELAY (ns)	RELATIVE DELAY BETWEEN OFF-LINE AND ON-LINE (ns)
Up-Path Through 12-kW HPA 1, Off-Line	750.3	—
Up-Path Through 12-kW HPA 2, On-Line	746.7	3.6
Up-Path Through 3-kW HPM (1-2)	606.0	—
Up-Path Through 3-kW HPM (1-3)	590.3	15.7
Down-Path LNA 1	557.3	—
Down-Path LNA 2	559.2	1.9

Substitution of the values for the time interval counter jitter, time base stability, and trigger error into the accuracy formula yields a predicted accuracy close to 1 ns and assures that the measured values will fall within the limits specified by INTELSAT. Details of these accuracy calculations are given in the Appendix, which is based on Hewlett Packard Application Note 200-4, "Understanding Frequency Counter Specifications."

Link stability test

The long-term stability of the link BER was examined by using the 12-kW on-line HPA and the 3-kW HPA-HPM system. The BER and the receive carrier power were recorded at regular intervals. A summary of this test is shown in Table 3. The processed data are shown in Figures 29 through 33 in which the BER and the receive carrier power variation are plotted as a function of time. The added noise power (thermal noise in the down-path) source is extremely stable (± 0.1 dB maximum variation), and therefore the variation in BER can be attributed to changing link conditions, such as receive carrier power variation, equalization perturbation, HPA gain fluctuation (giving rise to receive carrier fluctuation), and propagation turbulence.

Figures 29 through 33 reveal that the BER fluctuations followed the receive carrier power fluctuations. After completion of the BER stability test, the link amplitude and group delay responses were reexamined, and it was observed that link responses remained unperturbed during the 1-week test. Therefore, fluctuations in BER were mainly due to receive carrier power variation, which is a cumulative effect of the following:

TABLE 3. SUMMARY OF LINK STABILITY TEST (TRANSPONDER AT 2-dB IBO)

OPERATION NUMBER	LINK CONFIGURATION	DATES		CONTINUOUS OPERATION (Hr)	SET VALUES		MAXIMUM VARIATION ^{a,b}		COMMENTS
		FROM	TO		E_b/N_0 (dB)	BER	CARRIER POWER (dB)	BER	
1	Through 12-kW on-line HPA 1	9/24	9/27	62	12.8	1.1×10^{-6}	+0.5	4×10^{-7}	HPA ran on commercial bus and was subjected to gain variation by ± 1.5 dB
							-0.1	6×10^{-6}	
2	Through 12-kW on-line HPA 2	9/27	9/28	13	12.8	1×10^{-6}	+0.3	6×10^{-7}	HPA has electronic beam voltage regulation and does not experience significant gain fluctuations
							-0.1	1.5×10^{-6}	
3	Through 3-kW on-line HPA-HPM (1-2)	9/28	9/29	14	12.8	2×10^{-6}	+0.4	1.4×10^{-6}	HPA has electronic beam voltage regulation and does not experience significant gain fluctuations
							-0.1	3.5×10^{-6}	
4	Through 3-kW on-line HPA-HPM (1-2)	9/30	10/7	150	12.8	1.4×10^{-6}	+0.3	8.5×10^{-7}	HPA has electronic beam voltage regulation and does not experience significant gain fluctuations
							-0.3	5×10^{-6}	

^a Measurements were recorded at a regular interval by printers attached to the BER test set and to the digital power meter that recorded instantaneous values of receive carrier power.

^b The receive carrier power recorded included added thermal noise, which is subject to a ± 0.1 -dB variation.

- a. earth station e.i.r.p. fluctuations attributable to HPA gain variation,
- b. drift in satellite TWTA power output,
- c. satellite repeater gain fluctuations,
- d. satellite antenna pointing error,
- e. earth station antenna movement caused by wind loading, and
- f. atmospheric turbulence (if any).

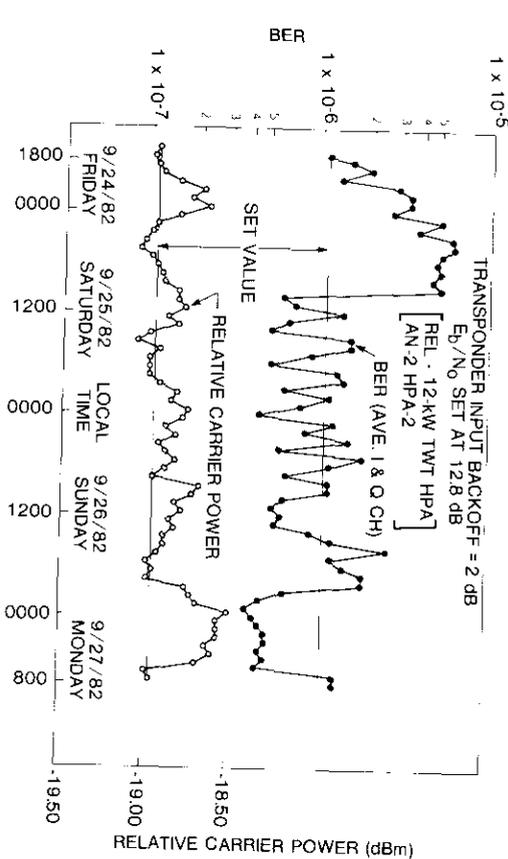


Figure 29. Link BER Stability Observation (9/24/82-9/27/82)

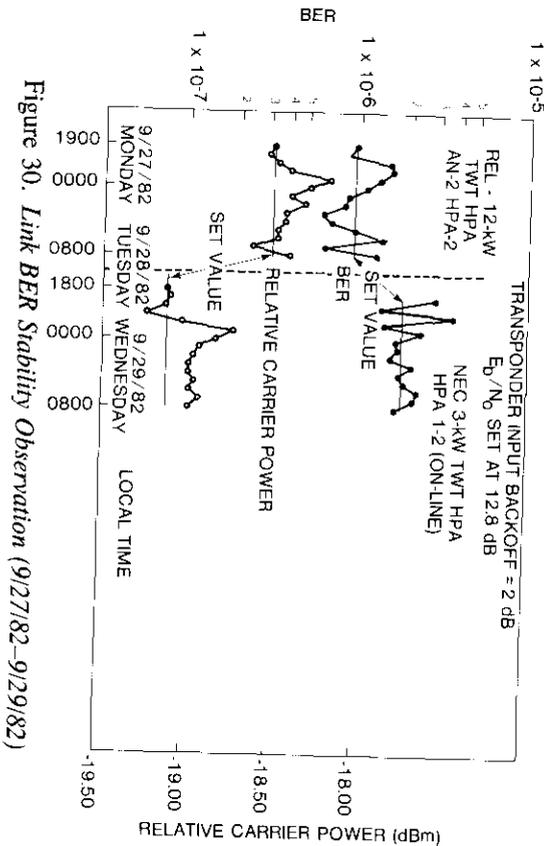


Figure 30. Link BER Stability Observation (9/27/82-9/29/82)

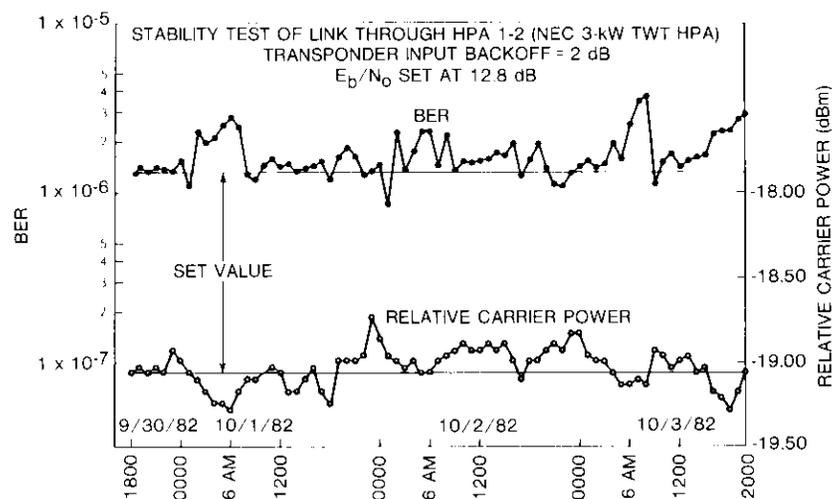


Figure 31. Link BER Stability Observation (9/30/82-10/3/82)

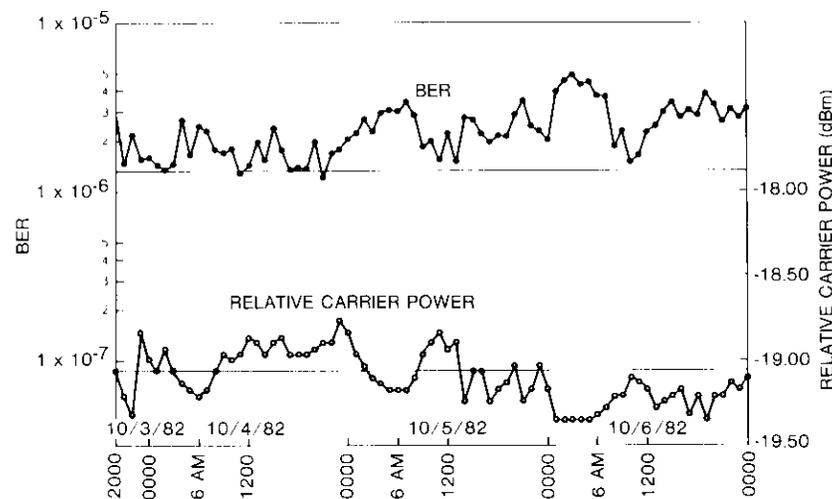


Figure 32. Link BER Stability Observation (10/3/82-10/6/82)

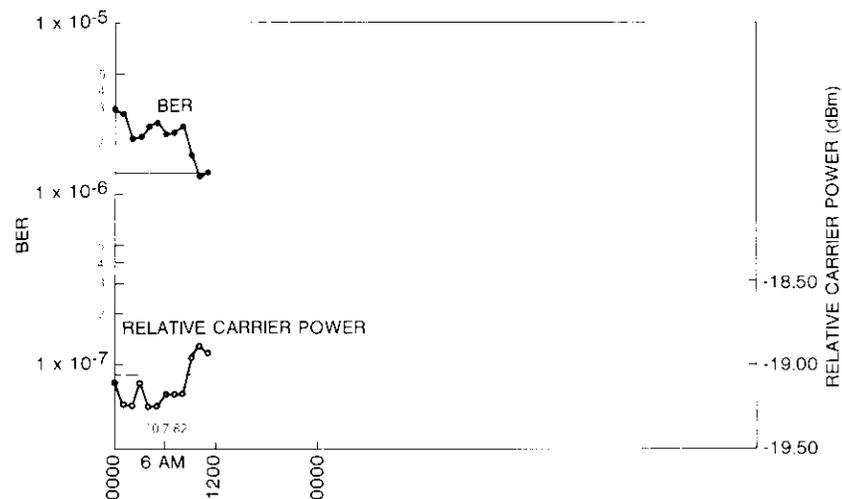


Figure 33. Link BER Stability Observation (10/7/82)

Conclusions

As the result of a systematic evaluation of the 120-Mbit/s QPSK channel via INTELSAT V (F2) Transponder 41 (Ch 1-2) and the AN-II link, it was concluded that the INTELSAT TDMA/DSI Specification (BG-42-65) for up- or down-path amplitude and group delay masks can be met by using fixed IF (140 MHz) equalizers in conjunction with 3-tap (IF-passband) transversal equalizers. The optimally equalized channel yields a BER of 10^{-6} at $E_b/N_0 \cong 12.9$ dB, with the transponder operating at 2-dB IBO and the HPA operating in a linear region. Link equalization can be repeated with consistent BER performance. Adjacent channel interference arising from PSK spectrum regrowth degrades the top channel NPR of an adjacent transponder 972-channel FDM/FM test carrier by about 0.5 to 1 dB when the PSK carrier transponder is operated at 2-dB IBO. This deficiency can be corrected by a marginal increase in the FM carrier e.i.r.p., if required. The channel equalization sensitivity study revealed that, in an optimally or near-optimally equalized channel, nominal amplitude or group delay perturbations in the up- or down-path do not significantly affect the channel BER. However, the channel is more sensitive to parabolic group delay perturbation in the down-path.

A 1-week link BER stability test revealed that the maximum variation in BER ranges from 1.0×10^{-6} to 5×10^{-6} when the set value of BER is 1.1×10^{-6} . The mean value of BER is around 1.3×10^{-6} .

Acknowledgments

This program was initiated and supported by the Project Management Office of the World Systems Division. Thanks are due to the headquarters staff for their coordination efforts and to the earth station staff for their support. Acknowledgment is made to INTELSAT for providing the Atlantic basin spare INTELSAT V satellite for use in this test program. Finally, thanks are due to M. Kappes and G. House for their help.

References

- [1] Contribution of the U.S. Signatory, "TDMA Amplitude and Group Delay Equalization and Electrical Path Length Measurements," BG-T-44, November 12, 1982.
- [2] P. R. Geffe, *Simplified Modern Filter Design*, New York: Hayden Book Co., Inc., 1963.
- [3] Members of the Technical Staff, Bell Telephone Laboratories, *Transmission Systems for Communications*, Revised Fourth Edition, December 1971, pp. 713-746.
- [4] R. W. Gruner and A. E. Williams, "A Low-Loss Multiplexer for Satellite Earth Terminals," *COMSAT Technical Review*, Vol. 5, No. 1, Spring 1975, pp. 157-178.
- [5] H. H. Ricker III, private communication.

Appendix. Accuracy of time delay measurements

The Hewlett-Packard 5370A Time Interval Counter (TIC) was used to measure the time delay through various earth station paths at Andover. The operating manual for this counter defines the accuracy for time interval measurements as

$$\alpha = \text{Jitter} \pm 700 \text{ ps systematic} \pm \text{time base} \pm \text{trigger error}/\sqrt{N}$$

where Jitter = 35 ps rms typical, 100 ps rms maximum, and N = number of samples (1,000 was chosen for these tests). Time base stability is a combination of aging, temperature variation, and line voltage variation,

where

- aging $< 3 \times 10^{-7}$ per month
- temperature $< 5 \times 10^{-6} 0^{\circ}\text{C}$ to 55°C
- line voltage $< 1 \times 10^{-8} \pm 10$ percent nominal.

Trigger error can be calculated by

$$\text{Trigger error} = \frac{\sqrt{(150 \mu\text{V})^2 + e_n^2}}{\text{input voltage slew rate at trigger points}} \text{ s rms}$$

The typical input amplifier noise on the 5370A is $150 \mu\text{V}$, and e_n is the rms noise of the input signal for a 500-MHz bandwidth system. Time base error, trigger error, and accuracy are calculated as follows.

Time base error

A worst-case condition will be assumed. Therefore, if the counter has not been calibrated for 1 year, the total $\Delta f/f$ attributable to aging (from manufacturer's data given above) is $3 \times 10^{-7}/\text{month} \times 12 \text{ months} = 3.6 \times 10^{-6}$. For temperature variations, the worst-case $\Delta f/f$ is 5×10^{-6} . By including worst-case line voltage variation, the total $\Delta f/f$ attributable to all sources is 8.7×10^{-6} . From measurements, the approximate delay time through the earth station link is $1 \mu\text{s}$. The error attributable to the time base is then $(\pm 8.7 \times 10^{-6})(1 \times 10^{-6}) = \pm 8.7 \times 10^{-12} = \pm 0.0087 \text{ ns}$.

Trigger error

To determine trigger error, knowledge of the noise input to the counter is required. The amplifier used after the detector has a specified equivalent wideband input noise level of $40 \mu\text{V}$, maximum. With 40-dB gain, the maximum noise level at the output of the amplifier will be $4,000 \mu\text{V}$.

The voltage slew rate at the trigger level of the counter is also required for the trigger error calculation. Laboratory observations and field tests imply a slew rate of $0.5 \text{ V}/15 \text{ ns}$, or $3.3 \times 10^7 \text{ V/s}$.

Substitution of the above values into the equation for trigger error yields

$$\text{Trigger error} = \frac{\sqrt{(150 \times 10^{-6})^2 + (4,000 \times 10^{-6})^2}}{3.3 \times 10^7} \text{ s rms}$$

$$\cong 0.13 \text{ ns}$$

Since the trigger error is associated with both the start and stop inputs, the total trigger error for the measurement will be $1.4 \times 0.13 \text{ ns}$ (trigger errors are added on an rms basis). The actual trigger error is $0.182 \text{ ns (rms)} = 182 \text{ ps (rms)}$.

Accuracy

Accuracy also depends on the number of samples taken for each counter reading

displayed. Tests were made with the sample size set at 1,000. Substitution of the values for jitter, time base, and trigger error into the accuracy formula yields

$$\text{Accuracy} = 100 \text{ ps} \pm 700 \text{ ps} \pm 8.7 \text{ ps} \pm 182 \text{ ps}/\sqrt{1,000}$$

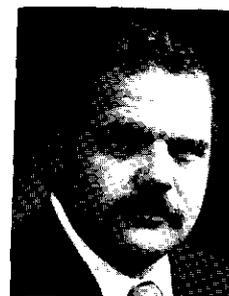
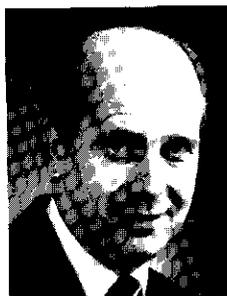
$$\alpha = 814 \text{ ps (maximum)}$$

The predicted accuracy is close to 1 ns and assures that the measured values will fall within the limits specified by INTELSAT.



Dayamoy Chakraborty received a Ph.D. in microwave physics from the University of Surrey, England, in 1967. Since 1968, he has been employed by COMSAT. From 1968 to 1971, he was involved with the satellite system optimization study and the analysis of transmission deviations. Dr. Chakraborty transferred to COMSAT Laboratories in 1971 and has been involved in developing high-speed digital satellite communications systems. Currently, as a Senior Staff Scientist in the Digital Communications Laboratory, he directs research in nonlinear satellite channel model design, maximum likelihood sequence estimation, and digital link technology. Prior to joining COMSAT, he was employed by the British Post Office, Research Department, Dollis Hill, London. Dr. Chakraborty is the author of numerous publications on microwave techniques and communications technology, and is a Senior Member of IEEE and a Corporate Member of IEE.

John W. Ehrmann received a B.S.E.T. from the Capitol Institute of Technology, Washington, D.C., in 1969 and a B.S.E.E. from the George Washington University in 1976. Since joining COMSAT in 1969, he has worked in the Systems Simulator Laboratory, the Earth Station Engineering Division, and the Monitor and Control Engineering Division. He is currently a Staff Engineer at COMSAT Laboratories' Maintenance and Supply Center, where he is responsible for coordinating various TDMA tasks.



Roger S. Miner received a B.S. from the University of New Haven in 1974. Since joining COMSAT in 1974, he has been active in areas of satellite communications, including TT&C, Maritime SATCOM, domestic systems satellite testing, computer systems, digital data processors, satellite video systems, and RF telecommunications systems. He is currently a Senior Technical Specialist in the Integration, Installation, and Test Department, where he is working on the integration of TDMA at COMSAT earth stations. Prior to joining COMSAT, Mr. Miner was employed in the broadcasting industry as Chief Engineer at AM and FM radio and television stations.

Gene Rosch received a B.S.E.E. and an M.S.E.E. from the State University of New York at Stony Brook in 1979 and 1980, respectively. He joined COMSAT in 1981 and is a member of the Headquarters Engineering Staff of the World Systems Division. He is currently involved with introducing TDMA/DSI technology into the INTELSAT satellite system. Mr. Rosch is a member of IEEE, Tau Beta Pi, and Eta Kappa Nu.



Serial implementation of Viterbi decoders

A. SHENOY AND P. JOHNSON

(Manuscript received June 17, 1983)

Abstract

This paper describes the implementation of two Viterbi decoders that handle input data serially. Internal computations are also performed in a serial rather than a parallel fashion. This allows for low-cost implementation with moderate complexity, using only standard Schottky TTL and memory chips. Synchronization circuits and input data buffers provide proper operation. Each decoder design was implemented with less than 72 chips.

One design is a rate 1/2 decoder with 3-bit input quantization and a path memory length of 64 bits. The other is a rate 2/3 decoder with 4-bit input quantization and a path memory length of 128 bits. Both designs have a constraint length of 7, with an input data rate of 64 kbit/s at the encoder. Bit-error rate (BER) tests indicate a coding gain of 4.5 dB for the rate 1/2 decoder and 3.0 dB for the rate 2/3 decoder at a BER of 10^{-4} . These findings correspond well with the expected theoretical coding gains with infinite quantization of 4.7 dB and 3.5 dB, respectively.

Introduction

Forward error correction (FEC) coding techniques have been used for several decades to improve the performance of digital communications systems. Depending on the system complexity and performance desired, digital data are either block or convolutionally encoded. While encoders for either approach are relatively straightforward, decoding procedures are far

more complicated. The development of practical decoding algorithms for convolutional codes has centered around probabilistic decoding algorithms and threshold or feedback decoding procedures.

Viterbi decoding

Probabilistic decoding algorithms usually calculate the likelihood of successful message transmission on the basis of reliability information extracted from present and past symbols. Therefore, a memory of suitable size and an arithmetic capability for message probability computation are required. Viterbi [1] first proposed a specific path decoding algorithm for convolutional codes which significantly reduced the number of computations needed to choose the most probable path. The Viterbi algorithm was shown to be an efficient and practical decoding technique for short constraint length codes by Heller [2], [3]. Omura [4] and others demonstrated that the Viterbi algorithm was, in effect, a maximum-likelihood procedure.

A maximum-likelihood decoder would calculate the likelihood of the received data for code symbol sequences on all paths through the code trellis. Figure 1 shows a typical code trellis diagram for a rate 1/2 convolutional encoder with constraint length 3.

A Viterbi decoder calculates the likelihood that each of the paths will enter a given state and chooses the most likely path leading to that state. This is done for each state at a given trellis depth. Thus, after each decoding operation, only one path remains leading to the given state. The decoder then proceeds one level deeper into the trellis, and the process is repeated.

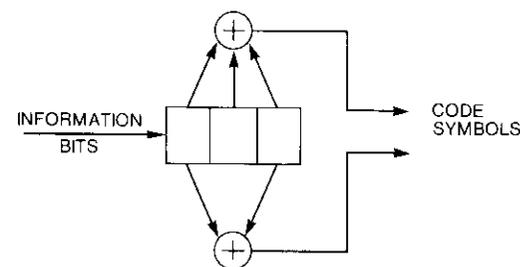
For a rate k/n convolutional encoder of constraint length K , the decoder always retains a set of $2^{K(K-1)}$ paths after each decoding step. It has been shown that, for a fixed amount of path history storage, the oldest bit on the most likely of the $2^{K(K-1)}$ paths can be output as the decoded bit [5]. The amount of path storage required is

$$m = h \cdot 2^{K(K-1)} \tag{1}$$

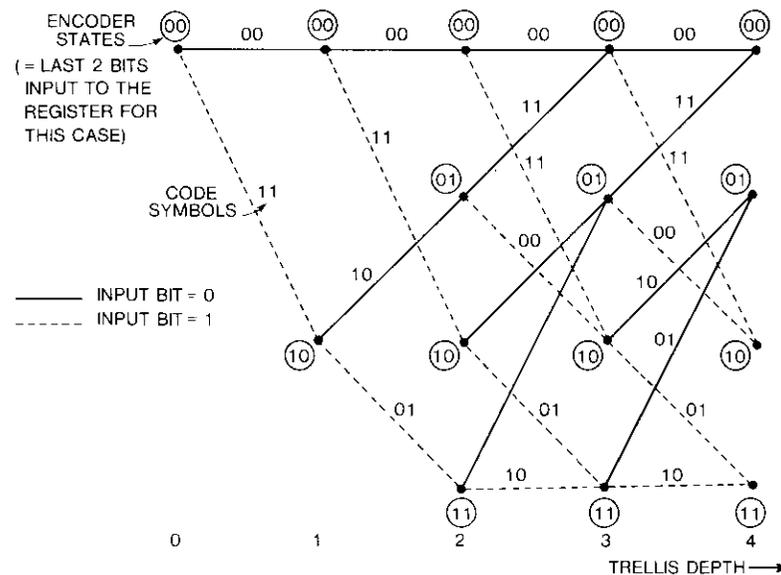
where h is the length of the information bit path history per state, which should be four or five times the constraint length.

Viterbi [6] has suggested using the log-likelihood function (called metric) to calculate the likelihood for each path entering a given state. Thus, when a path is extended by one branch, the metric of the new path is the sum of the new branch symbol metrics and the old path metric.

It is more convenient, however, to use integers as symbol metrics. These integers represent the distance, in terms of quantization size, of the received symbol from the reference code symbol positions on the signal map. The



(a) K = 3, RATE 1/2 CONVOLUTIONAL ENCODER



(b) CODE TRELLIS DIAGRAM

Figure 1. Example of a Typical Convolutional Encoder

smaller this distance, the greater the probability that the received symbol is indeed the reference code symbol from which the distance is being measured. Thus, rather than select as the most likely path the one with the largest value of the log-likelihood function, the selection is based on the path with the smallest state metric expressed in integer form.

The implementation described in this paper employs the square of the

integer distance [7]. Squared values are scaled so that their digital representation does not exceed a desired number of bits. Scaling leads to a loss of resolution; however, experimental results indicated that the degradation in performance is insignificant. Viterbi [6] and Heller and Jacobs [8] should be consulted for more detailed discussions of the Viterbi algorithm.

Serial implementation of Viterbi decoders

The basic Viterbi decoder can be divided into the following five functional units [8]:

- a. an input section,
- b. a branch metric calculation section,
- c. an add-compare-select (ACS) arithmetic section,
- d. a path memory, and
- e. an output section.

The branch metric calculation section accepts the input data and calculates the metric for each distinct branch. For a rate k/n code, 2^k possible branches are considered for each state. Each of these branch metrics must be added to the corresponding old state metric to yield the new state metrics from which the minimum state metric is selected. At this point, the following two modes of implementation are possible:

- a. a parallel form with dedicated arithmetic sections for each of the 2^k branch and state metric calculations; and
- b. a serial form with only one ACS arithmetic section, which sequentially receives each of the 2^k branch and state metrics.

The implementation described in this paper follows the serial architecture shown in Figure 2. This results in a decoder structure of moderate complexity, small size, and low cost. Two decoders are presented, with rates 1/2 and 2/3, which are designed for data rates up to 100 kbit/s.

The encoders for the rate 1/2 and 2/3 codes are relatively straightforward and are implemented by using serial shift registers that have exclusive-OR logic gates, with a connection pattern following Odenwalder's [9] and Paaske's [10] suggestions. Both encoders have a constraint length of 7.

Input section

The two decoders implemented run on fast internal clocks, typically 10 MHz. The input section consists of buffers for the input symbols and logic circuitry to synchronize the start and stop of the decoding process. After all calculations for one trellis depth are completed, the decoder halts and waits for the next symbol to arrive before proceeding to the next trellis depth.

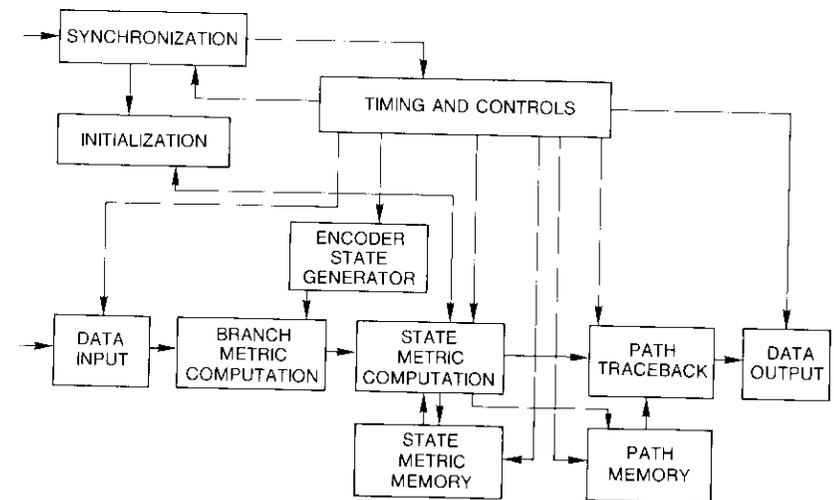


Figure 2. General Serial Architecture for the Viterbi Decoder

Power-on initialization is provided, which prevents the state metric memory from holding arbitrary digital words when power is applied to the circuit. This is accomplished by loading all the branch metrics corresponding to the first received symbol into the state metric memory. The decoding process then begins with the next received symbol. If, during the present decoding process, the next received symbol arrives too early, the decoder does not recognize it until the present decoding step is completed. Therefore, there is a maximum data rate that the decoder can tolerate. Figure 3 is a block diagram of the input section.

Branch metric calculation section

The rate 1/2 and 2/3 decoders use 3- and 4-bit quantizations, respectively. Figures 4 and 5 show the signal maps and code symbol locations with respect to the quantization levels for quadrature phase shift keying (QPSK) and octal phase shift keying (OPSK) modulation formats, respectively.

Branch metrics are calculated as scaled integers and are stored in a PROM look-up table. For a given received symbol, all 2^k branch metrics ($k = 1$ or 2) for all encoder states at a given trellis depth are read out sequentially from the PROM and sent to the ACS section. An encoder replica generates all the possible encoder states.

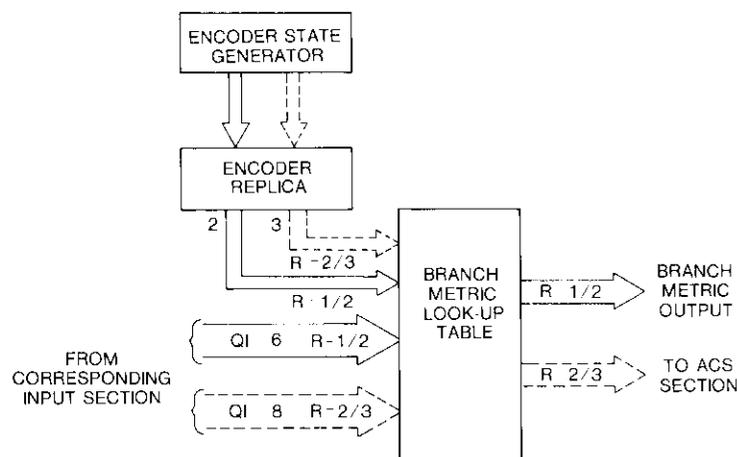


Figure 5. Branch Metric Calculation Section

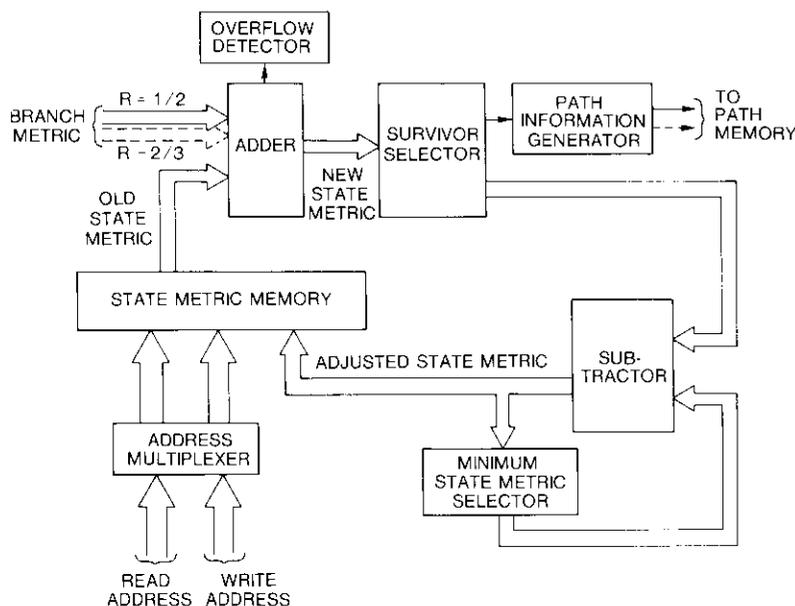


Figure 6. Add-Compare-Select Arithmetic Section

rate 1/2 decoder has two branches leading to each encoder state; so a single bit (1 or 0) can indicate which of the two was chosen as the most likely path. The rate 2/3 decoder has $2^2 = 4$ branches for each state; thus, a 2-bit designation (00, 01, 10, or 11) is needed. There are $2^6 = 64$ encoder states to be considered at each trellis depth, with a constraint length of 7 for both the codes. Thus, there are 64 groups of path bit data after each decoding step.

The minimum path history length should be at least four or five times the constraint length [5], which gives 35 sets for both codes. The implementation considered here uses 64 sets, which is nearly twice the required minimum. Each set consists of 64 paths containing 1 or 2 information bits.

A path traceback is done every time new path data are written in. The traceback starts from the location of the overall minimum state metric for the previous trellis depth. This location identifies the position of the most likely path. As the 64 present path data bits are written in for a given trellis depth, the decoder traces back through 63 past path information sets. When the last present path bit data are written in, the bits that were read out on the last traceback are output as the decoded symbol, and the decoder then disables all operations and waits for the next received symbol to be clocked in.

Path storage and path traceback are done alternately, since they involve the same memory. After 64 sets, the present path information starts overwriting old information, since path history that is more than 64 sets old is not used. Figure 7 is a block diagram of the path memory and traceback section.

Output section

Since the original data were input as a serial binary stream at the encoder, a serial binary stream must also be sent out by the decoder. The output section, shown in Figure 8, converts the parallel output that is read out of the path history RAM to a final serial output. The output clock and data are appropriately buffered.

Experimental results

Prototypes of the two codecs were built and tested. The rate 1/2 codec was built on three 7.5×5.5 in. (18.3×13.8 cm) wire-wrapped boards with 60 chips of Schottky integrated circuits, 1 PROM, and 3 RAMs. The rate 2/3 codec was built on four such wire-wrapped boards with 71 Schottky chips, 1 PROM, and 4 RAMs. Figures 9 and 10 are photographs of these prototypes.

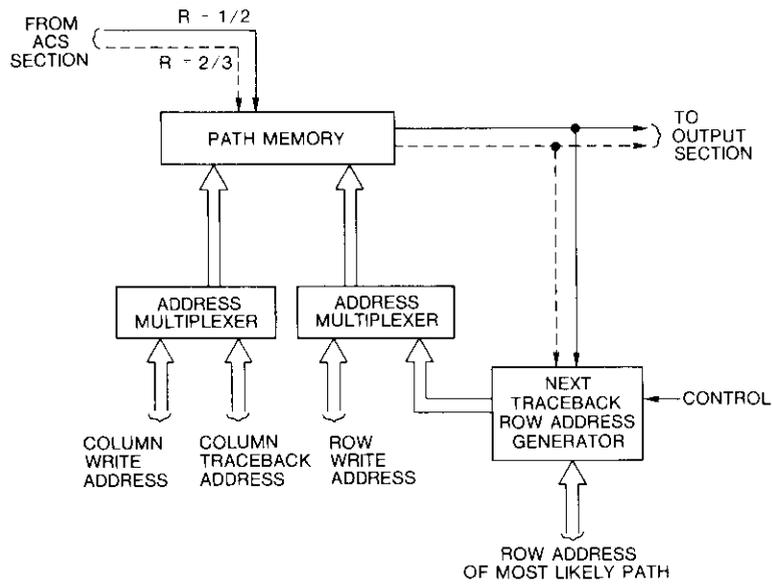


Figure 7. Path Memory and Traceback Section

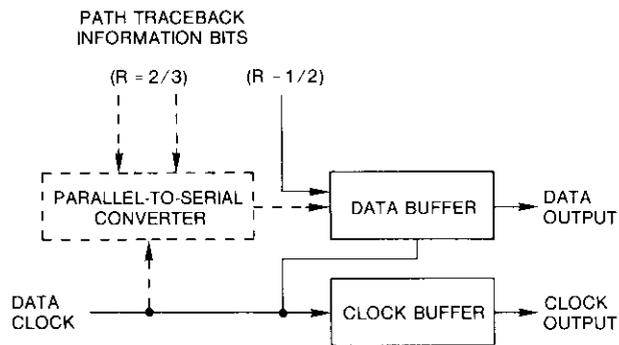


Figure 8. Output Section

The prototypes were connected to a soft-decision modem and test equipment to measure performance. Figure 11 is a general block diagram of the test setup. A pseudorandom bit stream was generated by the test set and sent to the FEC encoder. The encoder output was then modulated and transmitted on a channel. Thermal noise with a flat spectrum up to 150 MHz was used to corrupt the channel symbols. Soft-decision outputs from the demodulator

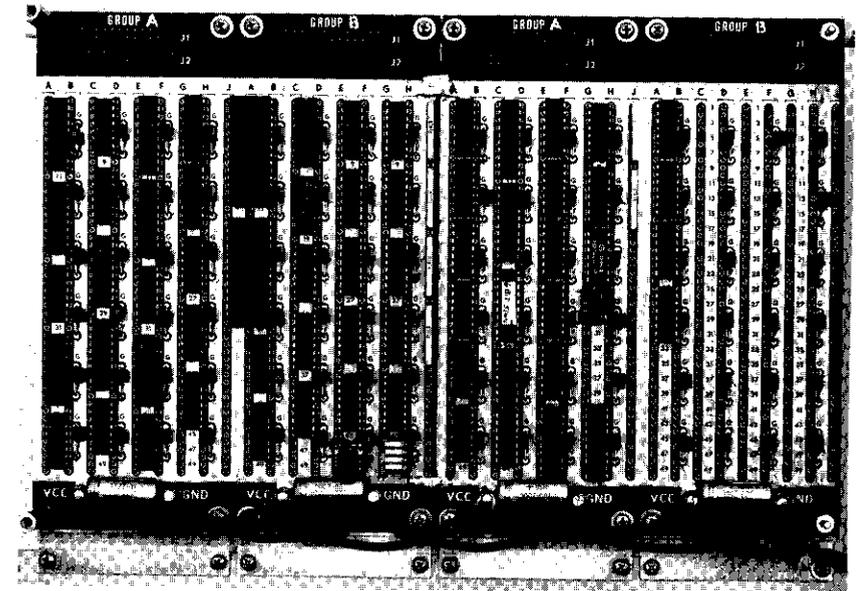


Figure 9. Rate 1/2 Viterbi Decoder

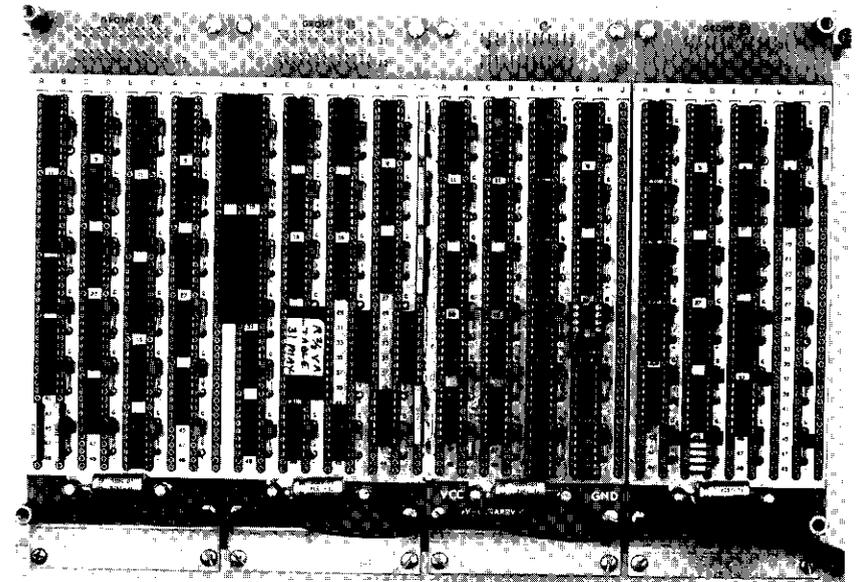


Figure 10. Rate 2/3 Viterbi Decoder

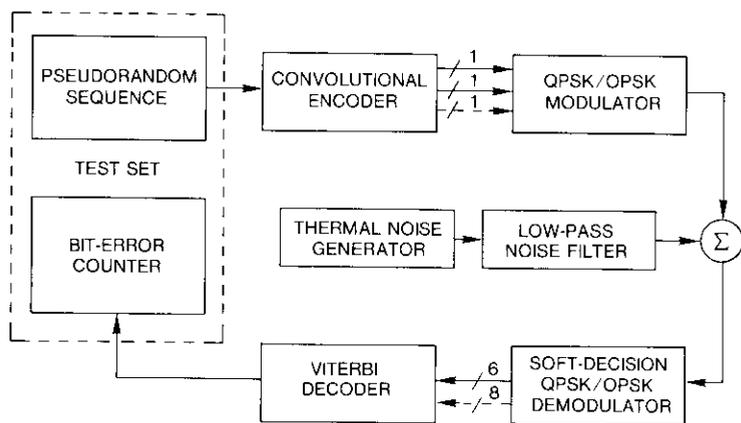


Figure 11. Test Setup Block Diagram

were then sent to the decoder, and the estimated transmitted sequence was routed back to the test set for BER calculations. Both decoders were tested at the design data rate of 64 kbit/s. They were also found to perform well at data rates up to about 104 kbit/s.

As shown in Figure 12, the rate 1/2 decoder had coding gain of 4.5 dB at a BER of 10^{-4} for 3-bit soft decision decoding (a 0.2-dB difference from the theoretical coding gain of 4.7 dB). The rate 2/3 decoder had a coding gain of 5.3 dB at a BER of 10^{-4} for 4-bit soft decision decoding, as illustrated in Figure 13. This coding gain was measured in comparison to uncoded QPSK and was 0.6 dB from the theoretical value of 5.9 dB. Compared to uncoded QPSK of the same symbol rate, the resulting coding gain was 3.0 dB (a 0.5-dB difference from the theoretical value of 3.5 dB).

Conclusions

Two implementations of the Viterbi decoding algorithm have been discussed. Both implementations follow a serial architecture, which results in a decoder of moderate complexity, small size, and low cost, because commercially available chips are used. Both decoders gave good BER performance; the degradation from the ideal was a small fraction of 1 dB. It must be remembered that the theoretical performance assumes infinite quantization and a full-range metric. Since the implementations presented have a finite quantization and a scaled metric, some degradation is expected. The prototypes discussed represent a good tradeoff between implementation complexity and BER performance.

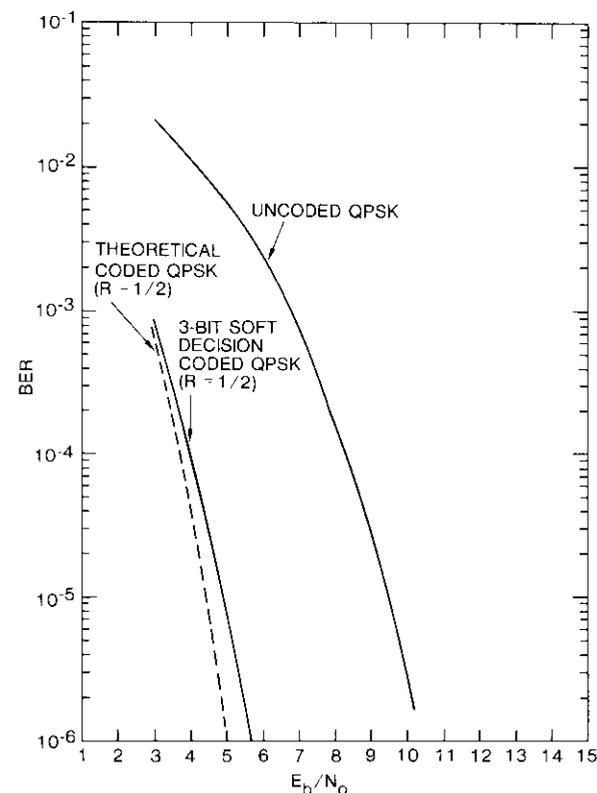


Figure 12. Rate 1/2 Viterbi Decoder Performance

While all subsections of the architecture were designed to minimize the number of chips required, advances in integrated circuit technology make it possible to readily upgrade the architecture to achieve higher speeds or lower chip counts. Replacement of the Schottky TTL chips with ECL chips would make data rates in megabits per second achievable with only slight modifications to the architecture. Alternatively, the replacement of several MSI chips by programmable logic arrays (PLAS) could reduce the chip count by 30 to 50 percent.

Viterbi decoders implemented with a parallel architecture and operating at megabits per second are commercially available but are expensive and would be excessive for applications that require low data rates. The decoders presented here were designed for low-cost, low-data-rate applications.

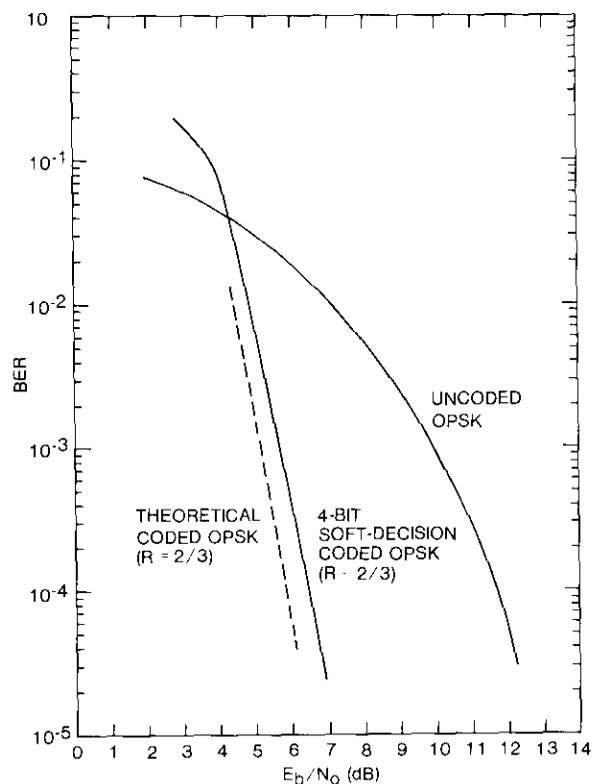


Figure 13. Rate 2/3 Viterbi Decoder Performance

Acknowledgments

The authors wish to thank Lin-Nan Lee for his constant guidance and encouragement; Neal Becker for his excellent design ideas; Chi-Yuan Wu, who conceived the architecture that was later refined and who did some of the preliminary work; and Ming-Jung Wu, who set up the test equipment and patiently gathered the test data.

References

- [1] A. J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," *IEEE Transactions on Information Theory*, IT-13, April 1967, pp. 260-269.

- [2] J. A. Heller, "Short Constraint Length Convolutional Codes," Jet Propulsion Laboratory, California Institute of Technology, *Space Programs Summary 37-54*, Vol III, October/November 1968, pp. 171-177.
- [3] J. A. Heller, "Improved Performance of Short Constraint Length Convolutional Codes," Jet Propulsion Laboratory, California Institute of Technology, *Space Programs Summary 37-56*, Vol III, February/March 1969, pp. 83-84.
- [4] J. K. Omura, "On the Viterbi Decoding Algorithm," *IEEE Transactions on Information Theory*, IT-15, January 1969, pp. 177-179.
- [5] LINKABIT Corp., "Coding Systems Study for High Data Rate Telemetry Links," Final Report for Contract NASA-6024, NASA Ames Research Center Report CR-114278, Moffett Field, California.
- [6] A. J. Viterbi, "Convolutional Codes and Their Performance in Communication System," *IEEE Transactions on Communications*, COM-19, October 1971, pp. 751-772.
- [7] L. N. Lee and C. Wu, private communication.
- [8] J. A. Heller and I. M. Jacobs, "Viterbi Decoding for Satellite and Space Communications," *IEEE Transactions on Communications*, COM-19, No. 5, October 1971, pp. 835-848.
- [9] J. P. Odenwalder, "Optimal Decoding of Convolutional Codes," Ph.D. dissertation, School of Engineering and Applied Science, University of California, Los Angeles, 1970.
- [10] E. Paaske, "Short Binary Convolutional Codes with Maximal Free Distance for Rates 2/3 and 3/4," *IEEE Transactions on Information Theory*, IT-20, September 1974, pp. 683-689.

Ajit Shenoy received a B. Tech. in electronics in the First Class, with Distinction, from the Indian Institute of Technology, Madras, India in 1979, and an M.S.E.E. from the Stevens Institute of Technology in 1981. He completed an additional year of graduate study in telecommunications and digital systems before joining COMSAT Laboratories in June 1982. He is currently a Member of the Technical Staff in the Transmission Systems Laboratory, where his work involves the design and implementation of advanced digital logic and microprocessor-based communications hardware, the study and analysis of transmission techniques, and the computer simulation of satellite transmission systems. Prior to joining COMSAT, he worked on electronic motor speed control with Philips India Ltd. and was a consultant for Synfax Manufacturing Inc., Belleville, N.J., where he designed instrumentation systems for quality control. Mr. Shenoy is a recipient of several awards and scholarships from the Government of India. He is a member of IEEE.





Peter N. Johnson is a Technical Specialist in the Transmission Systems Techniques Department at COMSAT Laboratories. His work involves the development, design, testing, and troubleshooting of digital communications systems. Prior to joining COMSAT in 1981, he specialized in instrumentation and high-voltage systems design at General Ionex Corporation. He is currently studying for a B.S.E.E.

Index: amplifiers, TDMA, traffic statistics, transmitters, transponders

Adaptive satellite power amplifier operation for TDMA down-links*

R. E. EAVES

(Manuscript received December 10, 1982)

Abstract

In conventional payload design and operation, communications satellite power amplifiers remain on and consume full power at all times. In the approach considered in this paper, each power amplifier supporting a time-division multiple-access (TDMA) down-link is turned off for those portions of the frame that do not carry traffic or other information. Potential reductions in power requirements are analyzed by decomposing satellite traffic into several components with different rates of variation. The time-varying components are characterized to estimate the effect of adaptive operation on power requirements. Practical issues of implementation, including on-board control and the implications for spacecraft subsystems, are considered to determine the degree to which time-varying power consumption can be exploited to reduce spacecraft requirements.

Introduction

Conventional satellite design is based on power amplifiers (PAs) that consume power at all times. In most applications, traffic approaches maximum capacity when the satellite is near saturation, and then only for relatively

* This paper is based upon work performed at COMSAT, INTELSAT Technical Services, under the sponsorship of the International Telecommunications Satellite Organization.

brief periods of the day. Consequently, a significant amount of the power consumed during the life of the satellite contributes in no way to the transmission of information.

The alternate approach considered in this paper reduces the average power consumption of TDMA or satellite-switched (SS)-TDMA down-links that are not filled to capacity. This is accomplished by rapidly switching PAS on and off in a controlled manner to conform to the transmission patterns actually present within the frame. Thus, the demands on the spacecraft power supply and other support systems are reduced. The larger objective is to make unused spacecraft resources available for additional payload functions, or to reduce spacecraft requirements.

Adaptive operation may be practical for both traveling wave tube amplifiers (TWTAs) and solid-state power amplifiers (SSPAs). Application of such an approach to earth station TWTAs has been investigated [1]. In this paper, only SSPAs are considered, for two reasons. First, it is simpler to implement rapid on-off switching for SSPAs, and second, SSPAs can be expected to increasingly replace TWTAs in most spacecraft applications.

A generalized configuration proposed for the implementation of adaptive PA operation is shown in Figure 1. Power supplies provide the DC bias voltages required for normal operation in the on mode. For adaptive operation, selected DC bias voltage control is employed to rapidly switch off the PAS.

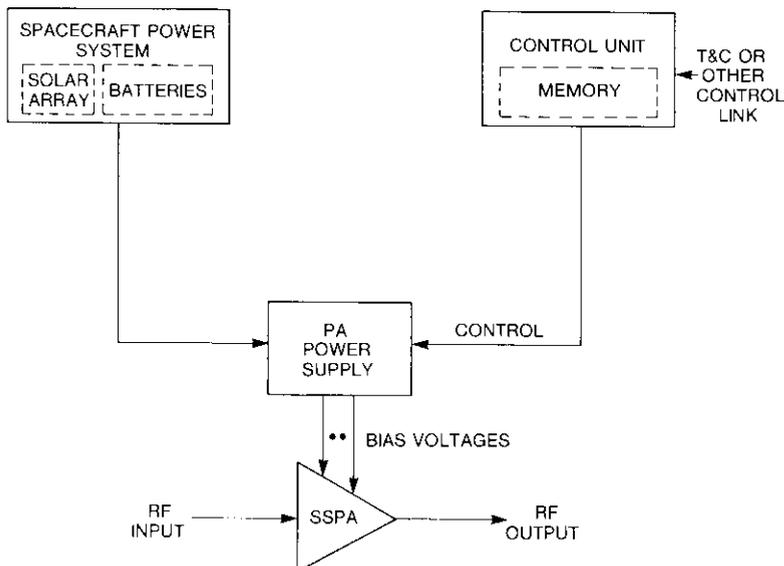


Figure 1. Generalized System Configuration for Adaptive SSPA Operation

The control unit contains a memory that accepts and retains information on the exact moments within the TDMA (here understood to generically include SS-TDMA) frame that power amplifiers are to be turned on or off. This information is read into the control unit over the telemetry and command (T&C) link from the central ground control facility or over a link from each of the participating earth stations. With the transition times in memory, the control unit counts off time in the TDMA frame and, at designated moments, sends instructions to the specified power supplies to turn their associated PAS on or off. In general, each PA will have a different on-off plan, and power consumption will vary as the on-off pattern is updated to correspond to traffic changes.

If the power requirement varies over the course of the day, battery storage can provide a function beyond the conventional one of sustaining operation during eclipse. Batteries would permit energy to be stored during periods of light traffic and used during periods of peak traffic, so that solar arrays can be designed more for average demand than for peak demand.

In satellite communications, traffic variations can be decomposed into several components with different origins and characteristics. In the INTELSAT system, the traffic growth component has been exponential, with an annual rate of around 12 to 14 percent. Another significant variation in traffic that a satellite experiences is the diurnal variation from hour to hour over the course of the day. For example, night or early morning traffic can be expected to be much less than midday traffic. This diurnal variation is usually a smooth curve, based on calls aggregated over segments of time on the order of an hour.

Another component of traffic variation superimposed on the diurnal variation is in the form of rapid fluctuations about a mean value. These variations are caused by the random nature of call arrivals and departures, and occur over a scale of minutes or seconds, during which diurnal characteristics are essentially constant.

On a finer time scale, another component of variation corresponding to speech activity exists. During a short period, on the order of a millisecond, the number of calls in progress will not change substantially; however, the speech activity on those calls may vary significantly during this period.

Adaptive design considerations

INTELSAT TDMA and SS-TDMA plans are based on the fixed assignment of channels defined by a burst time plan (BTP) that specifies the location of bursts within the frame. The BTP is altered only occasionally (e.g., every

several months), when new links are established or old ones are expanded. The simplest form of adaptive operation is to turn off the PAS between bursts (Figure 2). This type of adaptive operation could be controlled by a unit similar to the digital control unit (DCU) already included in the INTELSAT VI design. In fact, the DCU in its presently planned or modified form could be used to provide both microwave switch matrix (MSM) and PA control. The control unit would store on-off instructions corresponding to the beginning and end of bursts within the frame. These instructions would be loaded into the control unit over the T&C link and modified only when the BTP is changed.

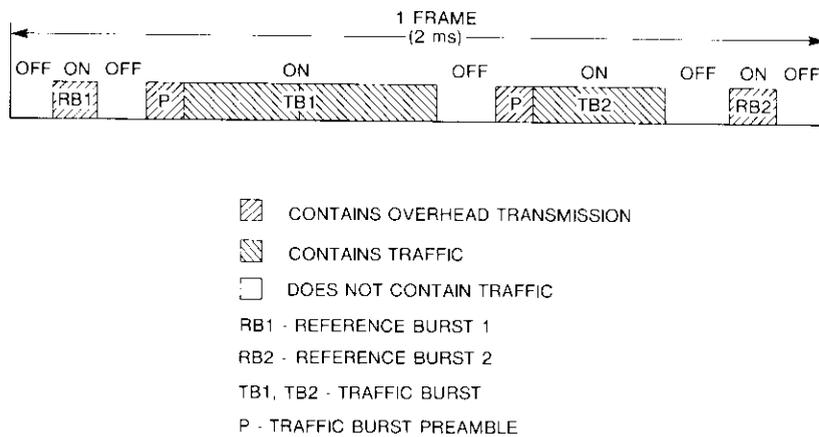


Figure 2. TDMA Frame With PAs Turned Off Between Bursts

A more adaptive operation that responds to the traffic content of bursts can be implemented by turning off PAs within the portions of a burst that do not contain traffic. This approach is facilitated if active channels are consolidated at the front of each subburst, as shown in Figure 3. The procedure is called herding and can be realized with an internal modification to the digital speech interpolation (DSI) unit. The earth station receiver would also require modification to avoid loss of synchronization caused by off periods within a burst.

Control links are critical in adaptive operation that is responsive in real time to traffic variations, since each earth station must measure its own current state of traffic and send this information to the satellite PA control unit. There are two ways to accomplish this. One approach is for each earth station to send information about its traffic to a central ground control facility,

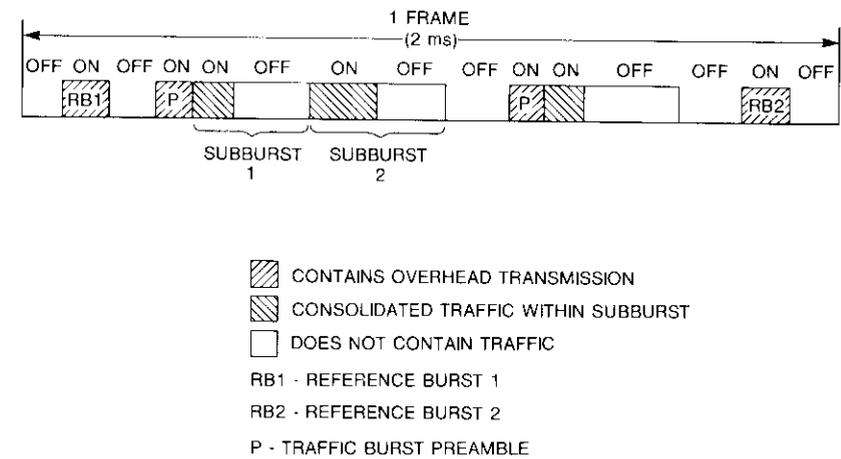


Figure 3. TDMA Frame With PAs Turned Off Between Bursts and at the End of Subbursts Not Containing Traffic

either by terrestrial or satellite link. From there, the information can be sent to the satellite over the T&C or other specialized link. The second approach is for each earth station to send traffic information directly to the satellite; for example, as part of the preamble preceding each burst. This kind of direct control requires on-board demodulation for the control messages, although the traffic itself can be transmitted by a conventional repeater transponder. An advantage of direct control is that it avoids the extra delay that occurs if the control message is relayed through a central ground control facility. In fact, if a control message is transmitted as part of the preamble for each burst, effectively instantaneous response is achieved. This permits complete adaptivity that can follow even variations in speech activity occurring from frame to frame over a 1-ms time scale. However, the implementation of any responsive control link has significant implications for both the earth and space segments. These implications require further investigation.

In contrast to the INTELSAT system, Satellite Business Systems (SBS) employs a demand assignment strategy in which control and traffic are all continuously consolidated at the front of each frame, as illustrated in Figure 4, with the unused resources at the end [2]. Consequently, for adaptive operation, each PA would be turned off during only one period in each frame. Furthermore, the central network control is always aware of the boundary between used and unused resources, so that adaptive control is more easily accommodated.

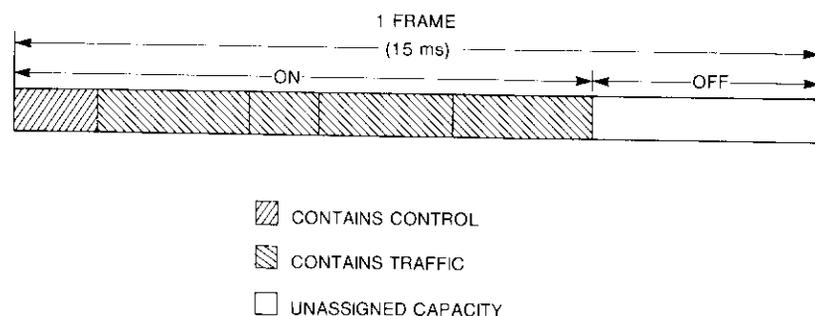


Figure 4. SBS TDMA Frame With PAs Turned Off at End of Burst

In the following treatment, the transient time characteristic of the PA and its switched power supply are assumed to be negligible relative to the time spent in the off state. The design of a switched power supply to achieve efficient operation will not be discussed.

BTP-adaptive operation

Operational plans for a satellite are generally based on the assumption that traffic growth will cause saturation after several years. Consequently, during earlier periods of life, the satellite can be expected to have significant excess capacity that requires power, even though it is unused. For example, if the traffic through a satellite grows at 14 percent per year, the capacity planned for the time of saturation would be only half utilized 5 years before saturation. The excess capacity available over much of the life of the satellite suggests that adaptive operation might be used to reduce power requirements during periods of underutilization, thus making this power available for other uses. Even when a satellite has reached saturation, most of the transponders in TDMA or SS-TDMA operation have large amounts of unused capacity. This occurs when the distribution of traffic among beams causes several transponders to saturate, while most of the transponders support much less than their full capacity. In addition, because of unavoidable inefficiencies in frame filling, even those transponders that have been declared saturated are typically only 90-percent full.

The adaptive approach considered in this section takes advantage of partial transponder filling by turning off each SSPA supporting a TDMA or SS-TDMA link for portions of the frame not occupied in the BTP. Whenever a new BTP is implemented to accommodate growth or other changes in traffic, PA operation is also modified to correspond to the new burst positions.

Power requirements at saturation

Estimates of potential power reduction for adaptive operation have been based on projected INTELSAT VI loading plans for three AOR roles and two IOR roles. The loading plans have been taken to a level of detail which estimates for each TDMA or SS-TDMA transponder the number of voice channels or half-circuits, the number of bursts, and an average DSI gain.

The fraction of a frame used, F , for a given transponder is determined principally by the number of satellite channels supported by the transponder. In addition to the portions of the frame directly devoted to traffic, PAs are assumed to be on for all overhead functions, including preambles, guard times, reference bursts, and the reference station window.

Assuming a switching efficiency of 100 percent, the power reduction for adaptive operation corresponds to the unused portion of the frame, $1 - F$. The peak power consumption is calculated by assuming 85-percent prime-power-to-regulated-DC efficiency and 26-percent DC-to-RF efficiency for SSPAs. On this basis, the potential time-averaged power reduction has been calculated for each transponder to show the power reduction that would be achieved if all hemi and zone beams employed SSPAs for channels (1'-2'), (1-2), and (3-4). For the AOR primary case, Table 1 shows the value of F and the power reduction for each TDMA and SS-TDMA transponder. The value of F averaged over all such transponders and the total PA power reduction are summarized for five roles in Table 2. These results are based on the assumption that the reference station window contains 5,000 symbols (83 μ s), or about 4 percent of the frame.

INTELSAT VI has a relatively small number of transponders (14 or less) devoted to TDMA and SS-TDMA. In the longer term, the use of TDMA and SS-TDMA will increase and dominate the system. Furthermore, SSPAs can be expected to eventually replace TWTAs at both 4 GHz and 11-12 GHz for applications up to a 20-W RF output, and perhaps beyond. As a result, the power reduction that can be achieved with the simple adaptive operation described in this section will represent an increasingly significant portion of the total spacecraft requirement.

Power requirements over the life of the satellite

Power consumption can be divided into the following four categories:

- a. spacecraft subsystems that are not part of the communications payload (including the T&C),
- b. communications subsystems excluding PAs,
- c. PAs that do not support TDMA or SS-TDMA down-links, and
- d. PAs that support TDMA or SS-TDMA down-links.

TABLE 1. INTELSAT VI TRANSPONDER FRACTIONAL USAGE AND POWER REDUCTION FOR AN AOR PRIMARY LOADING PLAN AT SATURATION IN 1993

TRANSPONDER NUMBER	DOWN-LINK BEAM	RF CHANNEL	NUMBER OF VOICE CHANNELS	NUMBER OF BURSTS	FRACTIONAL USAGE	PEAK POWER CONSUMPTION (W)	POWER REDUCTION (W)
10 ^a	WH	1-2'	3,625	4	0.893	79.5	8.5
11	WH	1-2	2,884	3	0.815	79.5	14.7
12	WH	3-4	3,070	4	0.866	79.5	10.7
20 ^a	EH	1-2'	3,625	1	0.885	65.9	7.6
21	EH	1-2	2,041	6	0.613	65.9	25.5
22	EH	3-4	1,388	9	0.445	65.9	36.6
41	Z1	1-2	2,998	1	0.839	10.0	1.6
42	Z1	3-4	3,117	4	0.879	10.0	1.2
51	Z3	1-2	2,504	9	0.745	13.6	3.5
52	Z3	3-4	2,136	12	0.668	13.6	4.5
91	Z2	1-2	2,358	8	0.703	45.5	13.5
92	Z2	3-4	3,117	1	0.870	45.5	5.9
101	Z4	1-2	2,299	13	0.715	45.5	13.0
102	Z4	3-4	2,136	10	0.657	45.5	15.6
					0.757 ^b		162.4 ^c

^a New band channel with conventional TDMA.
^b Fractional usage averaged over all transponders.
^c Total power reduction.

TABLE 2. FRACTIONAL USAGE AND POWER REDUCTION FOR DIFFERENT INTELSAT VI ROLES AT SATURATION

ROLE	NUMBER OF ADAPTIVE TRANSPONDERS	F AVERAGED OVER ALL TRANSPONDERS	TOTAL PA POWER REDUCTION (W)
AOR Primary, 1993	14	0.757	162
AOR Major Path 1, 1993	14	0.729	201
AOR Major Path 2, 1993	14	0.766	163
IOR Primary, 1994	10	0.676	220
IOR Major Path, 1994	10	0.541	207

To provide a representative example of breakdown, the percent of power devoted to (a) and (b), and to the sum of (c) and (d), have been chosen to be those for INTELSAT VI with SSPAS throughout. The breakdown of power between (c) and (d) will shift toward (d) as TDMA and SS-TDMA increasingly dominate the system. To show the implications for spacecraft in the post-INTELSAT VI era, 20 percent of PA power consumption has been assigned to (c) and 80 percent to (d). These components are expressed in Figure 5, which

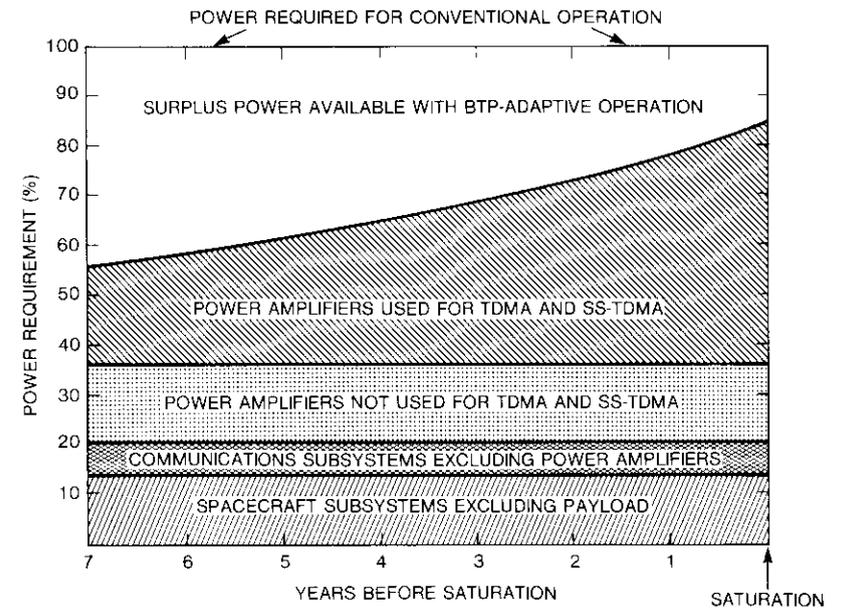


Figure 5. Components of Spacecraft Power Consumption

shows the power reduction that can be achieved with operation that adapts to traffic growth over a period of years.

Among the categories of power consumption, only (d) can benefit from adaptive operation. Adaptive operation is shown to reduce category (d) by 24 percent (25 - 1 percent for control overhead) at saturation, consistent with the projected loading plans for INTELSAT VI AOR roles. It is assumed here that the fill factors calculated for INTELSAT VI loading are not peculiar to that spacecraft, but are generally representative of TDMA and SS-TDMA in the INTELSAT system. Once the power consumption at saturation has been established, requirements for earlier years are calculated based on 14 percent per year traffic growth. The results given in the figure show that the total spacecraft power requirement is reduced by 16 percent at saturation and by a much larger amount in earlier years (for example, 39 percent 5 years before saturation).

The factors described in Figure 5 are applied to a realistic scenario in Figure 6. The satellite is assumed to be an INTELSAT VII (conventional power requirement ~2.1 kW) launched post-1993 into an AOR role. Traffic grows until it reaches saturation at the end of 5 years, when it is moved to an IOR

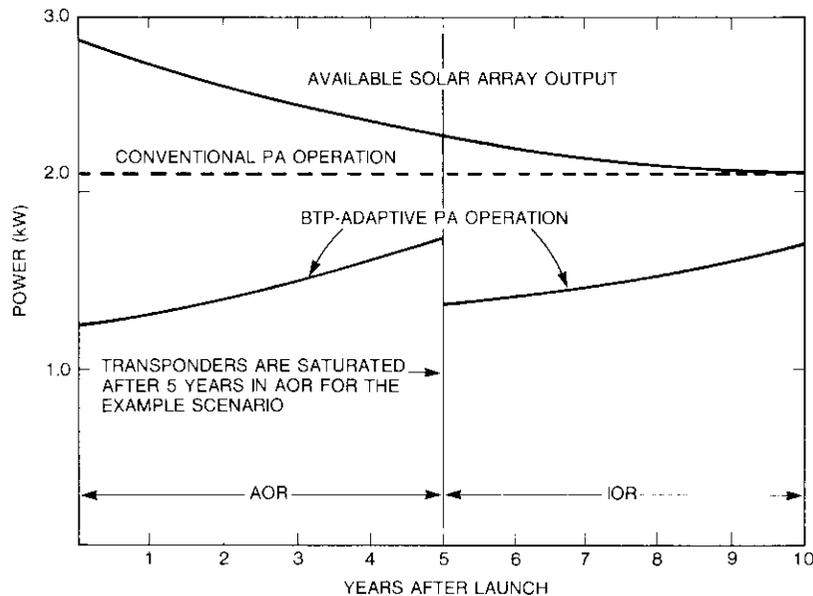


Figure 6. Comparison of Power Requirements for Conventional Operation and Adaptive Operation in an Example Scenario

role. The difference between the available solar array output and the power requirement for adaptive operation is the surplus power available for additional functions or payload packages (such as experimental and maritime), provided the associated increase in mass can be accommodated. The surplus power available for such an added function is 1,000 W for the first 3 years, 600 W for the first 8 years, and 400 W for the full 10 years.

Diurnally adaptive operation

The diurnal cycle related to fluctuation in human activity over a 24-hour day produces a recognized variation in telephone traffic. For local calls, traffic follows a pattern similar to that shown in Figure 7.

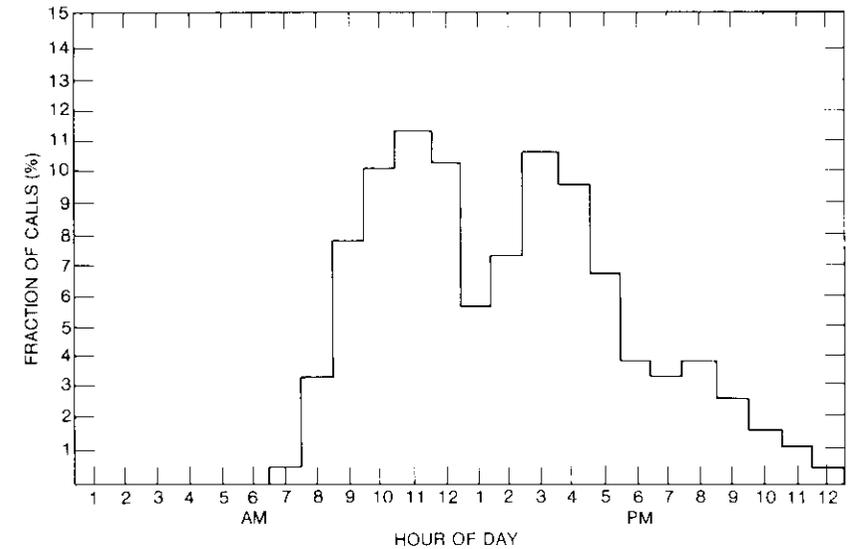


Figure 7. Traffic Density Aggregated in 1-Hour Intervals

The variation in traffic for calls between points in different time zones can be expected to be influenced by the human activity cycles at the ends of the link, which are offset from each other by the time zone difference. Direct measurements of this variation, with the time zone difference as a controlled parameter, do not appear to be available. In the absence of data on traffic between separated points, a model similar to one proposed by Palmer [3] has been used to relate traffic between separated points to available data dominated by local traffic. The model is constructed so that predicted traffic

is proportional to the product of activity functions at each end of the link.

Satellites in the INTELSAT system support traffic between many pairs of time zones, each pair with its own characteristic diurnal variation. The total traffic, which is a composite of these contributions, has its own diurnal variation with a peak during the day and reduced levels at night. Since the peaks of the time-zone-to-time-zone components do not all coincide, the diurnal peak of the total satellite traffic will be less than the sum of the component peaks.

The INTELSAT system is based on fixed assignment of circuit resources to earth stations, rather than on demand assignment. To provide adequate service at all times, this strategy must allocate enough circuits to meet the service criteria (e.g., blocking probability $B = 0.01$) during peak load. At other times of the day, surplus circuits go unused. The satellite design requires that all channels in the system be supported at all times. Effectively, this corresponds to the sum of worst-case traffic for all earth stations. Adaptive operation can take advantage of two traffic characteristics: the worst-case traffic for a given earth station is localized during the day, and the worst-case traffic does not occur at the same time for all earth stations.

The diurnal variation in total TDMA and SS-TDMA traffic through an AOR satellite at 335.3°E has been determined for a projected AOR scenario in the 1995 to 1999 time frame (Figure 8). The peak of the diurnal traffic is 8 percent less than the sum of the traffic peaks for the various time-zone-to-time-zone components, and is shown as a constant line for comparison.

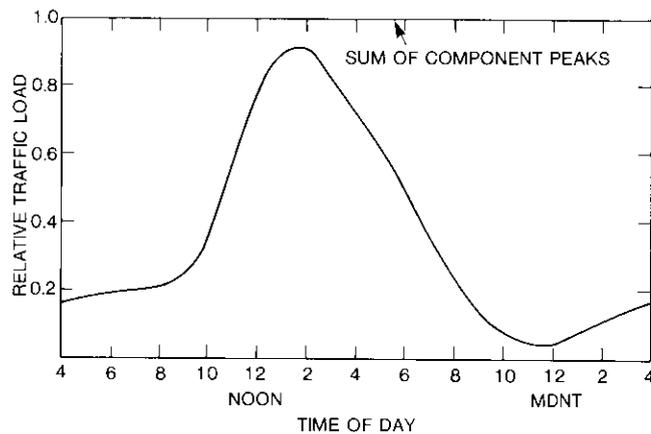


Figure 8. Diurnal Variation of Total Traffic Through an AOR Primary Satellite Based on Example Distribution of Traffic for 1995-1999

The power reduction achieved in a real system will not be perfectly proportional to reduced traffic levels because of decreased efficiency as the number of channels comprising each link decreases [4]. The network served by a satellite will consist of a distribution of link sizes; however, large links of 200 or more circuits can be expected to dominate efficiency statistics. Figure 9 shows the power requirements for networks that have an effective average number of circuits per link of 200 and compares these results to a hypothetical power reduction perfectly proportional to traffic level.

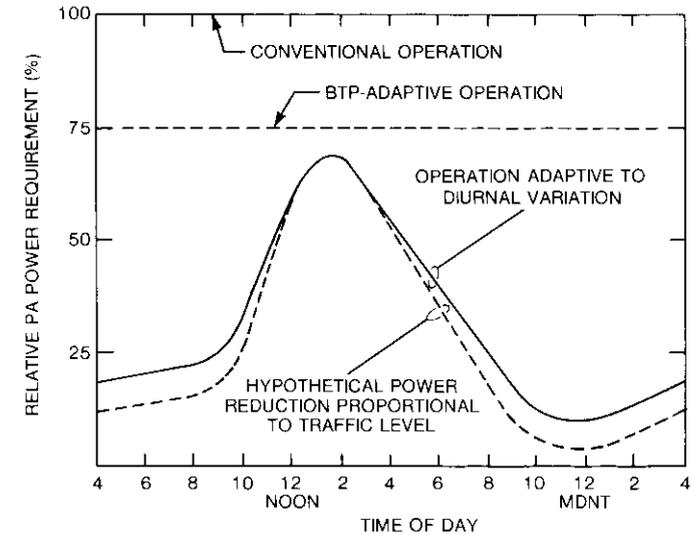


Figure 9. Relative PA Power Requirement for Adaptive Operation Following Diurnal Variation

The batteries on board the spacecraft permit energy to be stored during periods of reduced traffic and used to supplement the solar array output during periods of increased traffic. This is illustrated in Figure 10 by shaded areas A_1 and A_2 , which are proportional respectively to the energy stored by the batteries and that supplied by the batteries. The conditions necessary for such battery-supplemented operation are adequate solar array output ($A_1 \geq A_2$) and adequate battery storage capacity.

Solar array output and battery storage capacity can be traded off against each other to optimize spacecraft design. Results for diurnally-adaptive operation are based on minimizing the solar array. Batteries are chosen to be large enough so that no solar array output is lost; energy is either

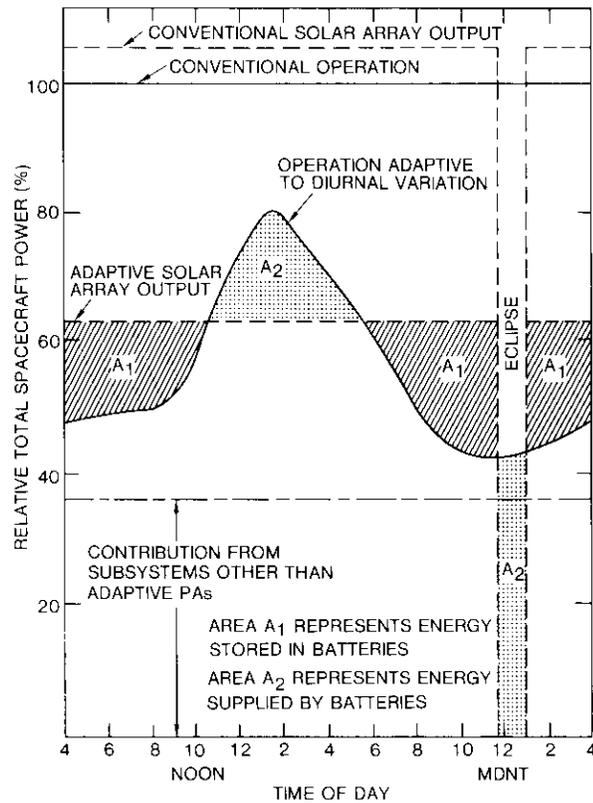


Figure 10. Benefit of Battery Storage in Adaptive Operation

immediately used by spacecraft systems or stored for later use. This corresponds to the condition $A_1 = A_2$. Equivalently, if $P_{SA}(t)$ is the power supplied by the solar array and $P_{SC}(t)$ is the power required by the spacecraft, then

$$\int_0^{24 \text{ hr}} [P_{SA}(\tau) - P_{SC}(\tau)] d\tau = 0, \quad (1)$$

which determines the minimum solar array size that can be achieved. The battery storage capacity required so that no energy is lost is given by

$$B_C = \max_{0 \leq t \leq 24 \text{ hr}} \int_0^t [P_{SA}(\tau) - P_{SC}(\tau)] d\tau \quad (2)$$

For this case, the power requirement for adaptive SSPAs is 38 percent of the conventional requirement. This corresponds to a total spacecraft power requirement that is 60 percent of conventional, assuming that 64 percent of the power in a conventional design would be assigned to TDMA and SS-TDMA links supported by SSPAs. As a result, the solar array output required for adaptive operation is 57 percent of that required for conventional operation. Furthermore, even the battery storage capacity has been reduced to 92 percent of that used for conventional operation. These reductions were obtained through a design approach that minimizes solar array requirements.

The analysis presented here has been based on a specific diurnal variation. In fact, peak load, and the diurnal variation in general, can vary somewhat from day to day for a given link [5]. Since a given satellite supports many such links, the law of large numbers suggests a much smaller day-to-day variation in total satellite traffic. Furthermore, battery storage can average over day-to-day changes in the diurnal variation.

The adaptive performance that has been discussed is based on operation that adjusts at intervals of 1 hour or less. Adaptive operation that is less responsive, for example, with adjustment intervals of several hours, will sacrifice efficiency in following traffic variations for some simplification. A particular form of adaptive operation has been suggested by Welti [6] to take advantage of reduced power requirements during eclipse. In a conventional design, batteries must sustain the full payload during eclipse, even though this occurs at night when traffic levels are greatly reduced. An adaptive approach to reduce this design requirement is to employ a reduced but adequate number of channels during eclipse. At all other times, the system operates in a conventional or a BTP-adaptive mode. This form of operation could allow the battery capacity required to sustain operation during eclipse to be reduced to less than half of the conventional requirement.

Random call variations

The diurnal variation appears relatively smooth when it is based on traffic aggregated over intervals of 1 hour or $\frac{1}{2}$ hour. However, superimposed on this smooth variation is a more rapid fluctuation caused by the random nature of user arrivals and departures. The number of calls in progress can fluctuate significantly over a period of seconds or minutes. In this section, more adaptive operation is considered in which the number of channels supported is actually the number of users on the system. The power consumed changes with random traffic fluctuations. Consequently, the power requirement smoothed by battery storage corresponds to the average over these fluctuations, rather than to the peak traffic in the interval.

The number of terrestrial channels in use on a link between two earth

stations (corresponding to a portion of the TDMA frame) at a given time is denoted by $n_T(t)$. If N_T terrestrial channels are supported during a given interval, service can be provided to any new user as long as $n_T(t) < N_T$. However, if $n_T(t) = N_T$, any new user requesting service is turned away, or blocked, in a blocked calls cleared (BCC) discipline. The probability of blocking is [7]

$$B = \frac{a^{N_T}/N_T!}{\sum_{k=0}^{N_T} a^k/k!} \quad (3)$$

where $a = \lambda\mu^{-1}$

λ = average rate of user arrivals (calls per second)

μ^{-1} = average call duration (seconds).

For voice service based on DSI, N_s satellite channels are required to support N_T terrestrial channels, related through the processing gain by $N_s = N_T/G_0$. The effective processing gain, G_0 , which is a function of the number of channels in a subburst, is influenced as N_T changes; however, its exact value depends on the details of the BTP. The N_T link channels can be spread among several subbursts, not filling all of them, or they can share a subburst with channels for other links to different destinations.

Adaptive operation that follows the random fluctuations in $n_T(t)$ must support $n_s(t)$ satellite channels, related by

$$n_s(t) = \frac{n_T(t)}{G(t)}$$

where the processing gain, $G(t)$, is influenced by $n_T(t)$. The power reduction that results from such adaptive operation is

$$\Delta P(t) = [P_{N_s} - P_0] \left[1 - \frac{n_s(t)}{N_s} \right]$$

where P_0 = power required for synchronization, orderwires, and other overhead

P_{N_s} = power required to support N_s satellite channels including overhead ($P_N \geq P_0$).

Thus, the time-averaged power reduction achieved, normalized to P_{N_s} , can be written as

$$\begin{aligned} \Delta p_{RC} &= \frac{\langle \Delta P(t) \rangle}{P_{N_s}} \\ &= \left[1 - \frac{P_0}{P_{N_s}} \right] \left[1 - \gamma \frac{\langle n_T(t) \rangle}{N_T} \right] \end{aligned} \quad (4)$$

where $\gamma = G_0 \langle G^{-1}(t) \rangle$.

If the system is in a steady state, the time-averaged value $\langle n_T(t) \rangle$ can be identified as the expected value $E[n_T]$. For a BCC discipline [7],

$$E[n_T] = a[1 - B(N_T, a)] \quad (5)$$

If B is fixed, the relative power reduction for adaptive operation, given by equation (4), can be related to N_T and the traffic load, a , through equation (5). Furthermore, the functional relationship given by equation (3) can be formally rewritten to express the variable a as a function of B and N_T . The result for the relative power reduction is

$$\Delta p_{RC} = \left[1 - \frac{P_0}{P_{N_s}} \right] \left[1 - \frac{\gamma(1-B)}{N_T} a(B, N_T) \right] \quad (6)$$

For the applications considered here, $G(t)$ decreases only slightly below G_0 as traffic fluctuates, so that γ can be assumed to lie in the range $0.96 < \gamma < 1.0$. Consequently, the representative value $\gamma = 0.98$ can be chosen with negligible effect on accuracy.

The relative power reduction, Δp_{RC} , is shown in Figure 11 as a function of N_T for $B = 0.1, 0.01, \text{ and } 0.001$. The fractional overhead, P_0/P_{N_s} , has been set to zero to provide reference curves; however, these results can be simply adjusted by the factor $(1 - P_0/P_{N_s})$ appearing in equation (6) to describe the effect of any overhead on Δp_{RC} . At peak traffic load, the fractional overhead averaged over all transponders on a satellite, can be expected to lie between 0.06 and 0.10, based on projected loading plans for INTELSAT VI.

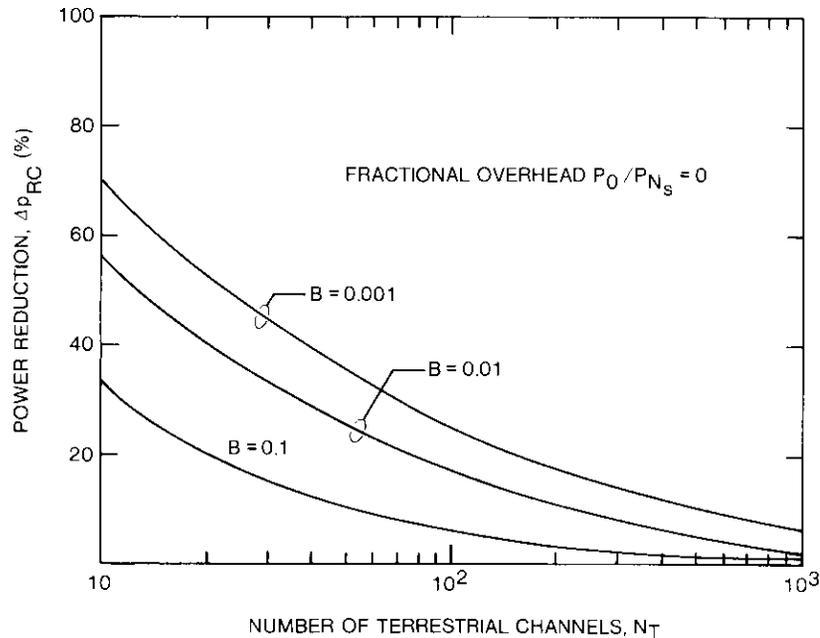


Figure 11. Relative Power Reduction for Operation Adaptive to Random Fluctuations

DSI activity variation

Over intervals of several tenths of a second, the number of calls in progress on a link between two earth stations will remain essentially constant. However, among a group of 200 terrestrial channels, a change in speech activity can be expected to occur every few milliseconds.

Although satellite channels are provided for roughly 50 percent of the terrestrial channels, on the average only 40 percent of the terrestrial channels contain speech activity at a given moment. Equivalently, 20 percent of the satellite channels are inactive, and a corresponding reduction in power is possible if adaptive operation is responsive enough to follow speech activity.

An accurate estimate of power reduction must be based on the statistics of speech activity [8]. For n_T terrestrial channels routed to a given DSI unit, the expected number of active channels is αn_T . If adaptive operation is sufficiently responsive that power can be turned on only for active channels, the time-averaged power consumed, disregarding overhead, will be proportional to αn_T . In contrast, adaptive operation that follows random call variations will require power proportional to n_s . Then the relative power

reduction for the DSI unit under consideration, excluding overhead, is

$$\Delta p_{SA} = \frac{n_s - \alpha n_T}{n_s} \quad (7)$$

For DSI based on a freeze-out probability of 0.005, the binomial distribution can be approximated as Gaussian, so that the expression

$$n_s \approx n_T \alpha + 2.575 \sqrt{n_T \alpha (1 - \alpha)}$$

can be substituted in equation (7) to give

$$\Delta p_{SA} = \frac{2.575 \sqrt{1 - \alpha}}{\sqrt{\alpha n_T + 2.575 \sqrt{1 - \alpha}}} \quad (8)$$

The relative power reduction, Δp_{SA} , for a DSI unit, given by equation (8), is shown in Figure 12 as a function of n_T . The value of n_T will vary for each

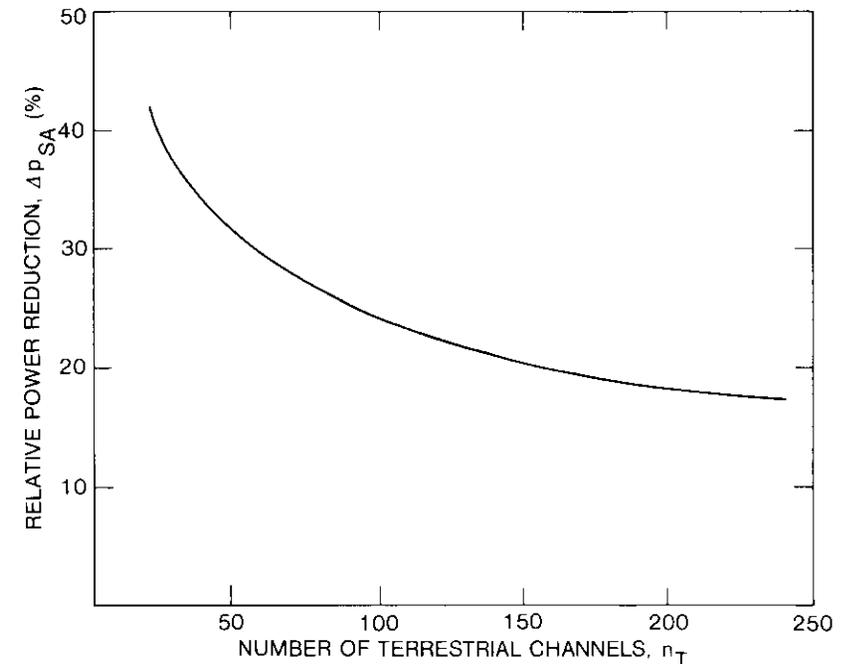


Figure 12. Relative Power Reduction in a Subburst for Adaptive Operation Following Speech Activity

DSI unit or subburst, but in general will tend to be larger (~ 200) at peak load and smaller as traffic decreases. The total power reduction for a transponder or a satellite will correspond to the aggregate power reductions of all subbursts. However, since overhead remains constant, results for the relative power reduction must be adjusted by the factor $(1 - P_0/P_{N_3})$, as in the case of random call variations.

The adaptive operation considered in this section requires extremely responsive control links. The only suitable candidate for this function is the direct approach previously described. The preamble of each burst contains bits that designate when the component subburst is turned off. These control bits arrive at the satellite in the same frame with the traffic they describe. The satellite demodulates the control bits on board and operates the PAS accordingly.

Comparison of adaptive operation

Adaptive operation that is responsive to the random call component reduces PA power requirements below that for diurnally adaptive operation, according to equation (6). Even more responsive operation that follows speech activity achieves the further reduction described by equation (8). The PA requirement for each class of progressively more adaptive operation that has been considered is compared to that for conventional operation in Figure 13.

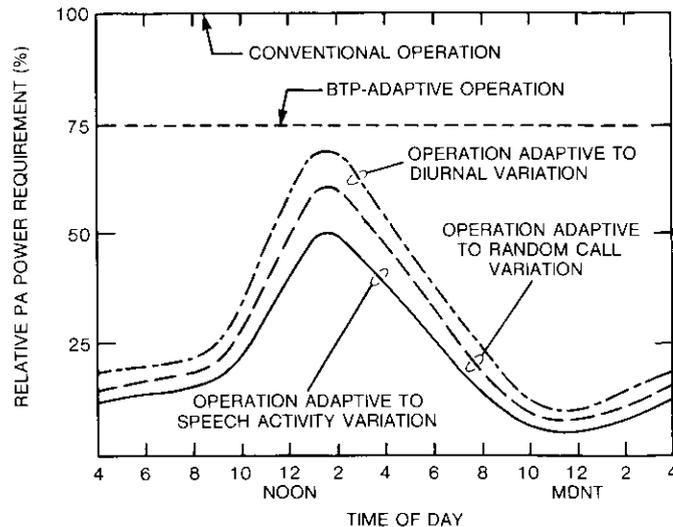


Figure 13. Relative PA Power Requirement for Adaptive Operation Following Speech Activity Variation

In BTP-adaptive operation, spacecraft power requirements are constant over the course of a day, as in conventional operation. The principal function of battery storage is to sustain operation during eclipse. In adaptive operation which follows variations that occur during the day (diurnal, random call, and speech activity), batteries play an essential role in storing energy during periods of light traffic to be used during periods of heavy traffic. This can permit solar arrays to be sized closer to the average power requirement rather than to the peak power requirement. Tradeoffs in spacecraft electrical power system design involving solar array size and battery size have not been investigated to optimize total spacecraft design. Instead, the design has been chosen so that battery storage is adequate to minimize solar array output. In this analysis, battery storage capacity is simply treated as well-defined energy corresponding to a specified maximum depth of discharge. A more involved analysis could recognize a statistical uncertainty in the diurnal variation and consider its effect on the depth of discharge.

Results based on the representative INTELSAT VII scenario are summarized in Table 3 for progressively more adaptive operation. These results include minimal overhead to describe on-board control. However, on-board demodulation of control bits will be required to implement direct control of the more adaptive types of operation (speech activity, random call, and perhaps diurnal). If the satellite is based on transponder repeaters, an additional overhead of power and mass not included in the table is required for direct control. However, for advanced designs that demodulate communications channels on board the satellite, direct control of PA operation is more integral to the communications system design.

TABLE 3. ADAPTIVE SYSTEM CHARACTERISTICS AT SATURATION FOR REPRESENTATIVE INTELSAT VII SCENARIO RELATIVE TO CONVENTIONAL SYSTEM

CHARACTERISTIC	DEGREE OF ADAPTIVITY			
	BTP	DIURNAL	RANDOM CALL	SPEECH ACTIVITY
SSPA Power Requirement (%)	76	38	31	26
Spacecraft Power Requirement (%)	85	60	56	53
Solar Array Output (%)	85	57	53	50
Battery Storage Capacity (%)	85	92	74	67

Conclusions

Four categories of traffic variation, differentiated by time scale, have been identified: growth (months/years), diurnal (hours), random call (minutes/seconds), and speech activity (milliseconds). These variations occur over progressively finer time scales, and adaptive operation that is more responsive will be able to follow increasingly rapid variations to achieve greater power reduction.

A simple form of adaptive operation that follows the growth component of traffic (BTP-adaptive) can be implemented with relatively modest modifications. No changes to the earth segment are required. More adaptive operation for future spacecraft will require greater system complexity but will permit further reduction in power requirements.

Results for BTP-adaptive operation have been based on projected transponder loading plans at saturation for INTELSAT VI and on a traffic growth rate of 14 percent. Estimates for operation that is adaptive to variations in channel use during the day have been based on a simplified description of traffic flow between different time zones. A representative example of projected 1995–1999 traffic distribution for an AOR primary satellite has been assumed. As detailed and accurate traffic projections become available, estimates for adaptive power requirements can be refined and extended to a wider variety of scenarios. Nevertheless, the results presented here indicate the potential for substantial power reductions. This provides a motivation for further study of the benefits of adaptive operation and the technical issues of implementation.

Acknowledgment

The author wishes to acknowledge the valuable contributions of A. E. Atia, W. J. Billerbeck, K. A. Chahine, J. Hahn, and N. M. Jacobus. The helpful comments from many other individuals are also gratefully acknowledged.

References

- [1] W. Billerbeck and M. A. Hulley, private communication.
- [2] W. H. Curry, Jr., "SBS System Evolution," *COMSAT Technical Review*, Vol. 11, No. 2, Fall 1981, pp. 267–291.
- [3] L. Palmer, private communication.
- [4] G. D. Dill and G. D. Gordon, "Efficient Computation of Erlang Loss Functions," *COMSAT Technical Review*, Vol. 8, No. 2, Fall 1978, pp. 353–370.
- [5] H. Kashper, S. M. Rocklin, and C. R. Szlag, "Effects of Day-to-Day Load Variation on Trunk Group Blocking," *Bell System Technical Journal*, Vol. 61, No. 2, February 1982, pp. 123–135.

- [6] G. R. Welti, "Some Design Concepts for INTELSAT VII," Sixth International Conference on Digital Satellite Communications, Phoenix, September 1983.
- [7] R. Syski, *Introduction to Congestion Theory in Telephone Systems*, London: Oliver and Boyd, 1960.
- [8] S. J. Campanella, "Digital Speech Interpolation," *COMSAT Technical Review*, Vol. 6, No. 1, Spring 1976, pp. 127–158.

Reuben E. Eaves received a B.E.S. from The Johns Hopkins University in 1964 and an Sc.M. and Ph.D. from Brown University in 1966 and 1969, all in electrical engineering. He is currently Manager of Satellite Design for INTELSAT Technical Services at COMSAT, with responsibility for system analysis and design studies. Prior to joining COMSAT, from 1976 to 1981, Dr. Eaves was at Lincoln Laboratory, Massachusetts Institute of Technology, where he was involved in the development of advanced satellite systems and concepts. From 1968 to 1976, he was at the NASA Electronics Research Center and the Transportation Systems Center, and concurrently held a visiting appointment at Brown University. Dr. Eaves is a member of Sigma Xi, Tau Beta Pi, Eta Kappa Nu, IEEE, and AIAA.



A computer program for communications channel modeling and simulation

A. HAMID, S. R. BAKER, AND W. COOK

(Manuscript received March 28, 1983)

Abstract

The Channel Modeling Program (CMP) is a general-purpose simulation program especially suited for simulation of satellite communications systems. This paper describes the capabilities of CMP and discusses the theory and concepts used in the implementation of the program. An illustrative example of CMP use is also given.

CMP is an effective tool for analyzing and comparing various communications channel configurations. A randomly modulated signal is first generated, which then may pass through various system elements, such as modulators/demodulators, nonlinear amplifiers, filters, and digital receivers. CMP calculates the average values of steady-state performance measures, including bit-error rates and signal-to-noise ratios. In addition, graphical outputs such as signal envelopes, cyc patterns, scatter diagrams, phase-plane trajectories, and power spectral density plots provide qualitative measures of the system performance.

Introduction

Computer simulation provides an economical means of estimating the actual performance of communications systems. The primary advantage of computer simulation is its flexibility; *i.e.*, once the simulation software has

been developed, a variety of configurations can be analyzed. Therefore, many computer programs have been developed for computer simulation of satellite communications systems since the late 1960s [1]–[14].

Because of the complexity of a satellite communications system, most simulation programs have been oriented toward specific studies, and are either too restrictive or cumbersome for general use. A simulation program for present-day satellite communications systems must be general enough to handle multiple signals having arbitrary characteristics. It must accommodate state-of-the-art modulation techniques and treat co-channel and adjacent-channel interference. Furthermore, a systems engineer with little knowledge of software should be able to use the program in design studies.

The development of channel modeling techniques has been an ongoing effort at COMSAT for several years. The Channel Modeling Program (CMP) represents the accumulated results of this effort. The program is written to allow the user maximum flexibility in specifying the characteristics of the communications link while eliminating the necessity of understanding the simulation process. Although the program is primarily intended for analysis of signal degradation in a satellite transmission link, the implementation can easily be generalized to other communications applications.

Early applications of CMP involved the comparison of digital modulation techniques, including phase-shift-keyed (PSK) modulation. Recently, the scope of the program has been expanded to areas for which no acceptable results were previously obtainable. One such area is the analysis of performance for configurations involving combinations of analog (FDM/FM) signals and digital signals. Another is the calculation of bit-error rate (BER) degradation due to imperfect synchronization.

Program features

CMP is specifically designed to run under VM/CMS (virtual machine/conversational monitor system) on IBM 370 series computers. An on-line graphics capability is available for a number of graphics devices, providing an effective tool for interactive system design. Some of the major program features are summarized below.

User friendly interface

The user communicates with the program by means of a free-field command language. The particular channel to be analyzed is specified by a list of commands describing signals, the characteristics of various link components,

and the nature of the analysis to be performed. The CMP commands are divided into the following categories: control, signal generation, modulation, filter, signal analysis, and output.

Tables 1 through 6 list the functions of commands in each category.

TABLE 1. CMP CONTROL COMMANDS

NAME	FUNCTION
END	Terminate the execution of the program
IDENTIFY	Assign an identification number to a run
INITIALIZE	Initialize number of samples and sample rate
LOOP	Loop over a set of commands
PULSE	Define a pulse shape to be used in generating PSK signals
REWIND	Rewind an auxiliary input data set
RETRIEVE	Retrieve a saved signal
SAVE	Save a signal
SWITCH	Switch control from the terminal to an auxiliary input data set and vice versa
VARY	Vary a parameter value in a loop

TABLE 2. CMP SIGNAL GENERATION COMMANDS

NAME	FUNCTION
FDM	Generate an approximate FDM telephony signal
FDM1	Generate an FDM signal of specified spectrum
NOISE	Generate a random noise signal
PSK1	Generate a random PSK signal
PSK2	Generate a pseudorandom PSK signal
PSK3	Generate a periodic PSK signal
PSK8	Generate a pseudorandom 8-phase PSK signal
TONES	Generate a summation of sinusoidal signals
MSK	Generate a minimum-shift-keyed (MSK) random signal

TABLE 3. CMP MODULATION COMMANDS

NAME	FUNCTION
DEMAM	Demodulate an AM signal
DEMFM	Demodulate an FM signal
DEMPM	Demodulate a PM signal
MODAM	Modulate by using amplitude modulation
MODFM	Modulate by using frequency modulation
MODPM	Modulate by using phase modulation

TABLE 4. CMP FILTER COMMANDS

NAME	FUNCTION
BFLT	Modify the bandwidth of a filter
CFLT	Cascade a filter with the currently defined filter
DFLT	Define a user-specified filter
EFLT	Equalize the currently defined filter
IFLT	Implement the currently defined filter
SFLT	Cascade an ideal $x/(\sin x)$ compensation with the currently defined filter
TFLT	Specify that the currently defined filter is the timing-phase receive filter

TABLE 5. CMP SIGNAL ANALYSIS COMMANDS

NAME	FUNCTION
AMPLIFIER	Simulate a nonlinear amplifier
BER	Calculate and print BERs for BPSK and QPSK signals
BER8	Calculate and print BER for 8-phase PSK signal
CALIBRATE	Calibrate signal for BER computations
CLIP	Simulate clipping effects in a modem
CSET	Reset the power level of the unmodulated carrier
DELAY	Introduce a time delay
DETECT	Detect a PSK signal
EXPRESS	Change the carrier frequency of a signal without changing the signal
JBER	Perform jittered BER calculations
NONLINEAR	Perform a nonlinear scaling on a signal

TABLE 5. CMP SIGNAL ANALYSIS COMMANDS (continued)

NAME	FUNCTION
NORMALIZE	Normalize the signal power to a specified level
OFFSET	Shift the imaginary component of a QPSK signal by one-half the symbol period
POWER	Calculate and print power of a signal
PRODUCT	Multiply two signals
SCALE	Scale a signal
SHIFT	Change the carrier frequency of a signal
SIGN	Perform the sign function on a signal
SPLIT	Split real and imaginary parts of a signal
SUM	Sum two or more signals
TBER	Calculate BER at a number of timing points

TABLE 6. CMP OUTPUT COMMANDS

NAME	FUNCTION
CALCOMP	Designate the Calcomp plotter as the plotting device
EYE	Plot the eye pattern of a PSK signal
HPPLOT	Designate the HP7221 as the plotting device
IBMPLOT	Designate an IBM 3279 terminal as the plotting device
PAMPLIFIER	Plot the currently defined amplifier
PBER	Plot BER curves
PFLT	Plot the currently defined filter
PLOT	Plot the complex envelope of a signal
PRINT	Print sample values of a signal
SCATTER	Plot the scatter diagram of a detected PSK signal
SPECTRUM	Plot the power spectral density of a modulated signal
TEKTRONIX	Designate the Tektronix terminal as the plotting device
TRAJECTORY	Plot the phase-plane trajectory of a modulated signal

CMP may be invoked in either an interactive or an automatic mode. In the interactive mode, the user types a command on the terminal, and the program

executes that command and prompts the user for the next command. The user may also store commonly used sequences of commands on a data file and switch the program control between the terminal and this file. In the automatic mode, all commands are stored on a data file, and the program automatically executes the commands on this file.

Flexibility

The user has complete flexibility in specifying the system configuration, including the number and types of link components, the number of signals sharing a common link, and the number of parallel paths a signal may follow. The program is capable of simulating a wide variety of signals having arbitrary characteristics.

Modularity

CMP is intended to be a research tool to investigate both present and future communications system configurations. Therefore, the program is designed and implemented in a modular form to simplify the task of modifying existing modules or adding new functions.

Efficiency

Time domain simulations can be very time consuming. Therefore, a major consideration in the design of CMP has been to maximize the efficiency of the computations involved. All signals are stored in core-resident arrays rather than on disk data sets. This approach significantly reduces execution time and simplifies the logic of the program. Very long signals can be handled by "looping" through the input commands a specified number of times, while modifying the random signal generator "seed word" for each pass. Throughout the program, efficient numerical algorithms are used to save execution time, and conversions between time and frequency domains are minimized.

Graphical outputs

CMP provides graphical outputs for a number of commonly used graphical devices. These outputs are helpful in verifying the parameters of the individual link elements and provide a qualitative measure of channel performance. Furthermore, these outputs can aid in assessing the validity and accuracy of the simulation model. Examples of CMP outputs are signal envelopes, eye patterns, scatter diagrams, phase-plane trajectories, power spectral density plots, plots of filter and amplifier characteristics, and plots of BER vs C/N_0 .

Model description

CMP models the sources of communications link impairments and calculates the average values of steady-state performance measures. A typical application of the program is the evaluation of steady-state BERS for PSK signals which are degraded by a combination of intersymbol interference and thermal noise on the down-link path of a satellite communications channel. The following subsections describe various components of the CMP model and the simulation techniques employed.

Signal representation

CMP is designed to handle any type of narrowband bandpass angle- or amplitude-modulated signal. The general form of such a signal is [15]

$$s(t) = \rho(t) \cos [2\pi f_c t + \theta(t)] \quad (1)$$

where f_c is the carrier frequency; and where the amplitude, $\rho(t)$, and phase, $\theta(t)$, characterize the modulation components of the bandpass signal.

A bandpass signal can also be expressed as

$$s(t) = p(t)e^{j\pi f_c t} \quad (2)$$

where

$$p(t) = \rho(t)e^{j\theta(t)} \quad (3)$$

The function $p(t)$ is called the complex envelope of $s(t)$. Thus, a bandpass signal is completely described by its carrier frequency, f_c , and its low-frequency complex envelope, $p(t)$. Furthermore, operations on bandpass signals may be simulated by the equivalent low-pass operations on the complex envelope [15]. The significance of this approach is that a computation that would be complicated by the presence of sinusoidal carrier terms can be reduced to an equivalent, but simpler, low-pass problem.

CMP stores signals in the form of samples of the complex envelope. Processes acting on the signals are modeled as equivalent low-pass processes acting on the complex envelope. Simulations are performed in both the time and frequency domains. The transformation between the two domains is accomplished by using a fast Fourier transform (FFT) algorithm which computes the discrete Fourier transform:

$$p(t_k) = \frac{1}{N} \sum_{i=0}^{N-1} P(f_i) e^{j2\pi k i / N}; \quad k = 0, \dots, N-1 \quad (4)$$

and

$$P(f_k) = \sum_{i=0}^{N-1} p(t_i) e^{-j2\pi k i / N}; \quad k = 0, \dots, N-1 \quad (5)$$

where N is the number of samples. The discrete times, t_k , and the discrete frequencies, f_k , are defined as

$$t_k = k/f_s \quad (6)$$

$$f_k = kf_s/N \quad (7)$$

where f_s is the sampling frequency or sample rate.

Because the signals are represented by discrete samples, no frequencies greater than one-half the sample rate can be represented. The frequency components that can be represented are integer multiples of the sample rate divided by the number of samples. Thus, all frequencies are quantized to this resolution in CMP. In cases for which greater resolution and accuracy are desired, the sample rate must be increased.

CMP simulates both linear and nonlinear processes. Examples of linear processes are amplitude modulation, summation of signals, and filtering; while nonlinear processes include frequency modulation, phase modulation, and nonlinear amplification.

A linear process can be represented exactly in CMP. In general, nonlinear processes cannot be exactly represented, and the output from a nonlinear process is an aliased version of the signal which would have resulted from the actual process. The effect of aliasing can usually be made negligible by using a high enough sample rate. For modulated signals, a useful rule of thumb is to select a sampling frequency high enough that signals are confined to the region $\pm f_s/4$.

Two other parameters are important in the simulation of digital communications signals: N_s , the number of samples per symbol, and R_s , the symbol rate. These parameters are related to the sample rate by

$$f_s = N_s R_s \quad (8)$$

or

$$N_s = \frac{f_s}{R_s} \quad (9)$$

The simulation software is simplified if N_s is selected to be an integer. Since it is necessary to define "half symbols" in some applications, it is

convenient for N_s to be even. Generally, N_s is selected as a power of 2.

A good practice is to pick the symbol rate first and then select the number of samples per symbol to achieve the desired sample rate. Because of the nature of the FFT routine used in CMP, the number of samples, N , must be a power of 2. Since the number of samples per symbol is also a power of 2, this requirement ensures an integer number of symbols in the signal.

In a significant departure from previous versions, the current version of the program stores signals in core-resident arrays, rather than on disk data sets. In addition to increasing the efficiency of the program, this approach avoids the problems associated with "windowing" and significantly simplifies the logic of the program.

Signal generation

CMP can simulate either digital or analog signals. The digital signals are the M -ary PSK signals (such as binary, quadrature and 8-phase PSK signals), MSK signals, and offset QPSK signals. Examples of analog signals are Gaussian noise, summations of sinusoids, and frequency division multiplexed (FDM) baseband signals. The baseband signals can be used to modulate a carrier by using amplitude, frequency, and phase modulation techniques and thus can simulate most common signals, including FM telephony signals.

Digital signals are generated in the time domain. The data sequence can be a pseudorandom sequence,* a random sequence based on a uniform distribution, or a user-specified periodic sequence. CMP then calculates and stores samples of the real and imaginary parts of the complex envelope evaluated at the sample rate. Figure 1 is an example of first 450 samples of an unfiltered QPSK signal generated by CMP.

Analog baseband signals are generated by specifying the frequency spectrum and then using the inverse FFT to generate the time domain signal. In contrast to direct generation of the signal in the time domain, which may introduce aliased spectral components, this approach is exact because only the representable periodic components are included in the time domain signal by the inverse FFT. Figure 2 shows the frequency spectrum of an FDM signal generated by CMP.

If only the amplitudes of the spectral components are specified, CMP assigns a phase to each, based on a uniform distribution on the range $(0, 2\pi)$. The resulting signal has a Gaussian amplitude distribution in time and the specified amplitude distribution in frequency if an adequate number of components are included.

* Pseudorandom means that the binary elements used to modulate the signal come from a maximal-length shift register sequence. The PSK signal thus generated has well-behaved statistical properties.

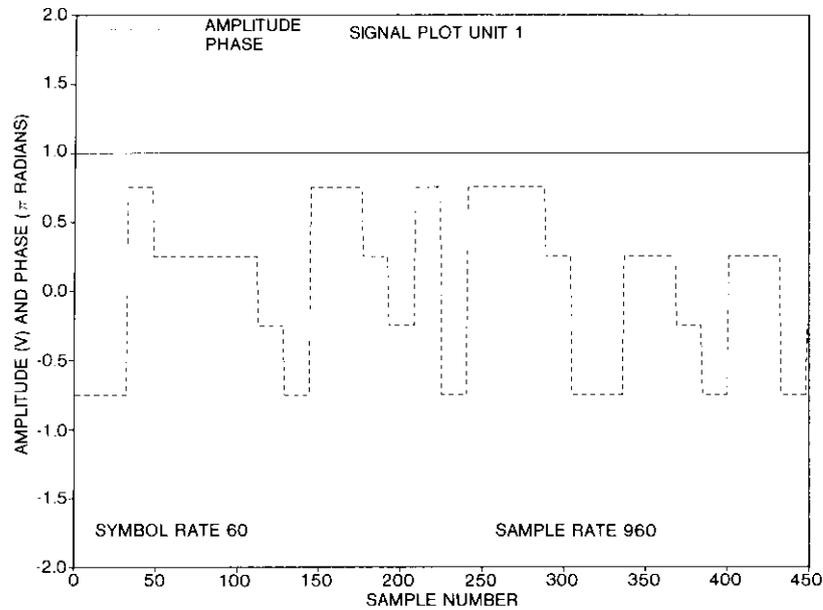


Figure 1. Example of a QPSK Signal Generated by CMP

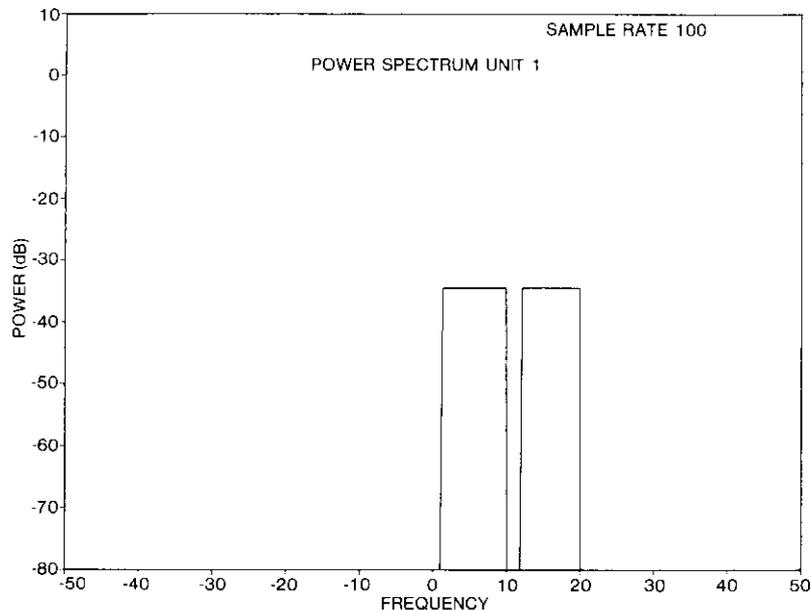


Figure 2. Spectrum of an FDM Signal

Baseband signals can be used to modulate a carrier using amplitude, frequency, or phase modulation. Only the real part of these signals is considered as the modulating waveform. Conversely, only the real part of the signal is recovered in the demodulation process and the imaginary part is set identically to zero. An option is provided for modeling nonlinear distortions in the frequency modulator.

Filters

CMP is capable of simulating a variety of bandpass filters. Currently the following types of filters can be simulated:

- Common generic filters such as Butterworth, Chebychev, inverse Chebychev, and Bessel;
- specialty filters such as Nyquist and raised cosine filters;
- filters defined by pole and zero locations;
- filters defined by their impulse response (e.g., tapped-delay-line filters); and
- filters defined by user-supplied gain and group delay characteristics.

Once a filter is defined, its characteristics can be subsequently modified by cascading it with another filter, ideal equalization, and $(\sin x)/x$ compensation. Thus, the three distinct phases of filter simulation in CMP are definition, modification, and implementation.

FILTER DEFINITION

The filtering process is defined by

$$Y(\omega) = H(\omega)X(\omega) \quad (10)$$

where $H(\omega)$ is the filter frequency response, and $X(\omega)$ and $Y(\omega)$ are the frequency domain representations of $x(t)$ and $y(t)$, respectively.

Frequency responses of generic filters can be described by analytical functions [16]–[20]. For these filters, the locations of poles and zeros are calculated to determine the frequency response,

$$H(\omega) = H_0 \frac{\prod_{i=1}^k (j\omega - z_i)}{\prod_{i=1}^l (j\omega - p_i)} \quad (11)$$

where ω = frequency

z_i = location of i th zero

p_i = location of i th pole

k = number of zeros
 ℓ = number of poles
 H_0 = normalizing factor such that the filter gain is unity at zero frequency.

Nyquist filters have the unique property of causing no intersymbol interference at the center of each symbol, since the impulse response of the filter is identically zero at multiples of the symbol period. This property is a result of the symmetry of the frequency response with respect to half the symbol rate.

One family of filters having this property is the raised cosine n -percent roll-off filter. Its transfer function is defined in terms of the half-amplitude frequency, ω_h , and the roll-off factor, R . The magnitude characteristic is

$$|H(\omega)| = \begin{cases} 1, & \omega \leq (1 - R)\omega_h = \omega_1 \\ \frac{1 + \cos [(\pi/2R)(\omega/\omega_h + R - 1)]}{2}, & \omega_1 \leq \omega \leq \omega_0 \\ 0, & \omega \geq (1 + R)\omega_h = \omega_0 \end{cases} \quad (12)$$

where ω_1 and ω_0 are the unity and zero transmission frequencies, respectively.

The filter frequency response is flat up to ω_1 , has a cosine-shaped transition band, and is zero above ω_0 . The group delay of these filters is identically zero. This definition corresponds to an idealized family of filters which are not practically realizable, but have useful properties for study. CMP implements a full Nyquist and a square root Nyquist filter with $\omega_h/2\pi$ equal to one-half the symbol rate.

The transversal filter, also known as the tapped-delay-line filter, was implemented in CMP specifically for the study of adaptive equalizers. Designed for use on PSK signals, this filter is defined by a complex impulse response. It implements an M -tap complex coefficient filter, with taps spaced in time uniformly at the symbol rate, by creating a complex impulse response with nonzero values at N points spaced at the symbol rate. The impulse response is then used to compute the frequency response by means of the inverse FFT. The coefficients can be specified by the user to create a known distortion or can be computed by one of the adaptive equalizer processes to minimize the specified error function.

A filter can also be defined in terms of the amplitude and group delay characteristics at specified discrete frequency points. CMP first interpolates between these values to determine the filter characteristics at the same frequency values at which the signal is defined. The phase characteristic is

then obtained by integrating the group delay. The filter frequency response is

$$H(\omega) = |H(\omega)|e^{j\theta\omega} \quad (13)$$

where

$$\begin{aligned} |H(\omega)| &= \text{filter gain} \\ \theta_\omega &= \text{phase.} \end{aligned}$$

FILTER MODIFICATION

Once a filter is defined, its characteristics may be subsequently modified. The modifications currently implemented include cascading of filters, ideal equalization, and "sinc" compensation.

A cascaded filter is implemented by combining two filter characteristics to produce a single filter whose effect on the signal is equivalent to that of the two filters in series. The advantages of this approach over the straight-forward implementation of the two filters in series are numerous. The first is increased computational efficiency, since the cascaded version will execute faster. If different signals are to be filtered by the combination, there is a further savings in not having to redefine the filters. The user can obtain plots of the cascaded filter characteristics, as well as the cascaded noise bandwidth. The latter is required in determining BER as a function of C/N_0 if the receive filter is actually a cascade of two or more filters.

Ideal equalization can be used to eliminate the degradation due to the phase response of the filter. It is implemented simply by setting the phase response to zero at each point while leaving the amplitude response unchanged.

Consider a digital transmission system in which the source generates a signal with a power spectral density given by

$$S(f) = \frac{\sin(\pi f t_s)}{\pi f t_s} \quad (14)$$

where t_s is the symbol time. The power spectrum of the signal may be equalized or "whitened" by applying a filter having a frequency response which is the reciprocal of $(\sin x)/x$ (*i.e.*, the sinc^{-1} function). In digital communication systems, such a filter can be used in conjunction with Nyquist filtering to minimize intersymbol interference.

In CMP, "sinc" compensation is implemented by filtering the rectangular pulse PSK signal with the "sinc" compensation filter. The implementation permits the filter to be defined only in cascade with some other filter. The filter characteristics are altered only between the range $-R_s$ to R_s , where R_s

is the symbol rate. This procedure results in a reasonable approximation of the desired behavior while avoiding the singularities in the sinc^{-1} function.

FILTER IMPLEMENTATION

The filtering process consists of transforming the signal to the frequency domain, multiplying by the filter characteristics, and transforming back to the time domain. This operation assumes that the filter and the signal are centered at the same frequency. To filter a signal at an offset center frequency, it is necessary to shift the filter definition in the frequency before the multiplication takes place.

The frequency domain implementation has the following advantages over a time domain implementation based on recursive digital filters:

- a. The frequency response is defined exactly in the frequency domain and does not suffer distortions introduced by the bilinear z -transform.
- b. The filtering in the frequency domain is more efficient in most cases when an FFT algorithm is used. Once the frequency response is defined, the speed of filter implementation is independent of the order or type of the filter.
- c. Recursive filtering in the time domain is less accurate because of accumulating errors.
- d. The time domain implementation can result in the loss of data at the beginning and end of the signal.

Whenever a filter is defined or its characteristics are modified, CMP calculates the equivalent noise bandwidth,

$$B_n = \int_0^{\infty} |H(f)|^2 df \quad (15)$$

for use in subsequent BER calculations.

Nonlinear amplifier

A nonlinear amplifier whose transfer characteristics are independent of frequency over the frequency band of the signals being amplified is considered to be memoryless. Such an amplifier is specified by a pair of curves which relate the instantaneous power and phase of the output signal to the instantaneous power of the input signal.

The operating point is calibrated by calculating a scale factor which would scale the unmodulated signal power to the user-specified input backoff. This scale factor is then applied to the instantaneous power of the input signal.

The amplifier also normalizes the output power to 1 W. The phase curve is shifted by a constant to cause the phase shift at the carrier backoff point to be zero.

Miscellaneous operations

A number of additional operations may be performed on signals within CMP. Examples of some of these operations follow.

FREQUENCY SHIFT

A signal may be shifted in frequency by a specified amount. The complex envelope of the resulting signal is

$$z'(n) = z(n)e^{j\Delta\omega t_n} \quad (16)$$

where

- $z'(n)$ = complex envelope of the output signal
- $z(n)$ = complex envelope of the input signal
- $\Delta\omega$ = frequency shift
- $t_n = n/f_s$
- n = sample number (varies from 1 to N)
- f_s = sample rate.

SUMMATION OF SIGNALS

The following conditions are necessary to sum two or more signals:

- a. all signals must have the same sampling rate, and
- b. all signals must be defined in terms of the same reference frequency.

In CMP, the first constraint is ensured, since the same sample rate is used for all signals in a single execution. The second condition is met by frequency shifting each signal to the carrier frequency of the first (the designated) signal. Thus, two or more signals can be added, as follows:

$$z'(n) = \sum_{i=1}^k z_i(n)e^{j\Delta\omega_i t_n}, \quad n = 1, 2, \dots, N \quad (17)$$

where

- k = number of signals
- $\Delta\omega_i$ = difference between the frequency of the i th signal and the frequency of the first signal.

SCALING

Scaling of a signal involves multiplying the signal by a complex constant; *i.e.*,

$$z'(n) = cz(n), \quad n = 1, 2, \dots, N \quad (18)$$

A number of options are available for defining the scale factor c . For example,

$$c = Ae^{j\theta} \quad (19)$$

where A = gain factor
 θ = phase shift.

A signal may also be normalized to a specified power level p ; *i.e.*,

$$c = \sqrt{\frac{p}{p'}} \quad (20)$$

where p' is either the modulated or unmodulated power.

An option is also provided for nonlinear scaling with a cubic polynomial. In this case, the scaling factor for each sample is

$$c(n) = \frac{A'(n)}{A(n)} \quad (21)$$

where $A(n)$ and $A'(n)$ are the input and output amplitudes, respectively, which can be expressed as

$$A(n) = |z(n)| \quad (22)$$

$$A'(n) = b_1A(n) + b_2A^2(n) + b_3A^3(n) \quad (23)$$

where b_1, b_2, b_3 are user-specified constants.

TIME DELAY

A signal may be delayed by a specified period of time. The complex envelope of the resulting signal is then expressed by

$$z'(n) = z(n - \ell), \quad n = 1, 2, \dots, N \quad (24)$$

where ℓ is the number of samples by which the signal is delayed. The sample

values outside the signal range are recovered by assuming that the signal is periodic with a period equal to N samples.

An option is available whereby an offset QPSK signal can be generated simply by delaying the imaginary part by one-half the sample period.

PRODUCT OF SIGNALS

The product of two signals is obtained by complex multiplication of two signals; *i.e.*,

$$z(n) = z_1(n)z_2(n), \quad n = 1, 2, \dots, N \quad (25)$$

The multiplication of the complex envelopes of the two signals generates components at either the sum or difference frequencies of the signals as requested by the user.

Signal power

The modulated signal power is the average power of the modulated signal and is calculated by

$$p_m = \frac{\sum_{n=1}^N [x^2(n) + y^2(n)]}{N} \quad (26)$$

where $x(n)$ and $y(n)$ are the real and imaginary parts of the signal, respectively.

The unmodulated carrier power is used to define the operating point of amplifiers and as a reference power for BER calculations. It is defined as the power which would result if the modulation were removed and a single tone at the carrier frequency were produced. The unmodulated carrier power is updated as each process transforms the signal. For example, filters cause this value to be multiplied by the gain at the carrier frequency. Amplifiers cause this value to be scaled according to the specified transfer curve.

Receiver modeling

Two alternative approaches to modeling digital receiver performance are available with CMP. The first approach assumes an "ideal" receiver, which may not be physically realizable. In the second approach, an estimate is made of the error probability with a physically realizable model of a receiver phase and symbol timing estimator.

CMP assumes that the only process acting on the signal between calibration and detection is a single linear receive filter, which can be defined as a cascaded combination of various filters. Implementation of the receive filter results in the signal being filtered and the noise bandwidth of the filter being calculated.

The ideal detection process for QPSK modulation is shown in Figure 3. A signal with power P is present at the input to the receiver, along with Gaussian noise having a power density N_0 . The receiver performs the following operations on the signal output by the receive modem filter:

- It determines the average zero-crossing position in the waveforms (*i.e.*, it determines symbol timing).
- It determines the average phase angle of the complex samples and then shifts the phase of the signal such that the complex sample pairs are centered in the four quadrants of the phase plane (*i.e.*, at 45° , 135° , 225° , and 315°).
- From the original samples (which are at a higher sampling rate, typically with 16 samples per symbol), it selects the (X_ℓ, Y_ℓ) pairs at the center of each symbol.

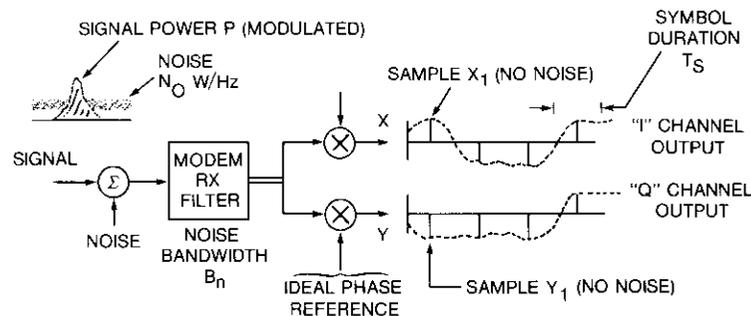


Figure 3. Coherent Demodulator

This process thus acts as an idealized coherent demodulator which obtains idealized symbol timing and phase estimation. It is also possible to introduce a constant bias in the sampling point or the phase of the detected signal. This bias would result in a shift of the sampling point from the center of the symbol in the first case and a shift of the average phase of detected samples in the second case. Introducing the bias is often necessary to compute a minimum value of BER, since the idealized detection algorithm does not, in all cases, result in minimum BER. In general, the most useful computation is the minimum BER with respect to phase and timing biases.

As an alternative to the ideal detection process, synchronization effects may be included by using a specific model for phase and timing recovery with noise added to the signal. The particular phase and timing recovery modules which are implemented were chosen to be typical and represent one possible receiver implementation. Timing is recovered by multiplying the signal by a half-symbol-time-shifted version of itself and filtering the resulting signal. Positive-going zero crossings are then used to determine the correct sample point. Phase is recovered by raising the signal to the fourth power, computing the arc tangent, and then filtering the resulting signal.

Performance estimation

Performance estimation capabilities within CMP include graphical outputs, spectral analysis, and BER calculations.

The on-line graphical display allows the user to determine the validity of the system model and to qualitatively assess system performance. Plots of the signal envelope, the filter characteristics, and the amplifier characteristics also serve as diagnostic tools. The eye pattern indicates the degree of intersymbol interference in PSK signals and is useful in determining the optimum sampling point. The scatter diagram is a useful qualitative indicator of the channel performance. In an ideal channel, exhibiting no intersymbol interference or other distortion, the sampled values should lie precisely on top of one another. The scattering of these points is a measure of the degree and type of channel distortion. The phase plane trajectory provides useful information about the time history of a modulated signal.

The power spectral density of a signal permits the analysis of baseband distortion in FDM/FM signals. A typical application is the determination of the noise power ratio (NPR) of the signal after transmission through the channel. To perform this analysis, a baseband signal is created with a narrow gap at the frequency at which NPR is to be determined. After demodulation, the spectrum of the baseband can be examined to determine the noise power at that frequency.

The CMP is specifically designed to calculate BER for PSK signals which are degraded by a combination of intersymbol interference and thermal noise on the down-link path of a satellite channel. By a proper design of the system model, a user can also evaluate co-channel and adjacent-channel interference or account for up-link thermal noise.

A hybrid simulation approach is taken in calculating the BER of PSK signals. In this approach, the received signal is first demodulated to determine the instantaneous sampled values due to intersymbol interference only. Then the

effects of thermal noise are accounted for analytically by superimposing the random noise distribution on each of the sampled values obtained from the simulation. This approach is valid for the class of coherent receivers for which the demodulation can be modeled as a linear process and the noise can be assumed to be Gaussian. This technique results in a considerable savings in computer time, since the entire error probability curve can be obtained from a single simulation run.

The analysis consists of three steps: calibration, detection, and BER computation. The calibration is performed on the signal just before the receive filter and consists of calculating the modulated and unmodulated signal powers.

The average BER of the signal is determined as a function of the ratio of energy per bit, E_b , to the noise power density, N_o . The quantity E_b is derived from either the unmodulated or modulated signal power at the point where the noise would be added. The noise power at the detection point is

$$\sigma^2 = N_o B_n \quad (27)$$

where N_o is the noise power density, and B_n is the noise equivalent bandwidth of the receive filter.

The BER is simply the average probability that a given bit will be incorrectly detected, with the average taken over the sample population of bits:

$$\overline{P(\epsilon)} = \frac{1}{L} \sum_{\ell=1}^L P(\epsilon|X_\ell, Y_\ell) \quad (28)$$

where $L =$ number of symbols
 $X_\ell, Y_\ell =$ voltage samples on the I and Q channel, respectively, on the ℓ th symbol
 $P(\epsilon|X_\ell, Y_\ell) =$ probability that an error occurs on a particular symbol for the values X_ℓ and Y_ℓ .

For a QPSK signal and Gaussian noise with zero mean, this probability is

$$P(\epsilon|X_\ell, Y_\ell) = \frac{1}{2} \left[\frac{1}{2} \operatorname{erfc} \left(\frac{|X_\ell|}{\sigma} \right) + \frac{1}{2} \operatorname{erfc} \left(\frac{|Y_\ell|}{\sigma} \right) \right] \quad (29)$$

where

$$\operatorname{erfc}(X) = \frac{2}{\sqrt{\pi}} \int_X^\infty e^{-u^2} du \quad (30)$$

For a BPSK signal, only the sample value Y_ℓ is used in detection, and the probability of bit error with the same definition of noise is

$$P_b(\epsilon|Y_\ell) = \frac{1}{2} \operatorname{erfc} \frac{|Y_\ell|}{\sigma} \quad (31)$$

The effects of co-channel and adjacent-channel interference can also be evaluated. The interfering signals are generated with the proper center frequencies and added to the designated signal. The analysis proceeds as in the single signal case. However, a number of factors must be accounted for in order for the computation to be done correctly. In specifying the frequency of the interfering signal (or signals), the user should be cognizant of the frequency shift which will be implemented when the signals are summed. Each signal which is added to the designated signal will have frequency components aliased from either the upper to the lower frequencies, or vice versa, depending on the sign of the frequency shift implicit in the summation. The total power misplaced by this aliasing can be decreased by increasing the sample rate. If the summed signal is passed through any nonlinear process, the aliasing will be much more severe for spectral components which are close to half the sample rate. This effect can also be alleviated by increasing the sample rate.

Up-link noise can be simulated in much the same way as co- and adjacent-channel interference, with similar problems. Noise can be specified as a separate signal which is spectrally flat over the entire sample rate. This signal is then added to the designated signal. Significantly longer signals will usually be required to obtain statistically significant results.

A sample application

An example of an application involving adjacent-channel interference will be given. Figure 4 is a block diagram of the simulation model. The model source consists of three QPSK signals which are generated at a symbol rate of 60 MHz. The number of samples per symbol was chosen to be 16, giving a sample rate of 960 samples per microsecond. The adjacent channels are centered at $\pm S$ in frequency with respect to the central channel. The earth station modem transmit and receive filters are represented by appropriate boxes; the transponder input and output filters are represented by F_{IN} and F_{OUT} , respectively. The ground station HPA and the satellite TWTA are assumed to operate at an output power of -5.0 dB relative to single carrier saturation. The detection is performed on the waveform of the center channel.

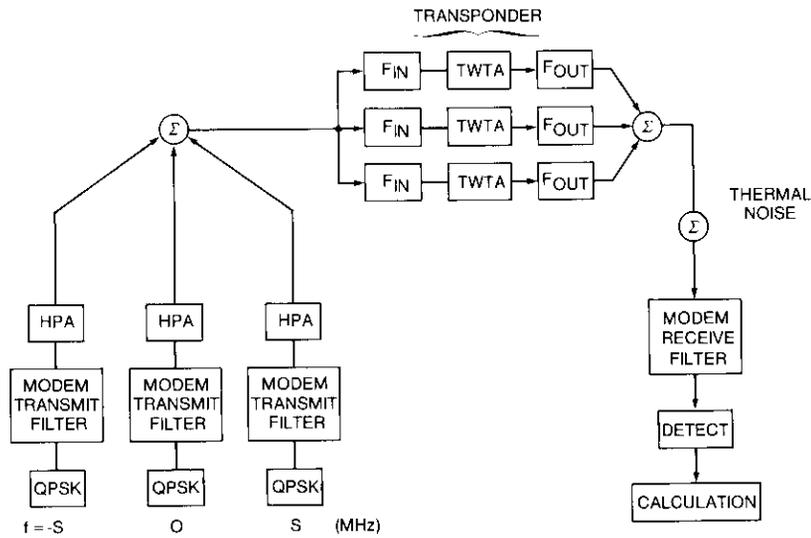


Figure 4. Adjacent Channel Simulation Model of a Satellite System

Calculations of BER for the center channel are made with respect to a perfect coherent receiver/demodulator. The ratio E_b/N_o is calculated relative to the unmodulated carrier power.

Figure 5 shows a BER table generated by the program. The table gives BERS for various values of E_b/N_o , as well as $\Delta E_b/N_o$, which is the performance loss, expressed in decibels, compared to ideal QPSK performance. Figure 6 shows the power spectrum of a QPSK signal prior to filtering. Figures 7 through 9 show the characteristics of various system elements. Figure 10 is the eye diagram of the received signal, and Figure 11 is a plot of BER vs E_b/N_o .

Conclusions

This paper has described a general purpose channel modeling and simulation program (CMP) for analyzing the performance of satellite communications channels. Work is currently underway to enhance the capabilities of CMP.

Future plans include the development of a command preprocessor which will validate an input file before execution. The preprocessor will ensure that the user input commands contain no errors which would terminate the execution of the program. This preprocessor will greatly reduce the number of runs required to achieve the desired results.

BER	E_b/N_o	DELTA
5.0E-03	7.03	1.84
2.0E-03	8.29	2.13
1.0E-03	9.16	2.39
5.0E-04	9.96	2.65
2.0E-04	10.96	2.99
1.0E-04	11.65	3.28
5.0E-05	12.32	3.55
2.0E-05	13.15	3.91
1.0E-05	13.72	4.16
5.0E-06	14.26	4.38
2.0E-06	14.95	4.70
1.0E-06	15.40	4.90
5.0E-07	15.84	5.08
2.0E-07	16.38	5.30
1.0E-07	16.76	5.47
5.0E-08	17.12	5.63
2.0E-08	17.54	5.78
1.0E-08	17.86	5.89
5.0E-09	18.15	6.01
2.0E-09	18.50	6.13
1.0E-09	18.76	6.24
5.0E-10	19.02	6.33
2.0E-10	19.30	6.40
1.0E-10	19.52	6.47

Figure 5. Sample BER Table

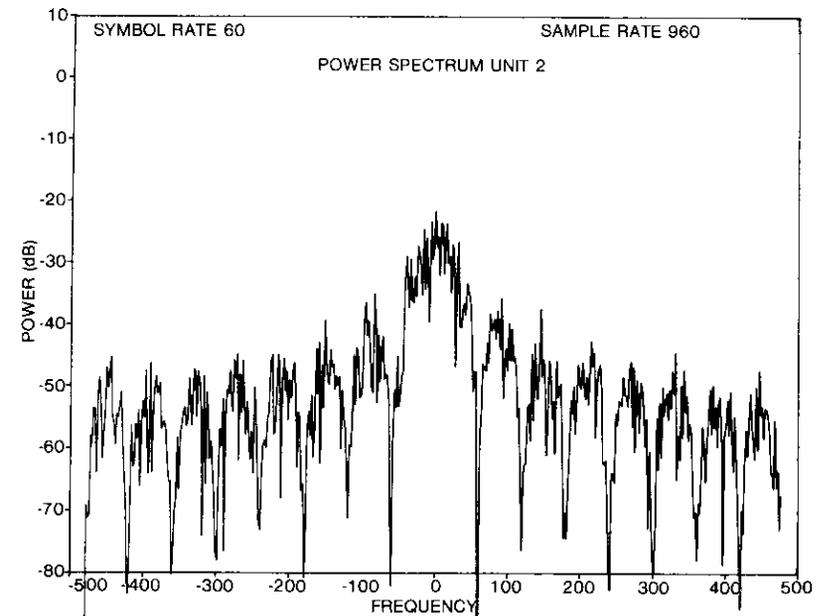


Figure 6. PSK Signal Spectrum

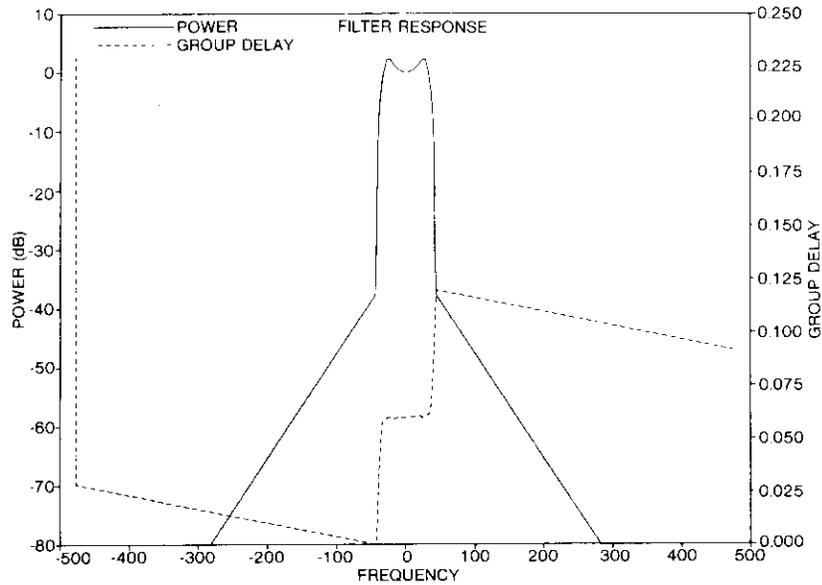


Figure 7. Transmitter Filter Characteristics

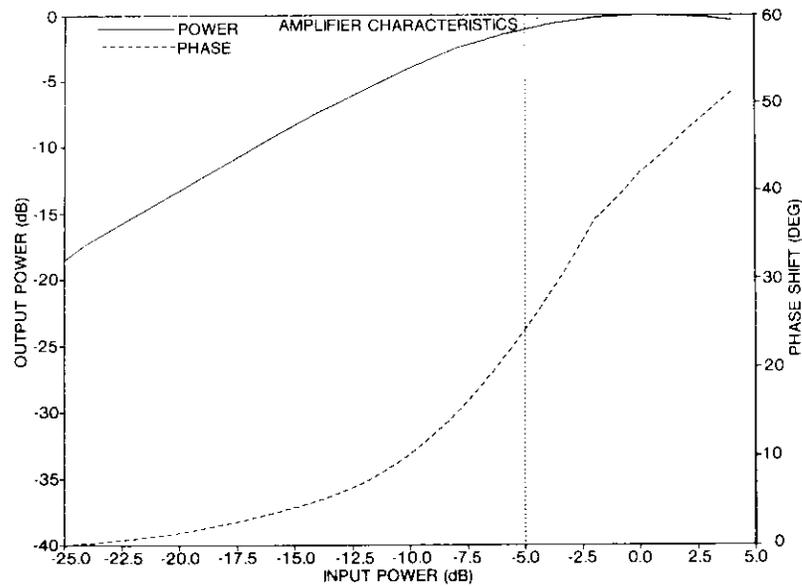


Figure 8. HPA Characteristics

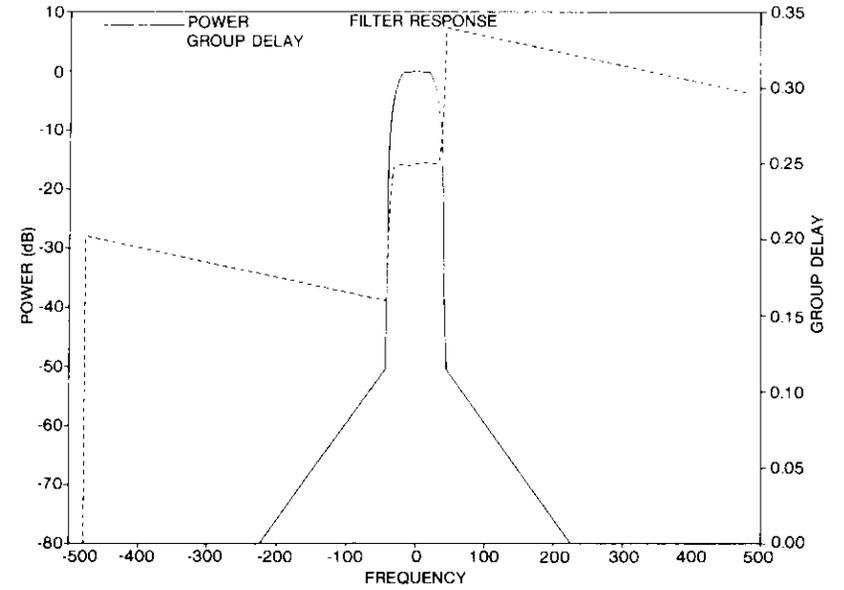


Figure 9. Receiver Filter Characteristics

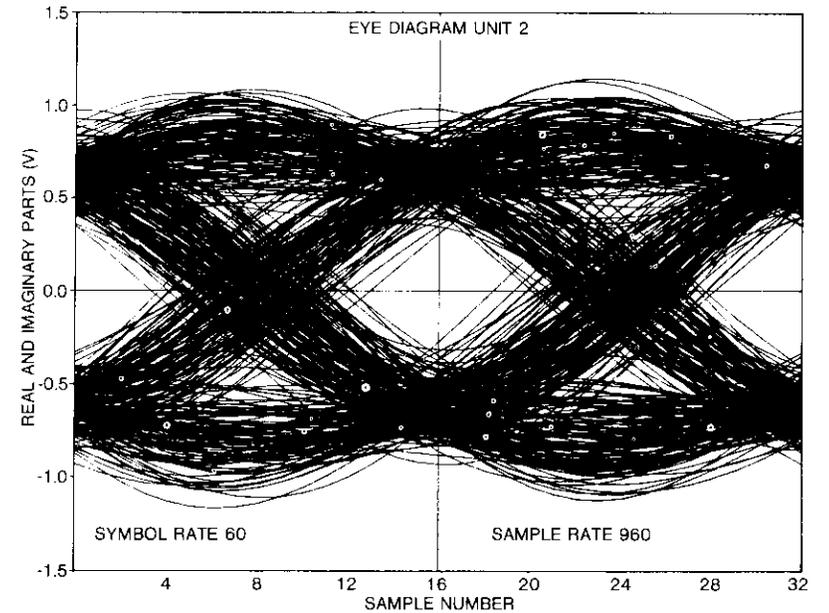


Figure 10. Eye Diagram of the Received Signal

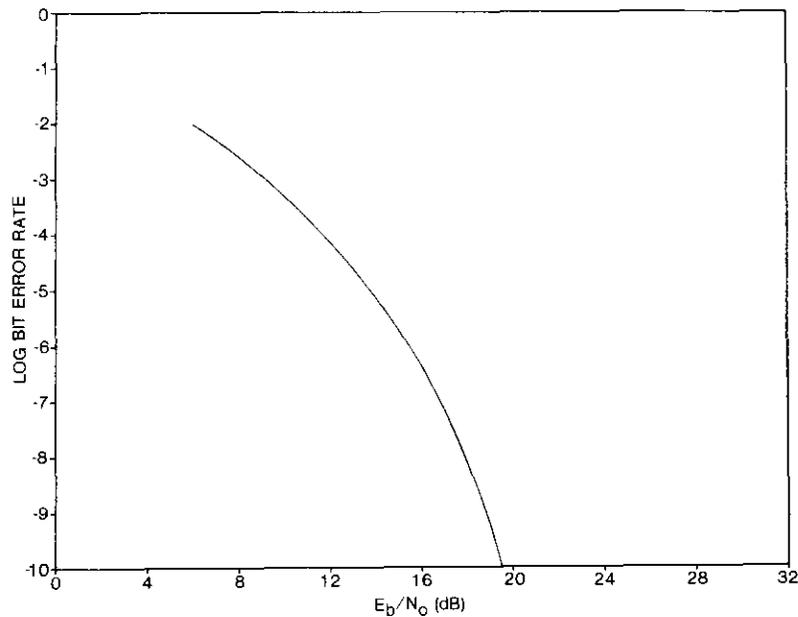


Figure 11. BER Performance of the Adjacent Channel Simulation Model

The scope of the program will be extended to handle very long signals. This development will greatly improve the ability of the program to evaluate baseband distortion in FDM/FM signal analysis. Other plans include the development of test patterns for TV simulation, implementation of practical receivers, and determining the effects on BER of various types of convolutional and block coding schemes.

Acknowledgments

The CMP has been under development at COMSAT for several years. The present version is the result of contributions of many individuals, including S. Lebowitz, D. Meadows, L. Palmer, S. Rhodes, and C. Wolejsza.

References

[1] S. A. Rhodes and L. C. Palmer, "Computer Simulation of a Digital Satellite Communications System Utilizing TDMA and Coherent Quadrature Signaling," International Conference on Communications, Philadelphia, June 1972, *Conference Record*, pp. 34-19-34-24.

- [2] M. Kato and H. Inose, "Differentially Coherent Detection Scheme With Error-Correcting Capability by Means of Decision Patterns," *Electronics and Communications in Japan*, Vol. 54-A, No. 5, 1971, pp. 1-8.
- [3] J. Dodo, S. Iwai, and K. Kawai, "Computer Simulation of Carrier Recovery Circuit in Microwave PCM-PSK System," *Fujitsu Scientific and Technical Journal*, September 1972, pp. 59-82.
- [4] U. Mengali, "Synchronization of QAM Signals in the Presence of ISI," *IEEE Transactions on Aerospace and Electronic Systems*, AES-12, No. 5, September 1976, pp. 556-560.
- [5] K. Feher and R. Cristofaro, "Transversal Filter Design and Application in Satellite Communications," *IEEE Transactions on Communications*, COM-24, No. 11, November 1976, pp. 1262-1268.
- [6] R. M. Fielding, H. L. Berger, and D. L. Lochhead, "Performance Characterization of High Data Rate MSK and QPSK Channel," International Conference on Communications 42-6, 1977, *Proc.*, pp. 3.2-42-3.2-46.
- [7] M. K. Klukis, "Hybrid Computer Simulation of Digital Transmission Over a Fading Microwave Channel," IEEE National Telecommunications Conference, 1977, *Conference Record*, pp. 16:6-1-16:6-9.
- [8] C. Cahn and C. May, "Bandwidth-Efficient QPSK for Multiple-Access Satellite Communication," IEEE National Telecommunications Conference, 1977, *Conference Record*, pp. 42:3-1-42:3-4.
- [9] A. R. Bens et al., "A Computer Simulation for the Analysis of Interference Problems Related to Satellite Communication Systems," IEEE Canadian Communications and Power Conference, Montreal, Canada, October 1976.
- [10] R. A. Plemel, J. A. Copeck and K. L. Krukewich, "A Computer Simulation for the Analysis of Interference Problems Related to Mobile Satellite Communications Systems," Maritime and Aeronautical Satellite Communication and Navigation Conference, London, England, 1978, *Proc.*, pp. 89-92.
- [11] Y. Furuya, F. Akashi and S. Murakami, "A Practical Approach Toward Maximum Likelihood Sequence Estimation for Bandlimited Nonlinear Channels," Fifth International Conference on Digital Satellite Communications, Genoa, Italy, 1981, *Proc.*, pp. 295-301.
- [12] G. Benelli et al., "Combined Amplitude-Phase Modulation for a VHF Communication Link," IEEE International Conference on Communications, Denver, Colorado, 1981, *Proc.*, Vol. 3, pp. 56.6.1-6.
- [13] P. Amadesi, "Clock Recovery for Onboard Regeneration in a TDMA Satellite Communication System," *CSELT Rapporti Tecnici*, Vol. VIII-N.4, 1981, pp. 261-270.
- [14] M. U. Chang et al., "Satellite Communication System Performance Evaluation by Computer Simulation," IEEE National Telesystems Conference, 1982, *Conference Record*, pp. C.3.2.1-C.3.2.5.
- [15] S. Stein and J. Jones, *Modern Communication Principles*, New York: McGraw-Hill, 1967.
- [16] L. Weinberg, *Network Analysis and Synthesis*, New York: Robert E. Krieger Publishing Company, 1975.

- [17] A. Peled and B. Liu, *Digital Signal Processing*, New York: John Wiley and Sons, 1976.
- [18] C. M. Kudsia and N. K. M. Chitre, "Transmission and Reflection Group Delay of Butterworth, Chebychev, and Elliptic Filters," *RCA Review*, 1969, Vol. 30, No. 2, pp. 248-267.
- [19] R. V. Churchill, J. W. Brown, and R. F. Verhey, *Complex Variables and Applications*, New York: McGraw-Hill, 1974.
- [20] A. Antoniou, *Digital Filters: Analysis and Design*, New York: McGraw-Hill, 1979.



Abdul Hamid received a B.S. from Faisalabad University, Pakistan, in 1972; an M.S. from the University of Manitoba, Canada, in 1974; and a Ph.D. from North Carolina State University in 1977, all in Engineering. Prior to joining COMSAT Laboratories, he participated in research and development in electronic instrumentation and contactless measurements at the University of Manitoba and was involved with design and development of mini/micro-based systems for monitoring and control of real-time processes at North Carolina State University. From 1978 to 1981, as a Staff Engineer at

ENSCO, Inc., he was engaged in engineering analyses and software development in the areas of digital signal processing, statistical characterization, time series analysis, and data base management systems.

Dr. Hamid joined COMSAT Laboratories in 1982 and is currently a Member of the Technical Staff in the Systems Analysis Software Department of the Computer Sciences Laboratory. His work at COMSAT has included software development in the areas of channel modeling and simulation for analysis of various communications configurations and direct-broadcast TV. He is a member of IEEE and has approximately 25 publications in professional journals.

Steven R. Baker received a B.S.E.E. in 1975 and an M.Eng. degree with honors in 1976, both from Cornell University. Before joining COMSAT Laboratories, he worked for 3 years on the design of computer-controlled navigation and guidance systems with IBM's Federal Systems Division. From 1980 to 1982, he was employed in the Computer Sciences Laboratory at COMSAT and was responsible for the continuing development of the Channel Modeling Program, with an emphasis on the development of user applications. In 1982, he transferred to the Communications Technology Division, where he has worked in the areas of acquisition and synchronization, demand assignment, and TDMA system development. Mr. Baker is a member of Tau Beta Pi.



William L. Cook received a B.S. in Engineering Mechanics from Lehigh University in 1964, an M.S. in Engineering Sciences from Purdue University in 1966, and a D.Sc. in Computer Sciences from The George Washington University in 1973. He is currently Director of the Systems Development Division at COMSAT Laboratories, where his responsibilities include research and development activities in packet communications, systems analysis software development, and architectural design and software engineering for complex computer-based distributed systems. While at COMSAT,

Dr. Cook has been responsible for the development of analysis techniques and computer software used in the evaluation and optimization of communications satellite systems and subsystems. He was also project manager for experimental programs in high-speed computer communications via satellite. Prior to joining COMSAT, he was a structural engineer with Goddard Space Flight Center and was a member of the group responsible for development of the NASA Structural Analysis Program (NASTRAN).

Multiple scattering in coherent radio-wave propagation through rain

D. V. ROGERS AND R. L. OLSEN

(Manuscript received May 24, 1983)

Abstract

Scattering quantities for coherent wave propagation through a sparsely populated, uniformly random medium of discrete scatterers are analyzed. Particular attention is given to interpreting the various scattering processes to show that the commonly used classical (or Mie-Ryde) theory for calculating the coherent transmitted field accounts for all multiple scattering interactions that involve purely forward scattering processes. Numerical calculations based on Twersky's "free-space" scattering formalism are presented to show that, for coherent transmission through rain, all other multiple scattering processes are negligible for frequencies up to at least 1,000 GHz. It is thus concluded that the classical formulation for computing specific attenuation and specific phase rotation is quite accurate for multiple scattering applications for the rain medium.

Introduction

The theory of wave propagation in random media of discrete scatterers is applicable to many practical problems [1]. One important application is in the analysis of radio wave propagation through rainfall [2]. For example, accurate estimates of amplitude and/or phase behavior for coherent propagation through rain are necessary for efficient, reliable design of many telecommunications systems, particularly those operating at frequencies above 10 GHz or employing dual orthogonal polarizations for frequency reuse. This

problem has been widely studied, and many scattering calculations for the rain medium have been performed.

Attention has also been directed to the question of multiple scattering in rain [3]–[9]. Neglect of multiple scattering is sometimes suspected of being responsible for differences between theory and experiment for rain attenuation [10]. Some researchers have attempted to establish criteria for the conditions under which multiple scattering in rain is important [3], [4], [9]. Others have argued on physical grounds that the most significant multiple scattering effects are accounted for in the classical theory, at least for attenuation [11]. The perception that the theory accounts only for single scattering often appears to be based on the derivation of a refractive index for discrete random media, ostensibly with the application of single scattering theory, by van de Hulst [12]. Additional confusion has arisen from incorrect use of the definition of single scattering. Incorrect or unclear interpretations of the effects are common [3]–[6], [9], [10], [13]–[15], including a previous one by the present authors [16].

This paper shows that the classical, or Mie-Ryde, approach for computing scattering parameters for coherent transmission through rain actually accounts for all significant multiple scattering processes. This result, which holds both for signal attenuation and for phase rotation, is in fact general, applicable to discrete random media other than rain. Because the discussion attempts to correct a long-held misconception on this topic, the approach is slightly tutorial and represents an extended presentation of previous work on this topic [8].

Scattering in a sparsely populated, uniformly random medium of discrete scatterers

Figure 1 illustrates several types of scattering that can occur upon transmission through a discrete random medium. Single scattering, by well-established definition, corresponds to the case in which every scatterer perceives the unattenuated incident field [12], so that scattering effects on the transmitted field are linearly proportional to the volume number density of scatterers. Conversely, multiple scattering applies whenever some or all of the effects of previous or subsequent scatterings are taken into account. By this definition, “forward multiple scattering” refers to multiple scattering interactions for which every segment in the scatterer-to-scatterer “chain” has a component in the forward direction; that is, each path between successive scatterers resides in the forward hemisphere, as defined by the direction of the incident field. “Backward multiple scattering” applies to multiple scattering processes for which at least one path between scatterers

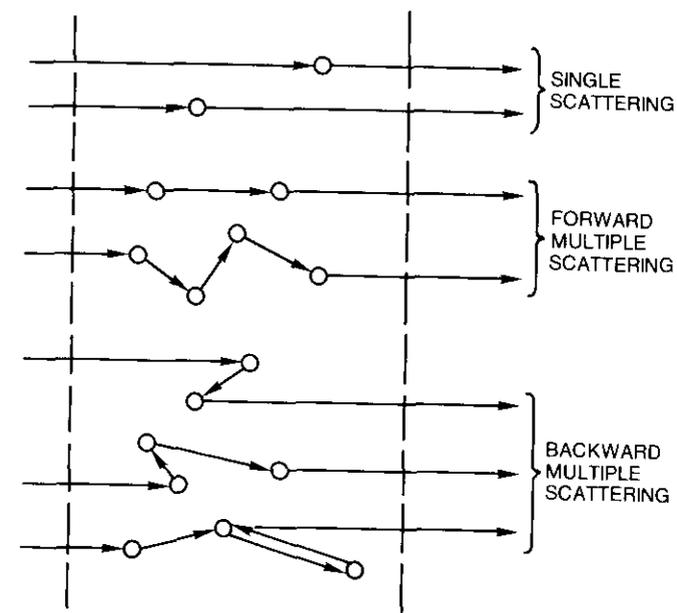


Figure 1. Schematic Diagram of Several Multiple Scattering Processes Contributing to the Coherent Transmitted Field

is in the backward hemisphere, and encompasses higher order effects, such as back-and-forth interactions between pairs of scatterers. Note that single scattering corresponds to zeroth-order forward multiple scattering.

Propagation through discrete random media

The scattering medium is represented by an infinite slab region of depth d , containing a sparse population (*i.e.*, scatterer separation \gg wavelength λ) of randomly located discrete scatterers. A field $\vec{E}_i(\vec{r}, t) = \vec{E}_i(\vec{r}) e^{i2\pi f t}$ of frequency f , wavelength λ , and direction \vec{k} ($k = 2\pi/\lambda$) is incident at an angle θ to the normal of the slab face. Typically, the incident field is assumed to be a plane wave, with $E_i(\vec{r}) = e^{-j\vec{k}\cdot\vec{r}}$, but it may also be an isotropic or an anisotropic spherical wave [6]. As shown in Figure 2, the effect of any particular fixed configuration of scatterers on the transmitted wave is described by a complex transmission coefficient, Y . The transmission coefficient for the medium is thus the average of Y over all possible scatterer configurations, *i.e.*, the ensemble average $\langle Y \rangle$, where $\langle Y \rangle E_i(\vec{r}, t)$ is the coherent transmitted field.

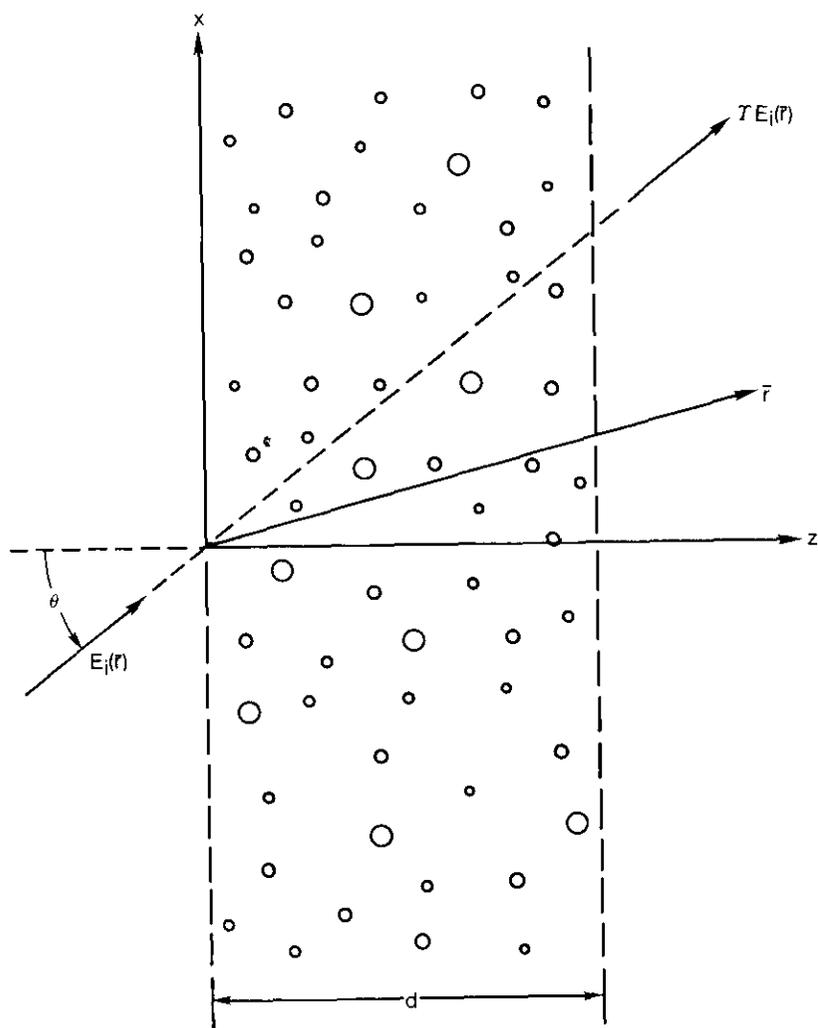


Figure 2. Geometry for Transmission Through a Slab Region of Discrete Random Scatterers

For simplicity, the scattering region is first assumed to comprise identical, homogeneous spherical scatterers. It can therefore be characterized as a polarization independent, monodisperse medium of average volume number density, N (scatterers per unit volume). Subsequently, the results will be generalized to polydisperse media by incorporating weighted averages of the

scattering amplitudes. Since the following formulation is in terms of a scalar theory, the conclusions are expected to apply to the characteristic (nondepolarized) polarizations [17] of media composed of nonspherical and/or nonhomogeneous scatterers [18], thereby permitting extension of the results to depolarizing media.

The slab region is now divided into n identical thin slabs of thickness Δd (*i.e.*, $d = n\Delta d$), with Δd sufficiently small that the single scattering approximation is applicable within each thin layer, but sufficiently large that the net scattering characteristics of the slab can be represented by an equivalent transmission coefficient, $\langle Y \rangle$. (Use of the single scattering approximation within successive thin layers is a basic assumption, either explicit or implicit, in previous derivations [12], [19], [20].) For identical scatterers with a forward scattering amplitude (*i.e.*, amplitude at zero scattering angle, parallel to the incident field) of $S(0)$, the single scattering transmission coefficient for a thin slab is [12]

$$\langle Y \rangle_{\Delta d} = 1 - \frac{2\pi N \Delta d}{k^2 \cos \theta} S(0) \quad (1)$$

For the entire scattering region composed of n slabs, the transmission coefficient $\langle Y \rangle$ may be presumed [21] to be the cascaded product, $\langle Y \rangle_{\Delta d}^n$. In terms of $d = n\Delta d$, therefore,

$$\langle Y \rangle_{\Delta d}^n = \left[1 - \frac{2\pi N d S(0) / k^2 \cos \theta}{n} \right]^n \quad (2)$$

The single scattering approximation for each thin layer theoretically can be made more accurate by decreasing Δd while keeping the transmission coefficient per unit length constant. By definition of the exponential function, in the limit as $\Delta d \rightarrow 0$ (*i.e.*, $n \rightarrow \infty$) the transmission coefficient for the total slab becomes

$$\langle Y \rangle = \lim_{n \rightarrow \infty} \langle Y \rangle_{\Delta d}^n = \exp \left(\frac{-2\pi N d S(0)}{k^2 \cos \theta} \right) \quad (3)$$

This derivation is believed to follow that implied by van de Hulst [12].

The specific attenuation, A (in decibels per unit distance), and the specific phase rotation, ϕ (in radians per unit distance), are obtained from equation (3) as

$$A = \frac{20}{\ell n 10} \frac{2\pi N}{k^2} \text{Re}[S(0)] \quad (4)$$

$$\phi = \frac{2\pi N}{k^2} \text{Im}[S(0)] \quad (5)$$

where, as customary, the attenuation in nepers is converted to decibels by multiplying by $20/\ell n 10$ (~ 8.686). Several variants of these expressions exist. For example, the bulk refractive index η_f of a continuous slab region of depth d having identical coherent transmission characteristics to the original slab composed of discrete random scatterers is [12]

$$\eta_f = 1 - j \frac{2\pi N}{k^3} S(0) \quad (6)$$

The subscript f denotes forward multiple scattering and will be justified shortly. With this expression, equations (3) to (5) become, respectively,

$$\langle Y \rangle = \exp \left[\frac{-jkd(\eta_f - 1)}{\cos \theta} \right] \quad (7)$$

$$A = 8.686k \text{Im}(\eta_f) \quad (8)$$

$$\phi = k \text{Re}(\eta_f - 1) \quad (9)$$

For polydisperse distributions of scatterers, where $N(D) dD$ is the number density of scatterers with equivalent diameter D in the size interval dD , the only modification required in equations (1) through (6) is

$$N \cdot S(0) \rightarrow \int_0^\infty S(0, D) N(D) dD \quad (10)$$

where $S(0, D)$ is the forward scattering amplitude for a scatterer of diameter D . For instance, in terms of the extinction cross section $C_e(D) = (4\pi/k^2) \text{Re}[S(0, D)]$, equation (4) becomes

$$A = 4.343 \int_0^\infty C_e(D) N(D) dD \quad (11)$$

which is the most frequently used expression for specific attenuation in rainfall.

Multiple scattering contributions contained in the classical formulation

The expressions for $\langle Y \rangle$ as given by equation (3) or (7), or the respective polydisperse forms, correspond exactly to the commonly used expressions for computing specific attenuation and phase rotation for coherent radio wave propagation through rain. Generally, this approach is presumed to be a single scattering theory. It will now be shown that this assumption is incorrect. A series expansion of the expression for $\langle Y \rangle$, *e.g.*, as given by equation (3), is

$$\langle Y \rangle = 1 - wS(0) + \frac{1}{2!} w^2 S^2(0) - \frac{1}{3!} w^3 S^3(0) + \dots \quad (12)$$

where $w = 2\pi Nd/k^2 \cos \theta$. Since single scattering contributions as given by equation (1) are completely accounted for by the first two terms of equation (12), the transmission coefficient $\langle Y \rangle$ clearly embodies additional scattering interactions.

This point is crucial. Many authors have erroneously referred to $\langle Y \rangle$ as given by equation (3), or to similar forms such as equations (4) through (11), as the single scattering approximation. This error seems to arise in general from two main sources. The first is a failure to recognize that single scattering interactions must be characterized by an expression such as equation (1), in which effects of individual scatterers on the coherent transmitted field are linear and certainly not exponential as in equation (3). The second is to assume equation (3) or one of the equivalent forms as the starting point for scattering analyses, while accepting the statement (Section 1.22 of Reference 12) that the theory is limited to single scattering.

Actually, Twersky [22] recognized that the term in equation (12) involving $S^n(0)$ accounts for the coherent contribution of all forward scattering processes of order n . Since equation (12) is an infinite series, the exponential form for $\langle Y \rangle$, equation (3) or equation (7) in conjunction with equation (6), therefore contains the coherent contribution of all orders of forward multiple scattering. Forward multiple scattering is thus taken into account for both the attenuation and phase rotation of the coherent transmitted field. Furthermore, it is accounted for "exactly," since the requirements that the total number of scatterers be Poisson distributed [7] or approach infinity [22] are both effectively met in an actual rain medium.

This result applies to waves of any polarization propagating through a sparse medium of discrete, homogeneous spherical scatterers. It should be emphasized, however, that the preceding development treats the transmitted

coherent wave only and does not apply to the direct backscattered component of the field as detected, for example, by monostatic radars.

Origin of the forward multiple scattering contributions

It will now be demonstrated why the above derivation of $\langle Y \rangle$ accounts for the forward multiple scattering interactions. Again consider n slabs of thickness Δd and retain all orders of forward multiple scattering within each slab (but ignore multiple scattering interactions between the layers). Then

$$\begin{aligned} \langle Y \rangle_{\Delta d}^n &= \left[1 - \frac{2\pi N \Delta d}{k^2 \cos \theta} S(0) + \frac{1}{2!} \left(\frac{2\pi N \Delta d}{k^2 \cos \theta} \right)^2 S^2(0) - \dots \right]^n \\ &= \{ \exp [- 2\pi N \Delta d S(0) / k^2 \cos \theta] \}^n \\ &= \exp [- 2\pi N d S(0) / k^2 \cos \theta] \end{aligned} \tag{13}$$

which is identical to equation (3). Thus, it is evident that retention of all orders of microscopic forward multiple scattering within individual slabs is equivalent to successive application of the single scattering approximation to infinitely thin layers of the scattering medium and that both approaches account for all orders of forward multiple scattering. In effect, the microscopic forward multiple scattering processes disappear within individual layers in the limit as $\Delta d \rightarrow 0$, but reappear for the overall slab through the macroscopic forward multiple scattering between the layers.

Multiple scattering terms neglected by the classical formulation

Twersky developed two multiple scattering theories [23]–[25], the “free-space” and “two-space” formalisms, that provide expressions for the bulk refractive index containing backward multiple scattering contributions. Both theories yield the same exact representation for the coherent transmitted field as the number density N approaches zero [7]. For propagation through rain, characterized by very small number densities [11], attention may be confined to the simpler free-space formalism, which is essentially exact for natural rainfall [8].

In the free-space formalism, for the geometry of Figure 2 and a monodisperse distribution,

$$\langle Y \rangle = \mathcal{D} (1 - Q^2) e^{-j(\Gamma - \gamma)d} \tag{14}$$

where

$$\gamma = k \cos \theta \tag{15a}$$

$$\Gamma^2 = k^2 (\eta_{bf}^2 - \sin^2 \theta) \tag{15b}$$

$$\eta_{bf}^2 = \epsilon \mu \tag{15c}$$

$$\mathcal{D} = (1 - Q^2 e^{-j2\Gamma d})^{-1} \tag{15d}$$

$$Q = \frac{CS(\pi)}{CS(0) + j(\Gamma + \gamma)} \tag{15e}$$

$$C = \frac{2\pi N}{k^2 \cos \theta} \tag{15f}$$

The quantity $S(\pi)$ is the backward scattering amplitude parallel to the plane of scattering, and η_{bf} is a bulk refractive index containing forward and some backscattering contributions. The bulk parameters Q , ϵ , and μ correspond respectively to the Fresnel reflection coefficient at the interface, the permittivity, and the permeability of a continuous slab region having identical coherent transfer properties as the discrete random slab.

For a random medium composed of arbitrary, but identical, discrete scatterers, Twersky derived for the free-space formalism [23]

$$\eta_{bf}^2 = 1 - j \frac{4\pi N}{k^3} S(0) - \left(\frac{2\pi N}{k^3 \cos \theta} \right)^2 [S^2(0) - S^2(\pi)] \tag{16}$$

Both $S(0)$ and $S(\pi)$ are replaced by their weighted averages for distributions of nonidentical scatterers [22], with equations of the form of equation (10). For normal incidence, equation (16) becomes

$$\eta_{bf}^2 = \eta_f^2 + \left(\frac{2\pi N}{k^3} \right)^2 S^2(\pi) \tag{17}$$

with η_f as given by equation (6). By using equation (16) for η_{bf} , it will be shown that the backward multiple scattering term is completely negligible for coherent transmission through rain.

Equations (14), (15a) through (15f), and (17) reveal that the free-space theory incorporates several backward multiple scattering processes that are not accounted for in the forward multiple scattering theory of equation (7). The term $\mathcal{D}(1 - Q^2)$ in equation (14) represents, macroscopically, the coherent multiple reflections between the interfaces of the bulk slab of refractive index η_{bf} , while the second term of equation (17) accounts for some of the coherent backscattering within the slab. There are yet higher orders of multiple scattering missing from equation (14). As discussed by Twersky [22], these processes contain one or more back-and-forth interactions between pairs of scatterers. Although the back-and-forth processes have been

numerically shown to become partially coherent for media in which the scatterers are closely spaced in terms of wavelengths, they may be assumed to be entirely incoherent for a low density medium such as rain [7].

Application to the rain medium

The previous treatment for the coherent transmitted field is actually quite general, and the results regarding multiple scattering contributions are applicable to any polarization-independent, sparse population of discrete random scatterers. Now, the analysis is specialized to the rain medium by using physical parameters of rain to compute the magnitude of the various multiple scattering contributions embodied in equation (14), with η_{bf} given by equation (17).

First, an approximation is made that is consistent with the small number densities of drops in real rain and that allows the exponent of equation (14) to be recast in a form identical to that of equation (7). Namely, for a slab region sparsely populated with discrete scatterers, the bulk refractive index η_{bf} is very nearly unity; thus, under this condition,

$$\begin{aligned} j(\Gamma - \gamma)d &= jkd[(\eta_{bf}^2 - \sin^2 \theta)^{1/2} - \cos \theta] \\ &\approx jkd(\eta_{bf} - 1)/\cos \theta \end{aligned} \quad (18)$$

This approximation, always exact for normal incidence ($\theta = 0$), is very accurate for modeled rain distributions, whether monodisperse or polydisperse, even at extreme rain rates [8].

Secondly, the Fresnel reflection coefficient at the slab face of a rain medium is sufficiently small that Q^2 can be neglected. This condition is easily confirmed by direct calculation of Q for the rain medium with the relevant expressions from equations (15a) through (15f). It was determined that both the real and imaginary parts of Q^2 are generally on the order of 10^{-12} . Even for worst-case monodisperse distributions at unrealistically extreme rain rates (several thousand millimeters per hour), the components of Q^2 never exceeded 10^{-9} . Thus, the multiplicative term $\mathcal{B}(1 - Q^2)$ of equation (14) is essentially unity for the rain medium, revealing coherent multiple reflections between interfaces to be negligible.

With the preceding approximations, equation (14) reduces to a form identical to equation (7), except that the bulk refractive index is given by η_{bf} instead of η_f . It now only remains to determine the quantitative importance of the backward multiple scattering term $(2\pi N/k^3)^2 S^2(\pi)$ in equation (17).

Values of η_{bf} and η_f were computed for monodisperse and polydisperse rain distributions at several rain rates, for frequencies between 1 and 1,000 GHz. The differences were very small in all cases, even for the highest rain rates.

As an example, Table 1 provides numerical values of η_{bf} and η_f in refractivity units $(\eta - 1) \times 10^6$ for the polydisperse Laws and Parsons [26] drop size distribution at a rain rate of 150 mm/hr for rain temperatures of 0°C and 20°C, calculated with the expression for the refractive index of water given by Ray [27] and the raindrop terminal velocities of Gunn and Kinzer as quoted by Medhurst [10]. These results reveal that the differences between the corresponding real and imaginary parts of $(\eta_{bf} - 1) \times 10^6$ and $(\eta_f - 1) \times 10^6$ are several orders of magnitude smaller than those caused by a temperature change of 20°C. Even for modeled rain of extreme intensity (5,000 mm/hr), composed of drops 1 mm in diameter (near worst case), the differences between the components of η_{bf} and η_f never exceeded 0.02 percent over the frequency range 1–1,000 GHz.

This comparison assumes that equation (17) is an accurate representation of the bulk refractive index of the medium, although in fact it is an approximation that neglects various backward multiple scattering components. Nevertheless, on the basis of the analysis of Olsen et al. [7], it is believed that the second term of equation (17) adequately accounts for the dominant backward multiple scattering contribution. Furthermore, the preceding numerical comparisons show that even if the backward multiple scattering effects were several orders of magnitude larger than indicated by equation (17), they would still be completely negligible.

Figure 3 illustrates the relative effects of the backscattering amplitude $S(\pi)$ and the term $2\pi/k^3$ on the backward multiple scattering contribution of η_{bf} . In this figure, $|Re(\eta_{bf} - \eta_f)| \times 10^6$ and $|Im(\eta_{bf} - \eta_f)| \times 10^6$ are plotted versus frequency for Laws and Parsons rain at 150 mm/hr and a drop temperature of 20°C. As is evident, the functions $S(\pi)$ and $2\pi/k^3$ offset one another to varying degrees at frequencies below about 30 GHz. Above this frequency, the $2\pi/k^3$ factor becomes dominant, and the already very small backward multiple scattering term in η_{bf} rapidly approaches zero.

Physically, the contribution of the backward multiple scattering processes to the coherent wave propagated through the rain medium is negligible because of the wide range of spacings, in terms of wavelengths, between the sparsely distributed raindrops. Hence, the phases of the backward multiply scattered components are approximately uniformly random, and the scattering processes are therefore effectively incoherent. The same relative incoherence will occur in forward multiple scattering components for which the paths between the scatterers are not aligned closely in the forward direction. It is

TABLE 1. COMPARISON OF MULTIPLE SCATTERING BULK REFRACTIVITIES FOR LAWS AND PARSONS RAIN AT 150-mm/hr RAIN RATE

FREQUENCY (GHz)	TEMPERATURE (°C)	$(\eta_r - 1) \times 10^6$	$(\eta_{br} - 1) \times 10^6$
1	0	8.3618999 - j0.033519870	8.3618654 - j0.033519641
	20	8.3325401 - j0.019890176	8.3325058 - j0.019890036
10	0	8.5851536 - j2.4044025	8.5851143 - j2.4044106
	20	8.6393766 - j2.8490678	8.6393383 - j2.8490815
50	0	2.2469141 - j4.3560448	2.2469115 - j4.3560420
	20	2.2760026 - j4.2207548	2.2759991 - j4.2207520
100	0	0.53816061 - j2.5724631	0.53816051 - j2.5724630
	20	0.58603329 - j2.5592917	0.58603315 - j2.5592915
200	0	0.081415141 - j1.3356303	0.081415137 - j1.3356303
	20	0.062313787 - j1.3501088	0.062313782 - j1.3501088
500	0	-0.037804831 - j0.48342210	-0.037804831 - j0.48342210
	20	-0.037362947 - j0.48212113	-0.037362947 - j0.48212113
1000	0	-0.019825341 - j0.22474476	-0.019825342 - j0.22474476
	20	-0.019906032 - j0.22492716	-0.019906032 - j0.22492716

the scattering interactions for which the paths are closely aligned in the forward direction (thus making the total path length through the medium nearly constant) that contribute predominantly to the coherent transmitted wave.

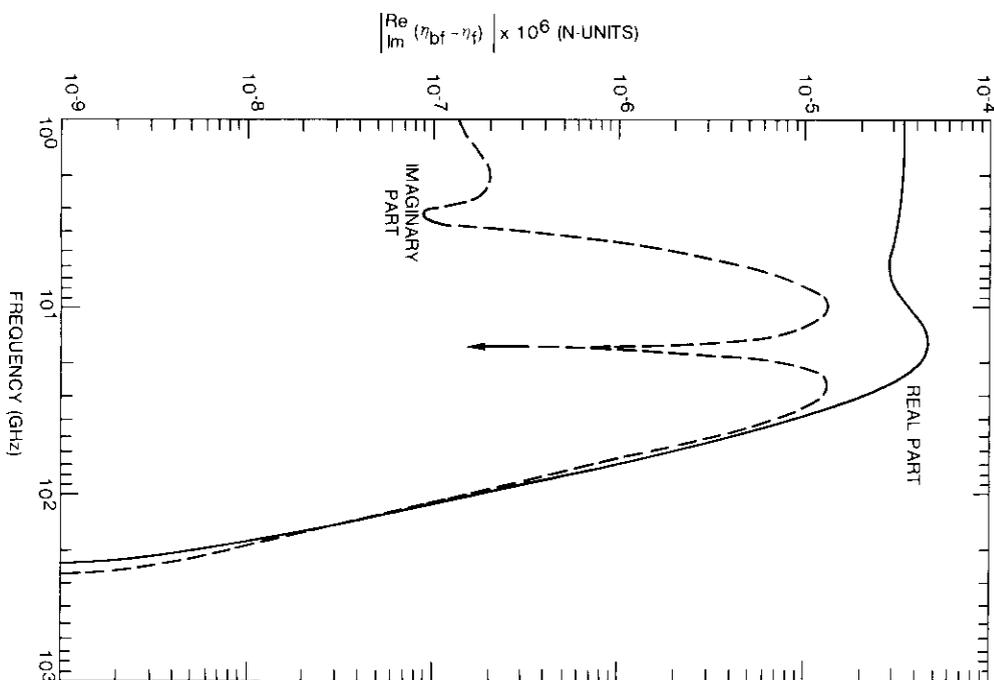


Figure 3. Difference (in N-units) Between η_{br} and η_r for Laws and Parsons Rain of Intensity 150 mm/hr and Temperature of 20°C (The imaginary component changes sign at $f \approx 15$ GHz)

Conclusions

By careful examination of the scattering processes involved in the classical theory for coherent wave propagation through a sparsely populated, uniformly random medium of discrete scatterers, it has been established that this theory accounts for all orders of forward multiple scattering, where forward multiple scattering includes all scattering interactions for which successive paths between scatterers have a component in the forward direction. Identical results are obtained either by retaining forward multiple scattering processes within layers of the scattering region or by applying single scattering theory to successive layers within the medium in the limit as the layer thickness $\Delta d \rightarrow 0$. Thus, it is evident that the microscopic forward multiple scattering processes disappear within individual layers as $\Delta d \rightarrow 0$, but reappear for the overall slab through the macroscopic forward multiple scattering between layers. This equivalence explains why the original derivations of the transmission coefficient $\langle Y \rangle$, which implicitly or explicitly assumed the single scattering approximation within thin slabs, actually incorporate all forward multiple scattering contributions, although this fact has not been widely recognized.

Comparisons between the classical approach and the free-space scattering theory of Twersky revealed the nature of the most significant backward multiple scattering contributions neglected by the classical theory. Numerical calculations showed that these backward multiple scattering contributions are negligible for coherent transmission through rain, even at extreme rain rates, for frequencies up to at least 1,000 GHz. Indeed, the backward multiple scattering contributions are several orders of magnitude less significant than the effects caused by a change of 20°C in rain temperature. It is therefore concluded that the classical approach for computing specific attenuation and phase rotation in the rain medium is quite accurate.

References

- [1] A. Ishimaru, "Theory and Application of Wave Propagation and Scattering in Random Media," *Proceedings of the IEEE*, Vol. 65, 1977, pp. 1030-1061.
- [2] R. L. Olsen, D. V. Rogers, and D. B. Hodge, "The aR^b Relation in the Calculation of Rain Attenuation," *IEEE Transactions on Antennas and Propagation*, AP-26, 1978, pp. 318-329.
- [3] R. K. Crane, "Microwave Scattering Parameters for New England Rain," Technical Report 426, M.I.T. Lincoln Laboratories, Lexington, Massachusetts, 1966.
- [4] R. K. Crane, "Propagation Phenomena Affecting Satellite Communication Systems Operating in the Centimeter and Millimeter Wavelength Bands," *Proceedings of the IEEE*, Vol. 59, 1971, pp. 173-188.
- [5] J. C. Lin and A. Ishimaru, "Propagation of Millimeter Waves in Rain," Technical Report 144, Dept. of Electrical Engineering, University of Washington, Seattle, 1971.
- [6] A. Ishimaru and J. C. Lin, "Multiple Scattering Effects on Wave Propagation Through Rain," *NATO/AGARD Conference Proceedings No. 107*, Brussels, Belgium, 1973, pp. 1/1-1/13.
- [7] R. L. Olsen, M. M. Z. Kharadley, and D. G. Corr, "On Certain Theories of Multiple Scattering in Random Media of Discrete Scatterers," *Radio Science*, Vol. 11, 1976, pp. 21-37.
- [8] D. V. Rogers and R. L. Olsen, "Calculation of Radiowave Attenuation due to Rain at Frequencies up to 1000 GHz," CRC Report 1299, Communications Research Centre, Ottawa, Canada, 1976.
- [9] N. K. Uzunoglu and B. G. Evans, "Multiple Scattering Effects in Electromagnetic Wave Propagation Through a Medium Containing Precipitation," *Journal of Physics, A: Mathematical and General*, Vol. 11, 1978, pp. 767-776.
- [10] R. G. Medhurst, "Rainfall Attenuation of Centimeter Waves: Comparison of Theory and Measurement," *IEEE Transactions on Antennas and Propagation*, AP-13, 1965, pp. 550-564.
- [11] D. C. Hogg and T. S. Chu, "The Role of Rain in Satellite Communications," *Proceedings of the IEEE*, Vol. 63, 1975, pp. 1308-1331.
- [12] H. C. van de Hulst, *Light Scattering by Small Particles*, New York: John Wiley, 1957, pp. 5-6 and pp. 28-36.
- [13] J. W. Ryde and D. Ryde, "Attenuation of Centimetre and Millimetre Waves by Rain, Hail, Fogs, and Clouds," Technical Report 8670, General Electric Company Research Laboratory, Wembley, England, 1945.
- [14] B. C. Blevis, "Microwave Attenuation and Scatter: Theory and Measurement," *Journal de Recherches Atmospherique*, Vol. 8, 1974, pp. 5-21.
- [15] K. Östburg, "Effect of Drop Canting Angle Distribution on Depolarization of Microwaves in Rain," FOA Rep. (Sweden), Vol. 10, 1976, pp. 1-10.
- [16] International Radio Consultative Committee (CCIR) Doc. 5/23 (Canada), "Calculations of Rain Attenuation for Frequencies up to 1000 GHz," Geneva: International Telecommunications Union, 1975.
- [17] P. Beckmann, *The Depolarization of Electromagnetic Waves*, Boulder: Golem Press, 1968, pp. 43-44.
- [18] R. L. Olsen, "A Review of Theories of Coherent Radiowave Propagation Through Precipitation Media of Randomly Oriented Scatterers, and the Role of Multiple Scattering," *Radio Science*, Vol. 17, 1982, pp. 913-928.
- [19] D. E. Kerr, *Propagation of Short Radio Waves*, Vol. 13, M.I.T. Radiation Laboratory Series, New York: McGraw-Hill, 1951, pp. 671-674.

- [20] T. Oguchi, "Statistical Fluctuation of Amplitude and Phase of Radio Signals Passing Through Rain," *Journal of the Radio Research Laboratories (Japan)*, Vol. 9, 1962, pp. 51-72.
- [21] G. C. McCormick, "Propagation Through a Precipitation Medium: Theory and Measurement," *IEEE Transactions on Antennas and Propagation*, AP-23, 1975, pp. 266-269.
- [22] V. Twersky, "On Propagation in Random Media of Discrete Scatterers," Symposium on Applied Physics and Engineering, American Mathematical Society, Providence, R.I., 1964, *Proc.*, pp. 84-116.
- [23] V. Twersky, "On Scattering of Waves by Random Distributions. I. Frec-Space Formalism," *Journal of Mathematical Physics*, Vol. 3, 1962, pp. 700-715.
- [24] V. Twersky, "On a General Class of Scattering Problems," *Journal of Mathematical Physics*, Vol. 3, 1962, pp. 716-723.
- [25] V. Twersky, "On Scattering of Waves by Random Distributions. II. Two-Space Scatterer Formalism," *Journal of Mathematical Physics*, Vol. 3, 1962, pp. 724-734.
- [26] J. O. Laws and D. A. Parsons, "The Relation of Raindrop-Size to Intensity," *Transactions of the American Geophysical Union*, Vol. 24, 1943, pp. 452-460.
- [27] P. S. Ray, "Broadband Complex Refractive Indices of Ice and Water," *Applied Optics*, Vol. 11, 1972, pp. 1836-1844.



David V. Rogers received a B.S. in Physics from Lamar University in 1967 and a Ph.D. in Physics from North Texas State University in 1973. Prior to joining COMSAT Laboratories in 1977, he worked at the Communications Research Centre, Ottawa, Canada, and at ORI, Incorporated, Silver Spring, Maryland. He is presently a Staff Scientist in the Propagation Studies Department at COMSAT, where he performs theoretical and experimental propagation research and analyses, primarily for frequencies above 10 GHz. Dr. Rogers is a member of Study Group 5 of the International Radio Consultative

Committee (CCIR) and Chairman of the U.S. CCIR Ad Hoc Committee on Radiometeorology. He has participated in several international CCIR Study Group 5 meetings. He is also a member of the IEEE.

Roderic L. Olsen received the B.A.Sc. and Ph.D. degrees in Electrical Engineering from the University of British Columbia, Vancouver, Canada, in 1965 and 1970, respectively. Since 1970, he has been employed as a Research Scientist at the Communications Research Centre (CRC), Department of Communications, Ottawa, Canada. His work has been primarily in the area of tropospheric propagation effects on earth-space and terrestrial communications at centimeter and millimeter wavelengths. He currently heads the Terrestrial Micro-



wave Propagation Group at CRC and is the Canadian national chairman of Study Group 5 of the International Radio Consultative Committee (CCIR). He has been a Canadian delegate to international CCIR meetings since 1977 and to the 1979 World Administrative Radio Conference. He is a member of the IEEE.

Rain and ice depolarization measurements at 4 GHz in Sitka, Alaska*

S. J. STRUHARIK

(Manuscript received May 16, 1983)

Abstract

Rain and ice depolarization effects and their variation with climate and path geometry are important factors in satellite system planning. This paper presents the results of depolarization measurements conducted at a low elevation angle of 11.7° in Sitka, Alaska (57°N latitude), which has a maritime climate with appreciable rainfall and snowfall. The slant path used was from the INTELSAT Pacific Ocean Region (POR) satellite at 174°E longitude. Cross-polarization discrimination (XPD) was determined for a 4-GHz continuous wave (CW) signal transmitted by the satellite. Radiometric sky noise temperature was measured concurrently at 11.6 GHz along the same path. Point precipitation was measured at ground level. In addition to providing attenuation information, the radiometer data were also used to distinguish between rain and ice depolarization effects.

Cumulative distributions of XPD showing separate and combined effects of rain and ice are discussed, along with attenuation and rain rate statistics. The 0.01-percent time levels of XPD attributed to rain only and to ice only were 20 and 22 dB, respectively. The lowest observed XPD was below 18 dB and occurred on two separate occasions, one attributed to rain depolarization and the other attributed to ice depolarization. It was determined that ice effects had a varying impact on the XPD statistics but were negligible at time percentages below about 0.01. The results are compared to current propagation models and literature and the effect of the low altitude of the local melting layer on the data is discussed.

* This paper is based on work performed under the sponsorship of the International Telecommunications Satellite Organization (INTELSAT). Views expressed are not necessarily those of INTELSAT.

Introduction

Frequency reuse by dual orthogonal polarizations has been implemented at 4 and 6 GHz in the INTELSAT system, as well as in several domestic satellite systems. In such systems, the polarization isolations of the satellite and earth station antennas and the depolarization effects in the propagation medium influence the coupling of energy between the two polarizations and thus the system carrier-to-interference (C/I) ratios. Characterization of the various depolarizing effects is important in satellite system planning, and much work has been done in this area [1].

Both ice crystals and rain along the slant path are known to produce degradations in XPD. One current consideration in the modeling of these effects is how ice depolarization may best be incorporated into propagation models. Measured data from diverse climates and locations are useful in testing this and other aspects of current models.

This paper presents the results of depolarization measurements conducted at 4 GHz in Sitka, Alaska, over a 12-month period beginning in 1980. Depolarization data collected in the northern Pacific region are scarce and are of interest by themselves [2], [3]. Because of the site location and the climate, it was expected that a significant amount of the observed depolarization would be caused by ice crystals along the slant path, a fact that proved to be true in the course of the data analysis.

The experimental results are presented, with emphasis on the rain vs ice crystal depolarization effects observed in the data. These results include cumulative percentage time exceedance statistics and examples of individual events. Comparisons are made to CCIR models and to the results of other experiments in different climates and geographical locations. Some inferences are drawn regarding the behavior of the melting layer, or zero degree isotherm, in the Sitka climate. Antenna measurements, hardware parameters, and the mathematics of the data processing algorithms are also presented as reference material that may be useful in the implementation of future XPD measurements.

Measurement configuration

A diagram of the overall measurement configuration is shown in Figure 1. A general discussion of the experiment is also presented in Reference 4.

Site description

The experiment site was located in Sitka, Alaska, which lies on the Pacific Coast in the southeastern part of the state. The site was situated only a few hundred meters from the water, with the slant path to the satellite looking

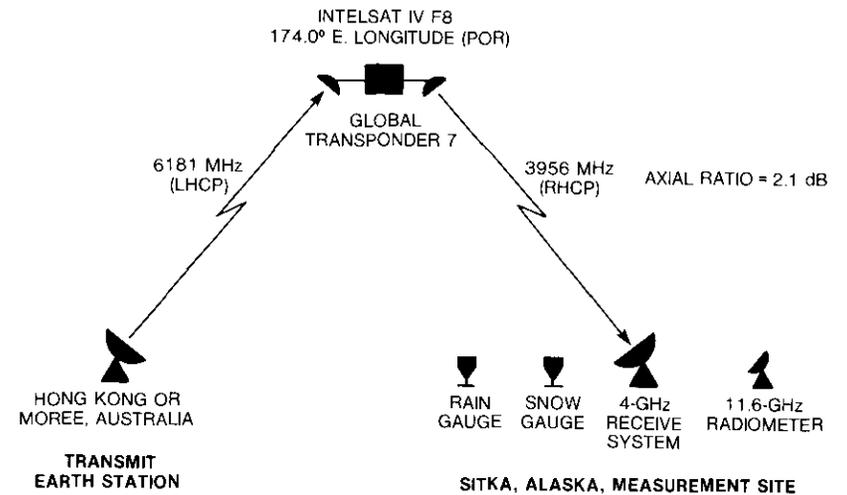


Figure 1. Measurement Configuration

directly out over Sitka Sound and the Pacific Ocean to the southwest at an elevation angle of 11.7° . The local weather pattern is such that storms tended to approach almost directly down the path, coming off the ocean and between mountains on either side of the path. The Sitka climate may be characterized as temperate maritime, with a mean annual rainfall of 2,286 mm [5], an annual snowfall of 1,016 mm [5], and a thunderstorm ratio, β , of 0.2 [6]. Table 1 gives details of the location, climate, and antenna pointing for the experiment site.

TABLE 1. CHARACTERISTICS OF THE EXPERIMENT SITE

Location	Sitka, Alaska (airport)
Latitude	57.05°N
Longitude	135.36°W
Elevation Above Sea Level	9 m
Slant Path Azimuth Angle	235.8°
Slant Path Elevation Angle	11.7°
Mean Annual Rainfall, M	2,286 mm
Mean Annual Snowfall	1,016 mm
Thunderstorm Ratio, β	0.2

The Sitka site was selected because of the high northern latitude, the low elevation angle, and the appreciable snowfall and rainfall characteristics. Low elevation angles represent a worst-case satellite system condition. The

portion of the slant path passing through the troposphere is longer, and thus more influenced by weather conditions, than slant paths at higher elevation angles. Also, low elevation angles imply that the particular site would be at the edge of satellite coverage of the earth, with correspondingly reduced satellite e.i.r.p. resulting from satellite antenna effects. The snowfall and rainfall characteristics, together with the latitude of the site location, indicated a likelihood that both rain depolarization and ice crystal depolarization effects would be present in the data base.

Space segment

The space segment of the experiment was provided by the INTELSAT POR primary path satellite, which during the measurements was INTELSAT IV F8, located at 174.0° east longitude. An unmodulated, left-hand circularly polarized (LHCP) carrier was transmitted to the satellite at 6,181 MHz by INTELSAT earth stations, initially from Moree, Australia, and subsequently from Hong Kong. The resulting 3,956-MHz down-link CW signal was in the global beam of the satellite, right-hand circularly polarized (RHCP) with an e.i.r.p. in the direction of Sitka of about 9 dBW. Measurements at the site established the voltage axial ratio of the arriving wave to be approximately 2.1 dB.

A link budget is shown in Table 2. The antenna used in the receive system at Sitka could be configured for circular or linearly polarized operation.

TABLE 2. LINK BUDGET

Spacecraft e.i.r.p. (RHCP, 3,956 MHz)	+ 12 dBW (beam center)
Spacecraft Antenna Pattern	- 3 dB (edge of coverage)
Free Space Path Loss	- 196.7 dB
Pointing Loss	- 1 dB
Receive System G/T	+ 22 dB/K
10 log k	- (-228.6) dBW/Hz-K
C/kT	61.9 dB/Hz
Receive System Bandwidth (10 log 150 Hz)	- 21.8 dB-Hz
C/N	40.1 dB

The carrier-to-noise (C/N) values measured at the site for both modes of operation are as follows:

a. Linearly Polarized Mode

$$C/N = 26 \text{ dB}$$

b. Circularly Polarized Mode

$$\text{Copolared (RHCP) } C/N = 29 \text{ dB}$$

$$\text{Cross-polarized (LHCP) null} = 39 \text{ dB}$$

As discussed later, in the linearly polarized mode the antenna was adjusted to deliver equal power to both receive channels. In the circularly polarized mode, the antenna polarization was matched to that of the arriving wave, thus delivering all signal power to the copolarized channel and yielding a 3-dB C/N improvement. The 39-dB cross-polarized null was the ratio of copolarized to cross-polarized signal levels in the circularly polarized mode. It was possible to achieve a null that exceeded the C/N because, in matching antenna polarization to that of the wave, the noise transmitted by the satellite (up-link thermal and intermodulation) was now all delivered to the copolarized channel. This lowered the noise floor of the cross-polarized channel, with down-link thermal noise being the only remaining contributor, and thus extended the dynamic range. The link budget shown also includes only down-link thermal noise and is in agreement with the measured cross-polarized null.

Data collection system

The data collection system in Sitka measured rainfall, snowfall, 11.6-GHz radiometric sky noise temperature, and orthogonal polarization components of the received 4-GHz signal. Data were recorded by a strip chart recorder. Figure 2 is a block diagram of the 4-GHz receive system.

A 4.5-m Cassegrain antenna was used, with a reflector and feed horn of the same type used in the RCA ALASCOM network. The feed network was of the type described in References 7 through 9 and allowed the axial ratio and tilt angle of the arriving wave to be modified prior to measurement. It consisted of two cascaded, independently rotatable polarizers and an orthogonal mode transducer (OMT). Each polarizer contained a phase shifting network that imparted a phase shift (180° and 90°, respectively) to the linear component of the field inside the waveguide that was aligned with the phase shift plane of the polarizer. The OMT coupled two linearly polarized orthogonal components of the resultant wave to the RF system.

The RF system consisted of two identical down-conversion chains. Two GaAs FET LNAs were mounted on the OMT at the antenna. The RF signals were routed via low-loss coaxial cables to the equipment shelter, where down-conversion occurred in two steps. The output of the second RF down-converter was phase-locked to a reference signal from the two-channel phase and amplitude IF receiver (Scientific Atlanta Model 1753). The outputs of the receiver and its associated logarithmic amplifiers were voltages proportional to the individual relative amplitudes of the two input signals and to the phase difference between them.

Two different modes of data collection were used to measure 4-GHz XPD. During the first 6 months of the experiment, spatially orthogonal, linearly

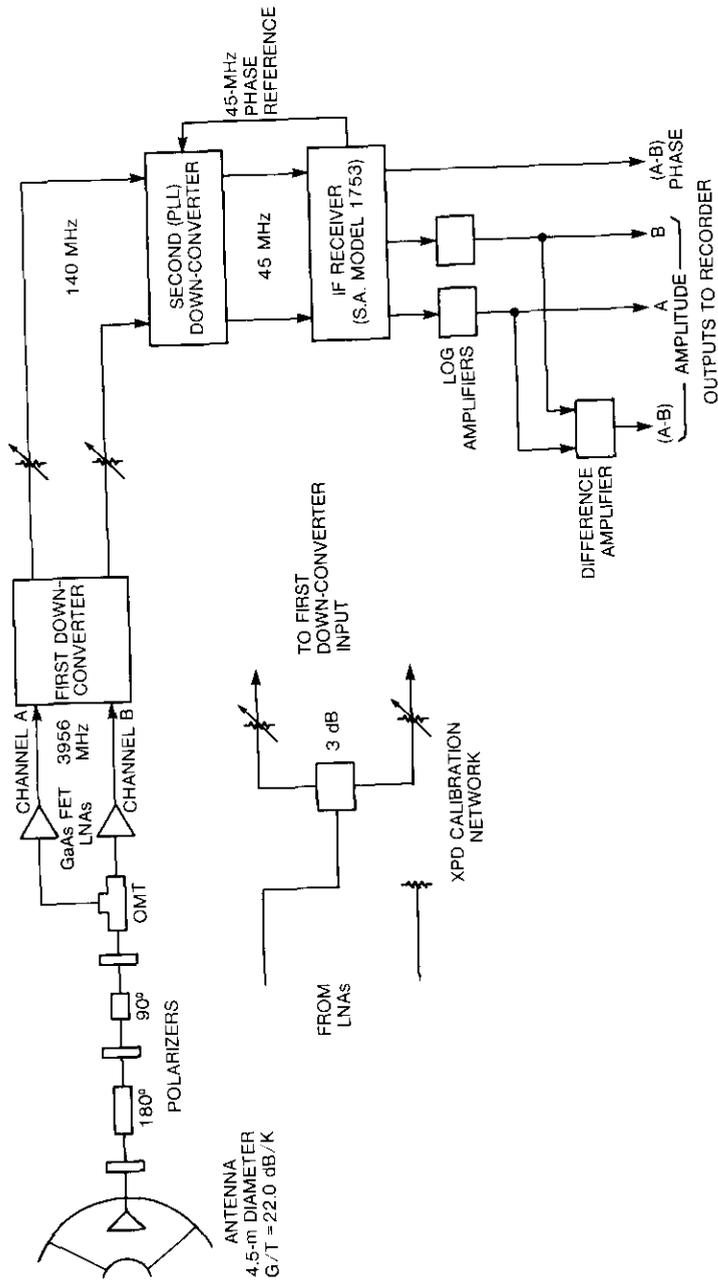


Figure 2. System Block Diagram

polarized components of the received wave were measured. The individual amplitudes and relative phase difference of the components were recorded. From these, transmission-medium-induced XPD data were calculated. During the second 6 months, the system was modified slightly; the antenna feed was readjusted so that the transmission-medium-induced XPD could be measured and recorded directly. The two measurement modes and feed network adjustments will be discussed further.

Ancillary measurement equipment consisted of two mechanical tipping bucket rain gauges, one of which was heated so that it responded to both snowfall and rainfall, and an 11.6-GHz radiometer of the type described in Reference 10. This equipment was used to assist in identifying the source of the observed depolarization, as well as to measure secondary propagation parameters. Specifically, the rain and snow gauges, in addition to measuring rain rate and accumulation, permitted the identification of liquid vs snow or ice precipitation at ground level. The 11.6-GHz radiometer measured radiometric excess sky noise temperature, from which excess attenuation was calculated. It also allowed a distinction to be made between rain and ice crystal depolarization in the 4-GHz data. The radiometer responded to the presence of liquid water (rain or wet snow) along the slant path by registering an increase in 11.6-GHz sky noise temperature. Ice crystals along the path produced no response in the radiometer, since they introduced negligible absorption. The radiometer was used to sense attenuation because it was a much more sensitive indicator than the 4-GHz copolarized signal level. An excess radiometric attenuation level of 1 dB was chosen as the threshold between rain and ice depolarization. That is, observed depolarization data (events or portions of events) occurring while the radiometric data indicated excess attenuations of less than 1.0 dB were attributed to ice depolarization. The remainder were presumed to be caused by liquid water along the slant path.

Measurement technique

Antenna parameters

The physical configuration of the feed system is shown in Figure 3. The angular positions of the feed horn and OMT were fixed, while the polarizers could each be rotated. Figure 4 shows the field components, parameters, and sign conventions associated with the feed network, as viewed looking toward the satellite (*i.e.*, wave approaching). All angles are referenced to local horizontal.

The input to the network is taken as the feed horn output flange. Vectors

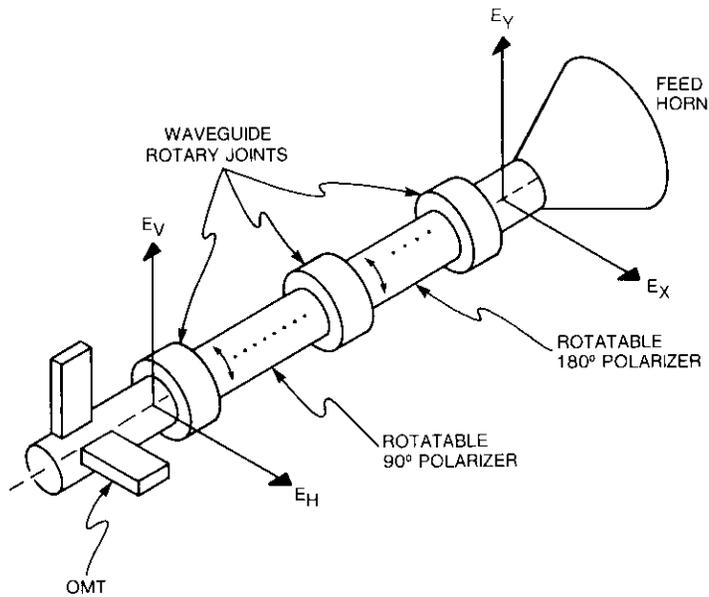


Figure 3. Physical Configuration of the Antenna Feed System

E_1 and E_2 are the field components at this point that are aligned with the semi-major and semi-minor axes of the wave polarization ellipse, the tilt angle of which is τ . Vectors E_X and E_Y are the field components aligned with local horizontal and vertical at the same point.

The phase shift planes of the polarizers are oriented at angles α and β . In each polarizer, the component lying in the phase shift plane experiences a net phase lag with respect to the component orthogonal to the plane, so that $\theta = -180^\circ$ and $\phi = -90^\circ$.

The output of the network is taken as the OMT input flange, between the OMT and the 90° polarizer. Vectors E_H and E_V are the components at this point that are aligned with the OMT output ports. The OMT is canted slightly at an angle γ .

The effect of the feed network on the arriving wave can be expressed as a matrix multiplication:

$$\begin{bmatrix} E_H \\ E_V \end{bmatrix} = [R(\beta - \gamma)] [S(\phi)] [R(\alpha - \beta)] [S(\theta)] [R(-\alpha)] \begin{bmatrix} E_X \\ E_Y \end{bmatrix} \quad (1)$$

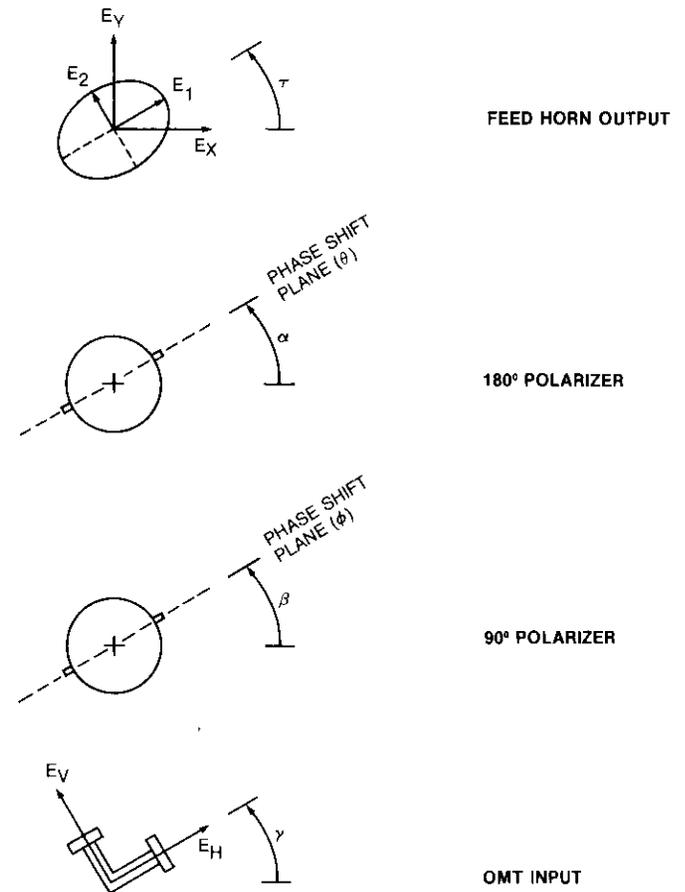


Figure 4. Antenna Feed Parameter Definitions and Sign Conventions

where

$$\begin{bmatrix} E_X \\ E_Y \end{bmatrix} = [R(\tau)] \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}$$

$$[R(\chi)] = \begin{bmatrix} \cos \chi & -\sin \chi \\ \sin \chi & \cos \chi \end{bmatrix}$$

$$[S(\psi)] = \begin{bmatrix} e^{j\psi} & 0 \\ 0 & 1 \end{bmatrix}$$

The matrix $[R(\chi)]$ provides coordinate rotations, and $[S(\psi)]$ allows the insertion of phase shifts. The values of the arguments χ and ψ depend on where the operations are applied. Since only the ratio $|E_1|/|E_2|$ is known, E_1 is taken to be unity, and, for RHCP,

$$E_2 = -j 10^{-0.05AR} \tag{2}$$

where AR is the voltage axial ratio of the arriving wave in decibels.

Similarly, the axial ratio and tilt angle of the arriving wave can be calculated from E_H and E_V by using the inverse of equation (1) to obtain E_X and E_Y . In this calculation, only the relative amplitudes and relative phase of E_V and E_H are known. E_H is taken as unity, and E_V is then

$$E_V = (|E_V|/|E_H|)e^{j\epsilon} \tag{3}$$

where ϵ is the phase of E_V with respect to E_H . The following expressions are then applied to the result [11], [12]:

$$\tau = 0.5 \tan^{-1} \frac{2|E_X||E_Y| \cos \delta}{|E_X|^2 - |E_Y|^2} \tag{4}$$

$$AR_{dB} = 20 \log \{ \tan^{-1} [0.5 \sin^{-1} (\sin [2 \cdot \tan^{-1} (|E_Y|/|E_X|) \sin \delta])] \} \tag{5}$$

where δ is the phase of E_Y with respect to E_X .

Equation (4), which contains an arctangent function, is not single valued and will always yield values of τ between $\pm 45^\circ$. If τ (defined here as the angle between local horizontal and the semi-major axis of the ellipse) is within these bounds, the expression yields the correct value. If τ is outside these bounds, the expression yields the angle between local horizontal and the semi-minor axis, and this angle must be corrected accordingly by adding or subtracting 90° . The need to apply the correction can be recognized by examining the ratio $|E_Y|/|E_X|$. Regardless of their phase difference, $|E_X|$ and $|E_Y|$ define a box which must bound the ellipse [11]. If $|E_Y|/|E_X| > 1$, the box is vertically oblong, τ is outside the bounds of $\pm 45^\circ$, and the correction must be applied.

A basic problem presented by the experiment is the separation of propagation effects from antenna effects. An exact formulation of their interaction is complex. However, an approximation of their interaction is possible with good accuracy under certain conditions [7]. This approximation consists of adding vectorially the axial ratios of the individual antennas and the medium

to arrive at the net observed axial ratio and the corresponding XPD. The accuracy of the result is within 0.35 dB of the true XPD if the individual axial ratios are better than 3 dB, which was the case during the Sitka measurements.

Figure 5 illustrates the approximation as applied to the experiment data. The length of each vector is equal to the axial ratio in decibels of the

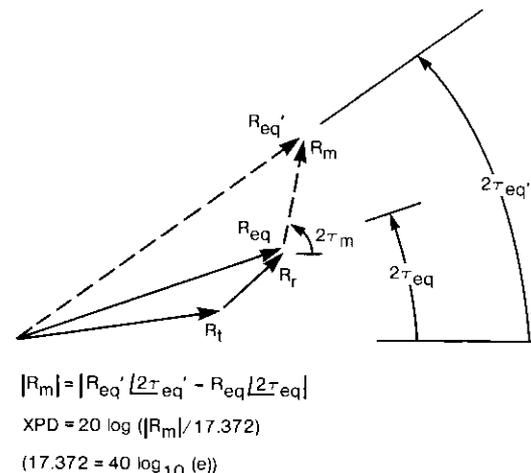


Figure 5. Calculation of Medium-Induced XPD From the Vector Combination of Axial Ratios

particular antenna (or to the equivalent axial ratio of the propagation medium). The angle of each vector is equal to twice the tilt angle of the particular polarization ellipse. The vectors R_t and R_r are the contributions of the transmit and receive antennas, respectively. Under clear-sky conditions, their sum yields a net value of R_{eq} . In the presence of depolarization along the path, a third vector, R_m , appears, which represents the effect of the propagation medium. The addition of R_m changes the observed axial ratio and tilt angle of the arriving wave into R_{eq}' and τ_{eq}' . Given the values of R_{eq} and τ_{eq} observed under clear-sky conditions and R_{eq}' and τ_{eq}' observed during a propagation event, the medium-induced XPD can then be calculated by using the equations shown in the figure.

XPD derived from orthogonal linear wave components

The first 6 months of depolarization data were gathered by measuring the linear orthogonal components of the arriving wave after it had been modified

by the feed network ($|E_H|$, $|E_V|$, and ϵ) and deriving from these components the medium-induced XPD, as discussed above.

In this mode, the feed network was adjusted in an attempt to achieve a 0-dB axial ratio at the OMT input, which would have simplified the data analysis. The resulting angular positions of the feed hardware were $\alpha = 1.5^\circ$, $\beta = 12.0^\circ$, and $\gamma = 3.5^\circ$. The attempt was not fully successful, however, because of hardware constraints during system installation. Of the two conditions necessary for 0-dB axial ratio (E_H and E_V equal in amplitude and in phase quadrature), the first was achieved but the second was not. To maintain consistency in the data, the angular positions were kept at these values for the first 6 months of the experiment.

Direct XPD measurements

The second 6 months of depolarization data were gathered by using the polarization-matched direct isolation measurement method [7]. In this technique, the polarizers are adjusted so that under clear-sky conditions, the receive antenna polarization is matched to that of the arriving wave at the copolarized (RHCP) OMT port. At the cross-polarized (LHCP) OMT port, a polarization then exists that is orthogonal to that of the arriving wave. Effectively, the polarization isolations of the satellite and receive antennas have been made to cancel each other out under clear-sky conditions. Departures from this null condition are then attributable to the appearance of depolarization in the transmission medium. The medium-induced XPD is measured directly by measuring the power ratio between the signals in the OMT ports.

Again, in Figure 5, the polarizers have been adjusted so that R_i and R_r cancel, leaving a very small clear-sky R_{eq} (ideally zero). During a propagation event, the net observed axial ratio, R'_{eq} , is essentially that of the propagation medium, R_m , with an uncertainty determined by how small an R_{eq} was achieved under clear-sky conditions.

Figure 6 shows the bounds of uncertainty placed on the data by the 39-dB cross-polarized null. The curves were generated by using the vectorial addition of axial ratios discussed earlier to calculate the best-case and worst-case combinations of the clear-sky XPD (39 dB) with a range of assumed medium-induced XPD values.

With the experiment reconfigured into this mode, phase measurements between the copolarized and cross-polarized signals were beyond the capabilities of the equipment because of noise effects and the large level difference between the two signals. Consequently, tilt angle information was no longer available. The angular positions of the feed network were $\alpha = 88.6^\circ$, $\beta = 131.2^\circ$, and $\gamma = 3.5^\circ$.

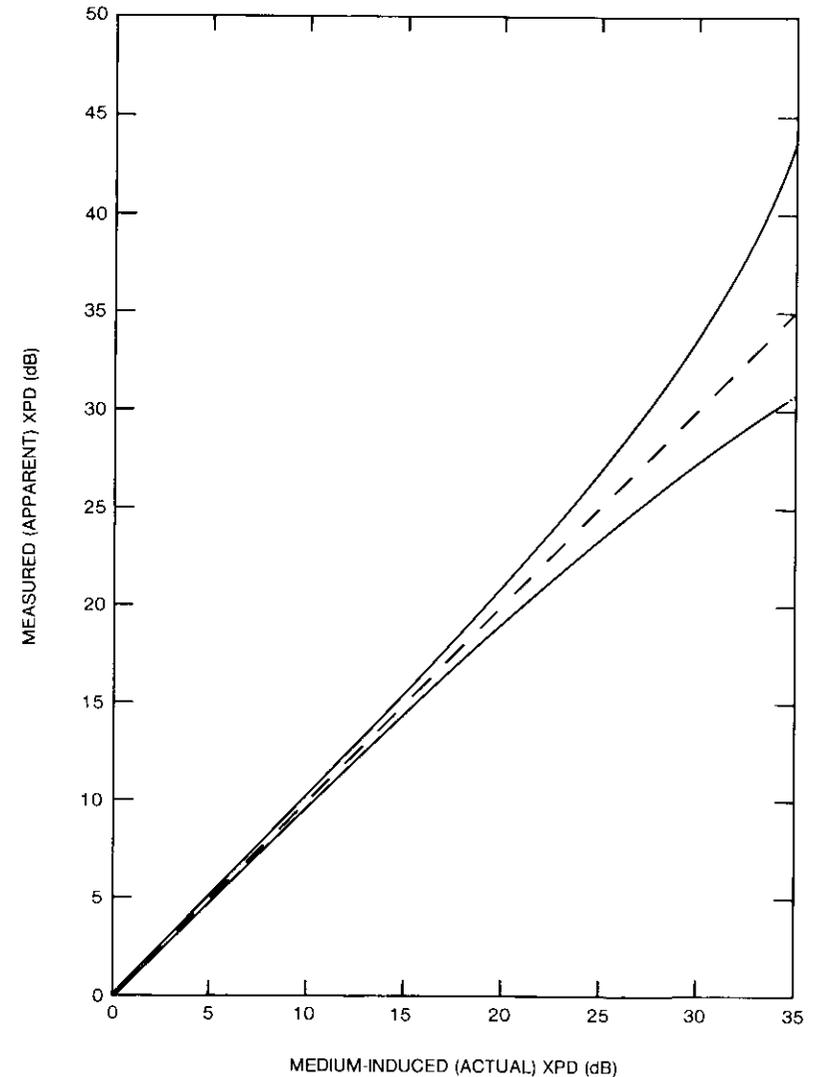


Figure 6. Uncertainty Bounds on Measured Medium-Induced XPD for Clear-Sky Isolation = 39 dB

Antenna measurements

The voltage axial ratio and tilt angle, τ , of the arriving wave, as seen at the feed horn output, were measured by aligning the 90° polarizer with the horizontal OMT port (taking the polarizer out of the circuit by setting $\beta = \gamma$) and then rotating the 180° polarizer through a full 360° while observing amplitude variations in the vertical and horizontal channels. This caused the polarization ellipse to rotate at twice the angular rate of the polarizer, with the peak-to-peak change in either channel being equal to the axial ratio. Several such measurements yielded an axial ratio of 2.1 dB.

The tilt angle of the wave was measured by noting the angular positions of the polarizer (α) at which the maxima and minima occurred, and calculating the corresponding value of τ . The basic principle is that, in passing through the polarizer, E_1 and E_2 of the input ellipse are each reflected spatially about a line normal to the polarizer phase shift plane because their components lying in the plane have been multiplied by -1 (acquiring a 180° phase lag). The average of several such measurements was $\tau = 40^\circ$.

When the system was reconfigured into the direct isolation measurement mode, a cross-check was made of the axial ratio measurements. The polarizers were set for perfect circular polarization (receive antenna axial ratio = 0 dB at $\beta = 135^\circ + \gamma$). The polarization isolation observed was 18 to 19 dB. This corresponds to an axial ratio in the arriving wave of 2.19 to 1.95 dB, which is in good agreement with the nominal value of 2.1 dB measured previously.

Precipitation and radiometer data

Rain rate and total rain accumulation were derived from the tipping bucket rain gauge. Each gauge tip represents an accumulation of 0.254 mm (0.01 in.) of rain. Rain rates are calculated from the time interval between tips; rain accumulation is calculated from the total number of tips. Snow gauge data were used only for examining events in other data channels.

The radiometer output is proportional to the 11.6-GHz antenna noise temperature, which includes the noise temperature of the system electronics, as well as noise arriving via the main, side, and back lobes of the antenna. The excess sky noise temperature, however, is that which occurs over and above the antenna noise temperature observed during clear-sky conditions and is presumed to be due to increases in sky noise temperature from the presence of attenuation along the slant path, *i.e.*,

$$T_e = T_s - T_{cs} \quad (6)$$

where T_e = excess sky noise temperature (K)
 T_s = antenna noise temperature (K) observed during an event
 T_{cs} = antenna noise temperature (K) observed during clear-sky conditions.

The corresponding excess attenuation levels are calculated by using the well-known equation

$$A_e = 10 \log \frac{T_m - T_{cs}}{T_m - T_s} \quad (7)$$

where A_e = excess attenuation (dB)
 T_m = effective medium temperature (K).

For Sitka, the effective medium temperature was assumed to be 275 K. The observed clear-sky antenna noise temperature was 58 K. The system was calibrated by using ambient temperature and liquid nitrogen (80 K) loads applied to the antenna feed.

Results and comparisons to model predictions

Examples of recorded data

Figure 7 shows data recorded during a large rain depolarization event that was measured by the linear orthogonal component technique. This event occurred on September 8, 1980, between 1647 and 1751 GMT. The minimum XPD observed was 18.1 dB, occurring at 1705 GMT. At this minimum, the differential phase had departed from clear-sky conditions by $+10.3^\circ$. At the same time, the amplitudes of E_V and E_H had changed by $+0.74$ dB and -0.60 dB, respectively. The radiometer registered a peak excess attenuation of 4.80 dB at 1742 GMT. During the event, a fairly steady rainfall occurred, averaging 9.1 mm/hr between 1708 GMT and the end of the event. The peak rain rate was about 22 mm/hr and occurred at about 1742 GMT, coincident with the peak radiometric attenuation. The total rainfall accumulated during the event was 7.9 mm.

Four inflections are visible in the phase channel as the event progresses. The XPD values associated with these, along with the time-coincident values of radiometric excess attenuation, are listed in Table 3.

Figure 8 shows the change from the clear-sky conditions of the arriving polarization ellipse at the time of the 18.1-dB minimum XPD. At the output flange of the feed horn, the observed ellipse changed from the clear-sky value of $AR = 2.10$ dB at $\tau = 40^\circ$ to a value of $AR = 4.13$ dB at $\tau = 47^\circ$. This change was the result of the addition to the clear-sky values of

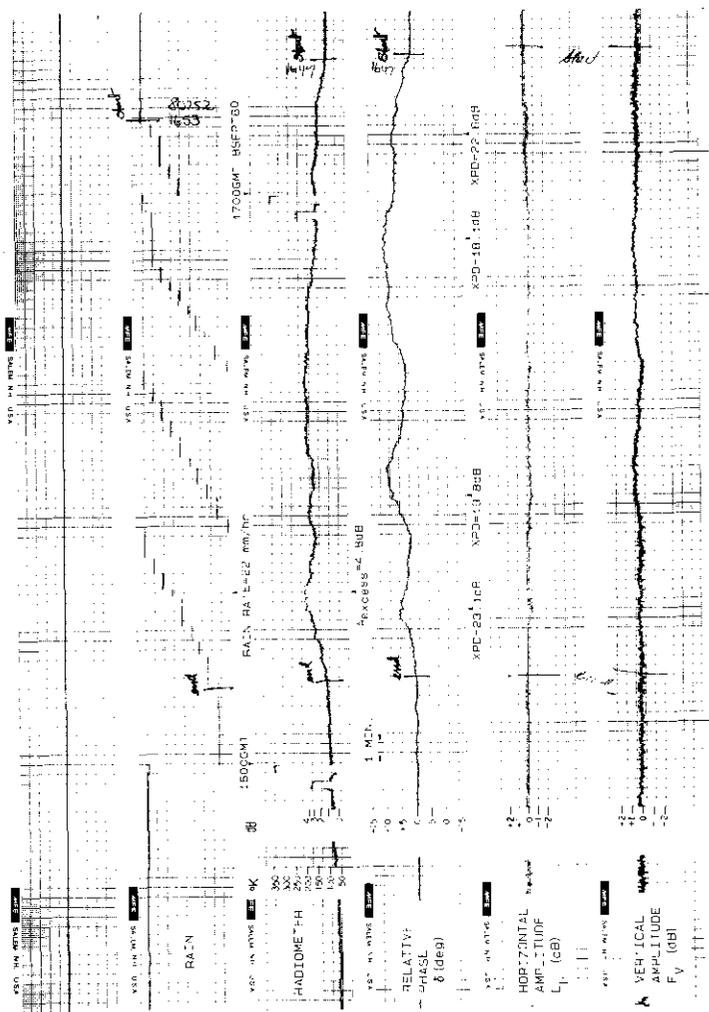
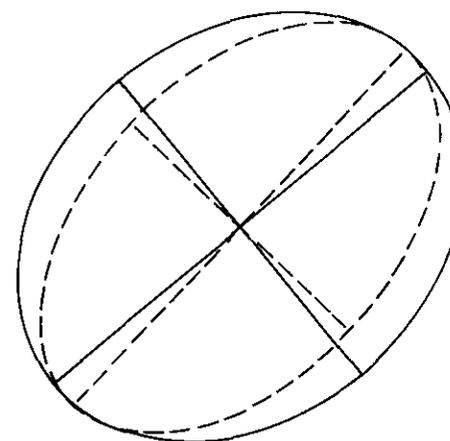


Figure 7. Rain Depolarization Event Data Recorded September 8, 1980

TABLE 3. MEASURED VALUES OF XPD AND RADIOMETRIC ATTENUATION FOR RAIN DEPOLARIZATION EVENT ON SEPTEMBER 28, 1980

TIME (GMT)	4-GHz XPD (dB)	11.6-GHz EXCESS ATTENUATION (dB)
1654	22.6	1.73
1705	18.1	2.05
1732	19.8	3.31
1744	23.1	4.18



— AXIAL RATIO = 2.1 dB, $\tau = 40^\circ$ (CLEAR SKY)
 - - - AXIAL RATIO = 4.13 dB, $\tau = 47^\circ$
 (EVENT, TRANSMISSION MEDIUM XPD = 18.1 dB)

Figure 8. Change in Polarization Ellipse Observed During the Event in Figure 7

a transmission medium contribution of $AR = 2.16$ dB (corresponding to $XPD = 18.1$ dB) at $\tau = 53.5^\circ$.

Figure 9 shows an example of an ice depolarization event. These data were recorded in the direct XPD measurement mode on November 27, 1980, between 0330 and 0453 GMT. The minimum XPD observed was 20 dB. This minimum occurred three times during the event, and each time the excess radiometric attenuation was below 1 dB. During most of the event, the

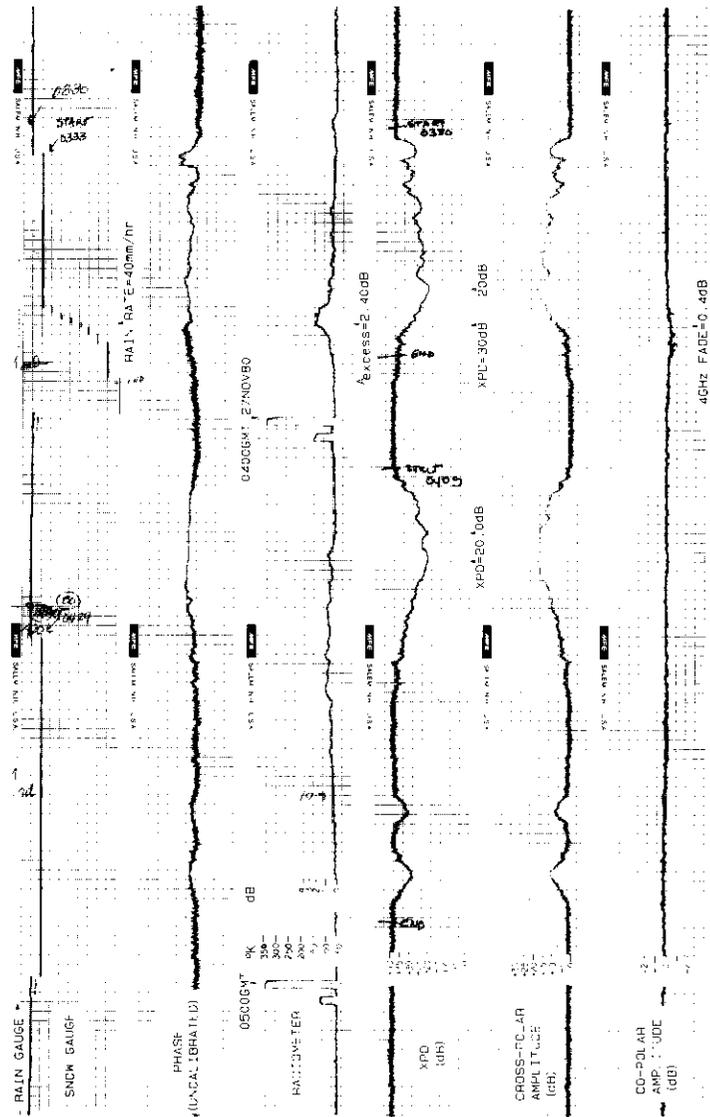


Figure 9. Ice Depolarization Event Data Recorded on November 27, 1980

radiometer registered excess attenuations ranging from 0.25 to 0.75 dB. The peak excess attenuation observed by the radiometer was 2.40 dB at 0350 GMT. This occurred at the same time as the peak rain rate at the site (40 mm/hr) and was accompanied by a fade of about 0.4 dB in the 4-GHz copolarized signal.

The rainfall measured at the site was of short duration and fairly high intensity. The bulk of the rain occurred over a 4-minute period from 0348 to 0352 GMT. The average rain rate over this period was 27 mm/hr, and the peak rain rate was 40 mm/hr. The total rain accumulation during the entire event was 2.5 mm. During the event, both the rain and snow gauges were active and in fair agreement, indicating rainfall rather than snowfall at ground level. The activity of the rain and snow gauges occurred well after the onset of the event, which is characteristic of the local weather pattern of storms tending to approach off the ocean from down the slant path.

Some of the measured XPD values, along with the time-coincident values of radiometric attenuation, are given in Table 4.

TABLE 4. MEASURED VALUES OF XPD AND RADIOMETRIC ATTENUATION FOR ICE DEPOLARIZATION EVENT ON NOVEMBER 27, 1980

TIME (GMT)	4-GHz XPD (dB)	11.6-GHz Excess ATTENUATION (dB)
0346	20.0	0.59
0350	30.0	2.40
0412	20.0	0.15
0415	20.0	0.70

Rain rate statistics

Figure 10 shows the cumulative distribution of rain rates observed during the experiment. The annualized rain accumulation was 2,409 mm. Distributions predicted by the Rice-Holmberg model [13] for the annualized accumulation are plotted for values of $\beta = 0.01$, which gives the best fit to the measured data, and for values of $\beta = 0$ and $\beta = 0.02$ for comparison. The distribution given in Reference 14 for the Sitka location (CCIR Climate Type D) is also plotted. The measured data are in fair agreement with the long-term average annual accumulation of $M = 2,286$ mm for the site, and they fit the CCIR data and the Rice-Holmberg model fairly well. They are, however, at some variance with the published value of $\beta = 0.20$ for the site [6]. Although no conclusions about long-term β values can be drawn from

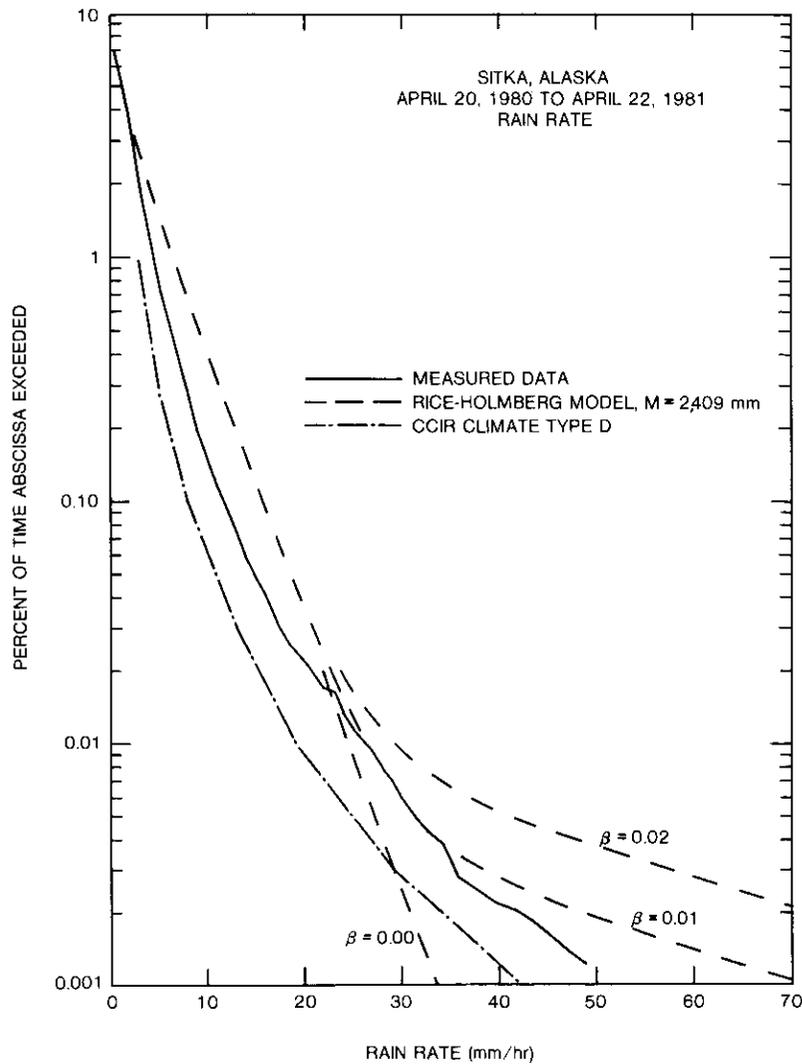


Figure 10. Cumulative Distributions of Rain Rate

a 1-year experiment, it is interesting to note that an earlier estimate of β for the Sitka region is 0.05 to 0.10 [13], which is closer to the rain characteristics observed in this experiment.

Overall, the comparison of measured vs expected values of M and β indicates that the rain characteristics during the experiment were typical in

terms of accumulation, but, due to the low observed β , the proportion of rain that occurred at high rain rates was smaller than expected.

Radiometric attenuation statistics

Figure 11 shows the cumulative distribution of 11.6-GHz excess sky noise temperature measured by the radiometer, and Figure 12 shows the corresponding cumulative distribution of excess attenuation for the assumed effective medium temperature (T_m) of 275 K. The corresponding distribution for $T_m = 220$ K and distributions predicted by a COMSAT model [15] and by a CCIR model [16] are also plotted. The CCIR model contains a latitude-

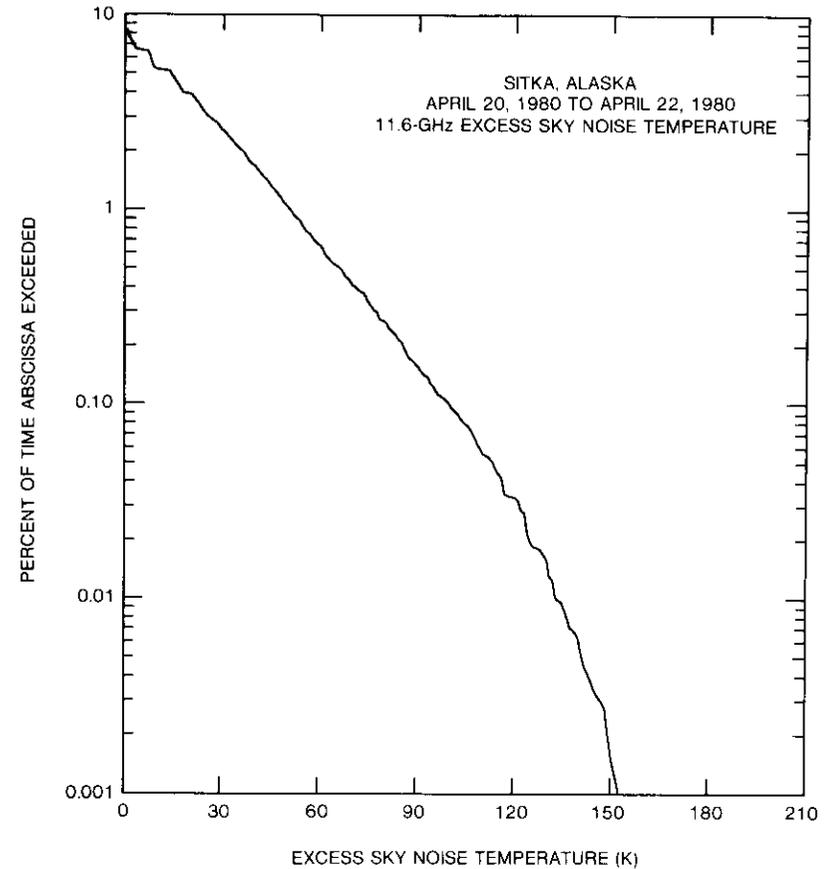


Figure 11. Cumulative Distribution of 11.6-GHz Excess Sky Noise Temperature

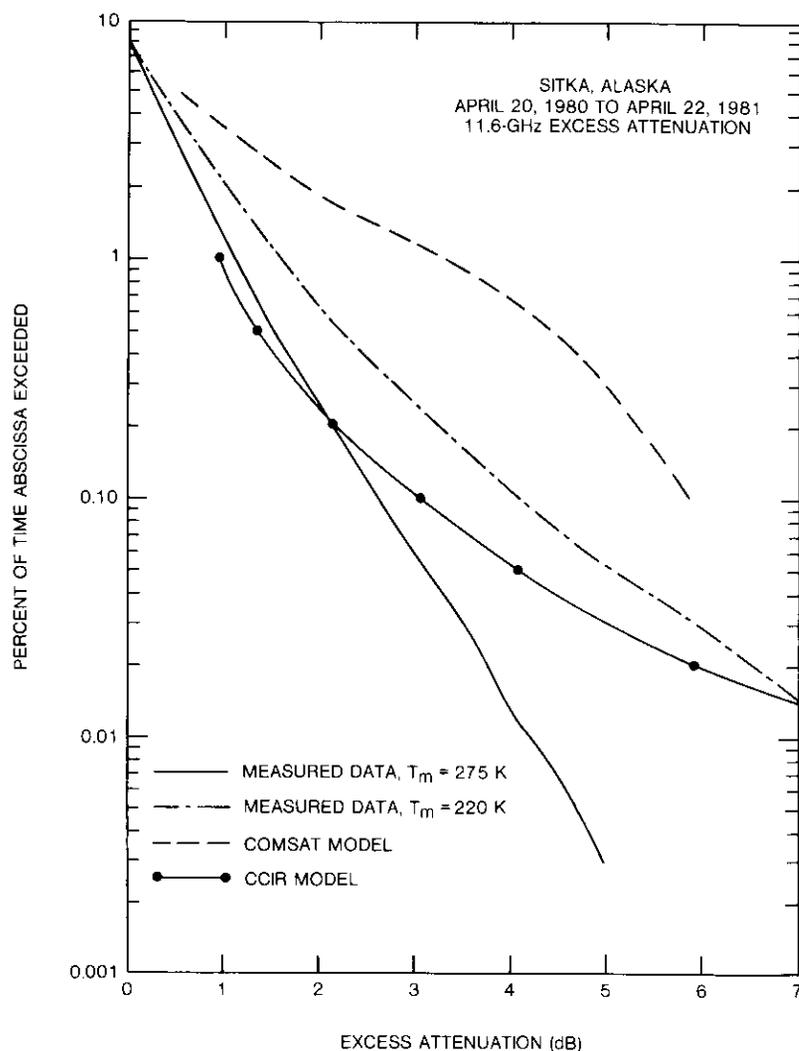


Figure 12. Cumulative Distributions of 11.6-GHz Excess Attenuation

dependent correction for the height of the 0° isotherm (melting layer); the COMSAT model does not. Both models utilize measured rain rate data.

As shown in Figure 12, significant disagreement exists at lower time percentages between the measured attenuation for $T_m = 275$ K and that predicted by either model. This may be due to a combination of two factors. First, the effective medium temperature may have been lower than the value

of 275 K assumed for the Sitka climate. Other experimenters have found $T_m = 220$ K to be a more suitable value at 11.6 GHz in northern climates [17]. The curve plotted for $T_m = 220$ K is in better agreement with the CCIR model; however, a disparity still exists between this curve and the COMSAT model. The second factor is the way in which the two models treat path length. Reference 18 discusses the variability in height of the 0° C isotherm and points out that the isotherm tends to be lower at higher latitudes and to display a marked seasonal dependence. The COMSAT model uses effective path lengths based on data that were measured in several locations, but mainly in temperate climates, where the melting layer may be expected to be higher than it is in Sitka. Thus, the model may use an appropriate specific attenuation, but an inappropriately long (for Sitka) effective path length, to arrive at a value of attenuation that would correspondingly be too large. The CCIR model, on the other hand, attempts to correct for melting layer altitude (and thus path length) and results in better agreement with the measured data.

XPD statistics

Figures 13, 14, and 15 show the cumulative distributions of measured 4-GHz XPD for the first half of the experiment (April 20 to October 5, 1980), for the second half (October 5, 1980, to April 22, 1981), and for the overall experiment, respectively. Each figure shows three distributions that correspond to the total XPD observed and the XPD components that were attributable to ice crystal depolarization only and to rain depolarization only. The ice and rain distributions were determined, as discussed earlier, by using concurrently measured 11.6-GHz radiometric excess attenuation. The XPD values measured while the radiometer indicated attenuation below 1.0 dB were attributed to ice crystal depolarization. This is an arbitrary guideline used to illustrate the degree to which depolarization occurred in the absence of significant attenuation, when there was presumably little liquid water along the slant path. Similar methods of identifying ice crystal depolarization have been used by other experimenters [1].

The data in Figures 13 and 14 were measured in spring/summer and autumn/winter, respectively. The relative contributions of ice and rain to the total observed XPD appear to be consistent with the seasonal time frame in each case. The most severe depolarization levels are dominated by rain in the spring/summer data, when the melting layer and thus the rain height would generally be higher, while the autumn/winter data at these levels are dominated by ice. Both data sets are dominated by ice at the less severe depolarization levels, which could reasonably correspond to events in which there was a cloud cover but no appreciable rain along the path. But even at

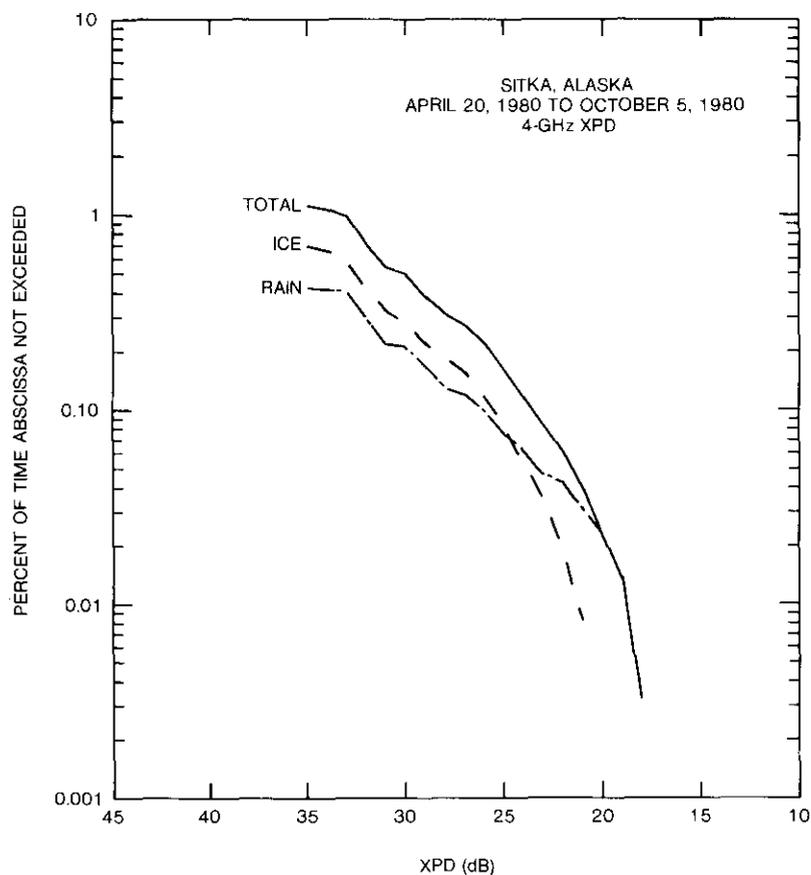


Figure 13. Cumulative Distributions of Ice, Rain, and Total 4-GHz XPD Data for the Period April 20 to October 5, 1980

these levels, the rain component of the distribution is larger in the spring/summer data.

The decibel offset between the rain and total distributions is due to ice effects. Earlier proposals [19] had suggested that, depending on geographic location, modeled total XPD (*i.e.*, including the effects of ice) could be estimated from modeled rain XPD by subtracting a constant factor of 2 to 5 dB from the rain XPD. The October–April data in Figure 14 are consistent with this theory. However, later proposals [16] have suggested that the correction factor should be a variable. The overall data shown in Figure 15

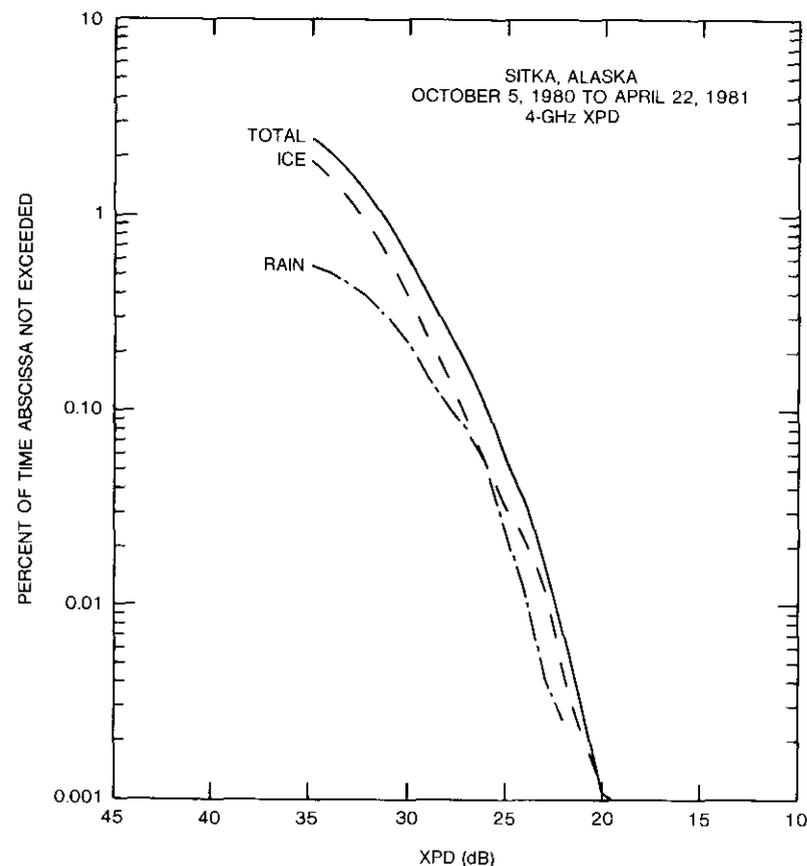


Figure 14. Cumulative Distributions of Ice, Rain, and Total 4-GHz XPD Data for the Period October 5, 1980, to April 22, 1981

indicate the latter to be a better approach, at least for the Sitka experiment. The separation visible in Figure 15 between the rain and total XPD distributions is 2 to 3 dB at higher time percentages; but below about the 0.01-percent level, the separation becomes negligible as rain effects are observed to dominate those of ice.

It has also been noted that in North America, the effect of ice depolarization in relation to rain depolarization appears to be greater in maritime regions than in interior regions [19], [1]. The significant ice component noted in the Sitka data is consistent with this observation.

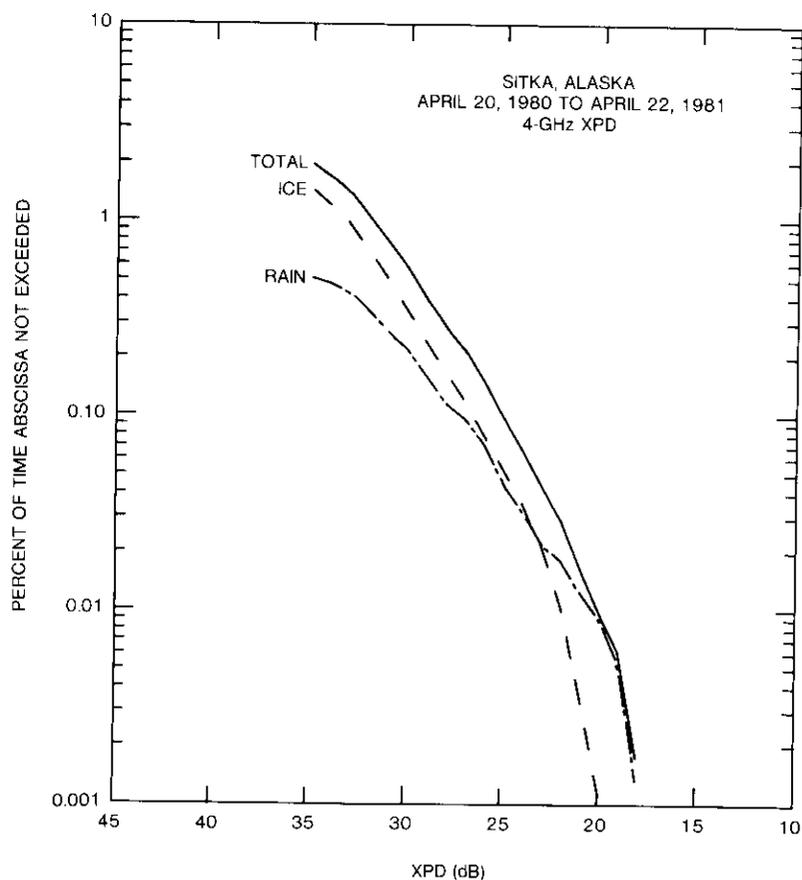


Figure 15. Cumulative Distributions of Ice, Rain, and Total 4-GHz XPD Data for the Overall Experiment (April 20, 1980, to April 22, 1981)

The most severe ice depolarization observed in the Sitka data resulted in an XPD in the 17- to 18-dB range at 4 GHz. The most severe rain depolarization observed resulted in an XPD that was also in the 17- to 18-dB range.

Figure 16 compares the total XPD statistics observed in the overall Sitka data base to cumulative distributions generated by applying modeling expressions given in Reference 19 to the measured radiometric excess attenuation data. The XPD at 11.6 GHz was estimated from the radiometric attenuation data, and the resulting value was frequency-scaled to obtain the estimated XPD at 3.956 GHz. The radiometer data, measured in horizontal polarization, were first corrected to the equivalent circularly polarized attenuation values

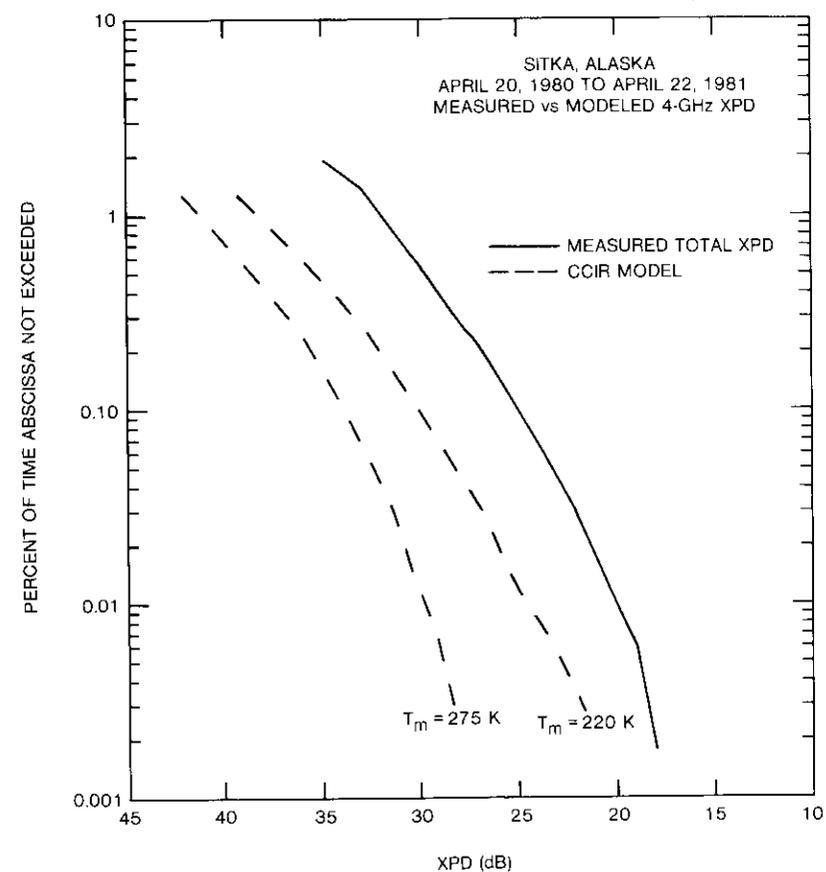


Figure 16. Comparison of Measured and Modeled XPD Statistics

by using the ratio of specific attenuation for the two polarizations, as derived from equations in Reference 20. The correction was small, the measured attenuations being reduced by 5.5 percent.

This approach was chosen because 4- and 11.6-GHz data were available over the same slant path. The alternate approach of employing relationships between XPD, point rain rate, and effective path length was considered to be highly theoretical and would not have been as desirable in this case because of the availability of the path-to-path data.

Even so, as can be seen in Figure 16, the modeling attempts significantly underestimated the observed depolarization. Two model curves are shown, one generated from radiometric data by assuming that $T_m = 275$ K, and the

other assuming $T_m = 220$ K. The latter offers some improvement but still falls short of the measured depolarization data. The underestimate is thought to be due, at least in part, to ice depolarization effects, which were not included in the modeling process.

Comparison to other experiments

The Sitka experiment was one of several INTELSAT programs in which long-term depolarization statistics were gathered at 4 GHz. The XPD statistics from three earlier experiments [21], conducted in Taipei, Taiwan; Ibaraki, Japan; and Lario, Italy, are shown in Figure 17, along with the Sitka results. Table 5 summarizes characteristics of the four experiments. The sites are listed in order of the relative severity of their XPD statistics. The M and β values cited are those of the "best fit" Rice-Holmberg rain rate distribution observed during the given experiment. (It should be noted that, for the Taipei and Ibaraki data, the Rice-Holmberg model did not give a good fit to the measured rain rate distributions).

The results of the four experiments match expected basic trends, with XPD degradations becoming more severe at higher values of M and β and/or at lower elevation angles. Various combinations of the three factors can give intermediate results. Lario, with the lowest M , a very low β , and a high elevation angle, might be expected to have the mildest XPD statistics. Ibaraki, with a higher M and β than Lario, had more severe XPD statistics despite the higher elevation angle. Sitka, with a higher M than Lario or Ibaraki and the lowest elevation angle of the four, had XPD statistics that were more severe, despite a low β . It is felt that the ice depolarization observed in Sitka contributed significantly to the severity of the XPD statistics, although, at lower time percentages, rain was still the dominant factor. While this may not be apparent from the M and β values for Sitka, it might be inferred from the maritime climate. Taipei, with a subtropical (wet) climate, the highest M and β , and a fairly low elevation angle, had the worst XPD statistics, as might be expected.

Summary and conclusions

Results have been presented for 4-GHz depolarization measurements conducted in Sitka, Alaska, over a 12-month period beginning in 1980. The measurements were performed in the northern Pacific Ocean Region at a low elevation angle (11.7°) in a maritime climate that had appreciable rainfall and snowfall.

Comparison of observed vs long-term average rainfall characteristics indicated that the rain accumulation during the experiment was typical for

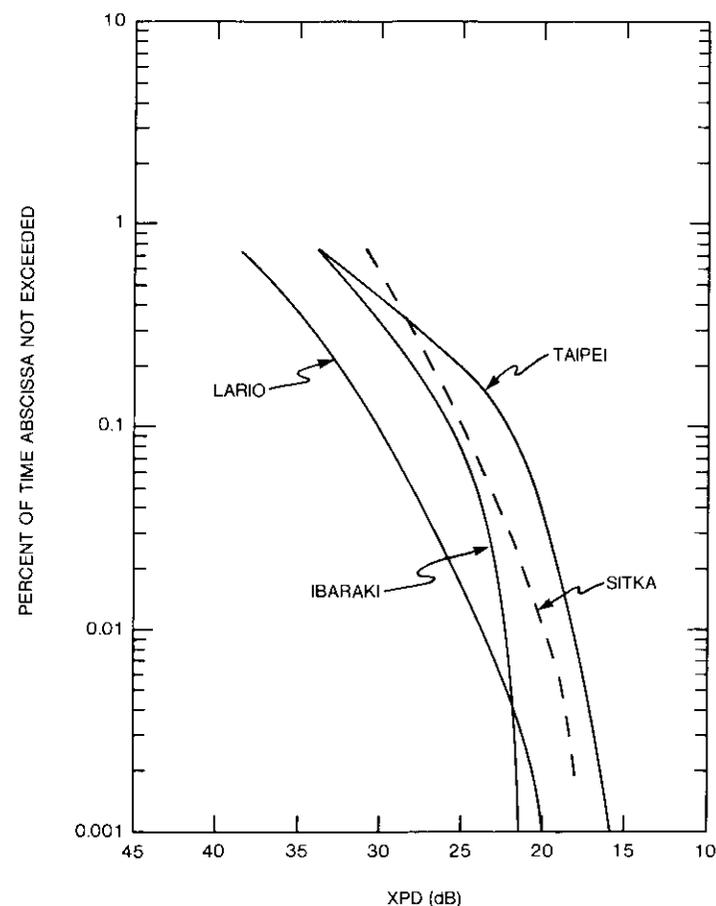


Figure 17. Comparison of Taipei, Ibaraki, Lario, and Sitka 4-GHz XPD Data

the site, but that a lower than expected proportion of rain occurred at high rain rates.

In the data analysis, rain depolarization vs ice crystal depolarization effects were examined. Relative contributions of rain depolarization and ice crystal depolarization to the measured XPD data were estimated by using concurrently measured 11.6-GHz radiometric attenuation data along the same slant path. It was found that, while ice depolarization had made a substantial contribution to the data, rain effects dominated the overall XPD statistics at time percentages

TABLE 5. CHARACTERISTICS OF LARIO, IBARAKI, SITKA, AND TAIPEI MEASUREMENTS

SITE	CLIMATE CLASSIFICATION	M (mm)	β	ELEVATION ANGLE (deg)	0.01-PERCENT XPD LEVEL (dB)
Lario	Temperate Continental (Mountain-influenced)	1,412	0.065	24.7	23.8
Ibaraki	Temperate Continental	1,680	0.2	35	22.4
Sitka	Temperate Maritime	2,409	0.01	11.7	20.1
Taipei	Subtropical (wet)	3,620	0.3	20	18.3

below about 0.01 percent. Measured XPD attributed to ice depolarization was below 21.9 dB for 0.01 percent of the time; the lowest observed value was in the range of 17 to 18 dB. The measured XPD attributed to rain depolarization was below 20.3 dB for 0.01 percent of the time; the lowest observed value was also in the 17- to 18-dB range. For the overall data, the measured XPD was below 20.1 dB for 0.01 percent of the time. Observed ice depolarization effects were consistent with seasonal variations in weather patterns, as well as with estimates proposed in the literature of the relative importance of ice vs rain depolarization.

Modeling attempts for both 11.6-GHz attenuation and 4-GHz depolarization did not agree with observed values. The differences were traceable to the assumptions made regarding effective medium temperature and melting layer altitude, both of which are lower in Sitka than in more temperate climates.

The Sitka data were compared with the results of three earlier measurements performed in Taiwan, Japan, and Italy. The relative effects of climate type, rainfall characteristics, and elevation angle on the XPD statistics of the four measurements were briefly explored and found to be in agreement with expected trends.

Finally, two different measurement techniques were used to collect the XPD data: measurement of linearly polarized orthogonal components of the arriving wave, and direct measurement of XPD with the antenna polarization matched to that of the arriving wave. Details of both measurement techniques were presented and discussed.

Acknowledgments

The author wishes to acknowledge the efforts of all those who contributed to the success of these measurements: D. Kennedy, who originally helped devise the experiment; W. Mercer, P. Ackermann, W. Kinsella, and C. Zaks for the design and fabrication of the measurement equipment; the State of Alaska for the loan of the antenna, and the COMSAT Laboratories Antenna Department for the design and construction of the feed network; G. Hyde for his technical support and advice; the staff of the University of Alaska Geophysical Institute, especially R. Hunsucker, R. Merritt, T. Roberts, B. Delana and P. Brooks, and numerous Sitka residents, for the field implementation of the experiment; the Baranof Composite Squadron of the Civil Air Patrol for providing facilities for the equipment; and M. Holmgren and J. Teas of the Sitka Electronics Laboratory, who operated the equipment in the field. Thanks are also due to K. Yeh, O. Piontek, and P. Fritz, who reduced and processed the experiment data; to S. Donnelly, who wrote and implemented the special-purpose software used in the data processing; and to T. Sullivan, who plotted the data.

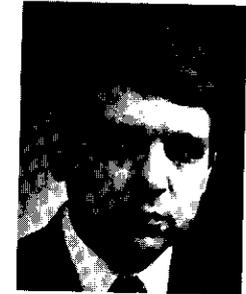
Special thanks are due to D. Rogers, as well as R. Gruner, W. English, R. Kreutel, R. Ott, D. DiFonzo, and I. Dilworth, for many beneficial technical discussions, and D. Fang and J. Allnutt for their encouragement and support throughout the measurement program.

References

- [1] D. C. Cox, "Depolarization of Radio Waves by Atmospheric Hydrometeors in Earth-Space Paths: A Review," *Radio Science*, Vol. 16, No. 5, September-October 1981, pp. 781-812.
- [2] P. K. Lau and J. Allnutt, "Attenuation and Depolarisation Data Obtained on 12 GHz Satellite-Earth Paths at Four Canadian Locations," *Electronics Letters*, Vol. 15, No. 18, August 30, 1979, pp. 565-567.
- [3] J. J. Schlesak and J. I. Strickland, "Geographic and Elevation Angle Dependence of Rain and Ice-Cloud Depolarization in Canada Along Earth-Space Paths at 12 GHz," IEEE International Conference on Communications, Seattle, Washington, June 1980, *ICC '80 Conference Record*, Vol. 3, pp. 40.1.1-40.2.4.
- [4] R. P. Merritt, T. D. Roberts, and R. D. Hunsucker, "Precipitation Depolarization of Satellite Signals in Sitka, Alaska," *The Northern Engineer*, Vol. 13, No. 3, pp. 24-27.
- [5] R. Hunsucker, R. Merritt, and T. Roberts, "Rain Depolarization Measurements From a Geostationary Satellite 4-GHz Down-Link Signal to an Earth Station in Alaska," Technical and Management Proposal, University of Alaska Geophysical Institute, Fairbanks, Alaska, July 1978.

- [6] E. J. Dutton, *Precipitation Variability in the USA for Microwave Terrestrial System Design*, OTR 77-134, Institute of Telecommunications Sciences, Boulder, Colorado, November 1977.
- [7] D. R. DiFonzo and W. S. Trachtman, "Antenna Depolarization Measurements Using Satellite Signals," *COMSAT Technical Review*, Vol. 8, No. 1, Spring 1978, pp. 155-185.
- [8] A. E. Williams, "A Dual-Polarized 4/6-GHz Adaptive Polarization Control Network," *COMSAT Technical Review*, Vol. 7, No. 1, Spring 1977, pp. 247-262.
- [9] D. DiFonzo, W. Trachtman, and A. Williams, "Adaptive Polarization Control for Satellite Frequency Reuse Systems," *COMSAT Technical Review*, Vol. 6, No. 2, Fall 1976, pp. 253-283.
- [10] D. Rogers and G. Hyde, "Diversity Measurements of 11.6-GHz Rain Attenuation at Etam and Lenox, West Virginia," *COMSAT Technical Review*, Vol. 9, No. 1, Spring 1979, pp. 243-254.
- [11] John D. Kraus, *Antennas*, New York: McGraw-Hill Book Company, 1950.
- [12] Max Born and Emil Wolf, *Principles of Optics*, Fifth Edition, New York: Pergamon Press, 1975.
- [13] P. L. Rice and N. R. Holmberg, "Cumulative Time Statistics of Surface-Point Rainfall Rates," *IEEE Transactions on Communications*, COM-21, No. 10, October 1973, pp. 1131-1136.
- [14] International Radio Consultative Committee (CCIR), "Radiometeorological Data," Report 563-1 (MOD F), *Documents of the XVth Plenary Assembly*, Vol. V, Geneva: International Telecommunication Union, 1982.
- [15] D. V. Rogers, "Propagation Analysis Package (PAP-2) User's Guide," COMSAT Technical Memorandum CL-9-80, January 1980.
- [16] International Radio Consultative Committee (CCIR), "Propagation Data Required for Space Telecommunication Systems," Report 564-2, CCIR Interim Working Party 5/2, Geneva, February-March 1983.
- [17] L. Hansson, "Correlating OTS Beacon Measurements With Radiometer and Rainfall Data," *Results of Tests and Experiments With the European OTS Satellite*, IEE Conference Publication No. 199, April 1981, pp. 33-36.
- [18] R. K. Crane, "Prediction of Attenuation by Rain," *IEEE Transactions on Communications*, COM-28, No. 9, September 1980, pp. 1717-1733.
- [19] International Radio Consultative Committee (CCIR), "Propagation Data Required for Space Telecommunication Systems," Report 564-1 (MOD F), *Documents of the XVth Plenary Assembly*, Vol. V, Geneva: International Telecommunication Union, 1982.
- [20] International Radio Consultative Committee (CCIR), "Attenuation by Precipitation and Other Atmospheric Particles," Report 721 (MOD F), *Documents of the XVth Plenary Assembly*, Vol. V, Geneva: International Telecommunication Union, 1982.
- [21] D. J. Kennedy, "Rain Depolarization Measurements at 4 GHz," *COMSAT Technical Review*, Vol. 9, No. 2B, Fall 1979, pp. 629-668.

Steven J. Struharik received B.E.E. and M.S. degrees in electrical engineering from the Ohio State University in 1971 and 1972. He joined COMSAT Laboratories in 1977, and for 6 years was a Member of the Technical Staff of the Propagation Studies Department, where his principal activities involved the design and implementation of propagation experiments. He is currently a Member of the Technical Staff of the Systems Simulation Department. Mr. Struharik is a member of Tau Beta Pi, Sigma Xi, and Eta Kappa Nu.



Ion implantation of Boron in GaAs devices

P. J. MCNALLY

(Manuscript received May 16, 1983)

Abstract

This paper describes some unique properties of ion implanted boron in GaAs which have been applied to improve the performance of Schottky barrier diodes and ion implanted low-noise MESFETs. Self-aligned high-resistivity guard rings for Schottky barrier diodes were formed by using low energy (40 keV) and low dose (10^{12} cm⁻²) implanted boron (¹¹B⁻). Reverse breakdown voltages up to 100 V were achieved with the ¹¹B⁺ implanted guard ring, which was used to improve the performance of a 13- to 18-GHz switching device,

GaAs substrate quality plays a major role in MESFET performance for devices fabricated by direct ion implantation of silicon atoms into liquid-encapsulated Czochralski material. High-energy (270 keV) ¹¹B⁻ implantation was used to improve the silicon impurity distribution near the substrate interface, reduce substrate effects, and produce a more abrupt interface. Since current conduction under low-noise MESFET operation occurs in this region of the channel, microwave performance is improved significantly.

Introduction

Ion implantation in GaAs has become an established doping technique in the fabrication of discrete devices and integrated circuits [1], [2]. Extensive characterization of both *n*- and *p*-type dopant elements has been done in the development of this technology [3]. Ion implantation of nondopant elements

has also contributed to GaAs device technology by providing isolation processes that use protons and oxygen for high-resistivity layer formation [4], [5]. High-resistivity regions are formed by free carrier compensation that results from radiation damage and/or introduction of stable deep levels in the forbidden band. Ion implantation damage resulting from neon ions has been used to getter impurities from semi-insulating GaAs prior to epitaxial layer growth [6]. Studies of implant damage compensation using nondopant elements of nitrogen, fluorine, and boron, and subsequent annealing characteristics, have been valuable in understanding defect production and annealing kinetics for heavy ions in GaAs [7]. In this paper, boron implantation has been used to improve the reverse breakdown voltage characteristics of Schottky barrier diodes and the microwave performance of GaAs MESFETS.

Implanted Schottky barrier guard ring

Planar Schottky barrier diodes without integral guard rings exhibit excess reverse bias leakage current and low breakdown voltage because the radius of curvature is zero at the edges of the metal semiconductor junction. Electric field crowding occurs in this region of the device, resulting in significant lowering of the barrier potential, which permits increased current flow at voltages below the breakdown voltage. Increased reverse leakage current adds to the power dissipation of a Schottky device, which places operating limits on its use in circuit applications. These effects became evident when attempts were made to produce high breakdown voltage Schottky barrier diodes on epitaxial GaAs with carrier concentrations in the range of 1×10^{15} to $1 \times 10^{16} \text{ cm}^{-3}$ and layer thickness of 3 to 5 μm .

The application and structure used in this work was the switchable attenuating medium propagation (SAMP) device, an RF switch for the 13- to 18-GHz region [8]. The Cr/Au Schottky diode in this device has a coplanar $5 \times 1,250 \mu\text{m}$ area and two AuGeNi ohmic contacts. Spacing in the gaps between the Schottky diode and the ohmic contacts is 7 μm . Device operation principles and geometry precluded the use of a conventional $p-n$ junction guard ring to increase breakdown voltage characteristics. Coplanar requirements also prevented the use of a mesa structure.

Experimental approach

High breakdown voltage devices were achieved by incorporating an ion implanted boron ($^{11}\text{B}^+$) guard ring in the gaps between the ohmic contacts and the Schottky metallization. A self-aligned structure was produced, and a cross section is shown in Figure 1.

Boron was selectively implanted, as shown, by using the Schottky barrier

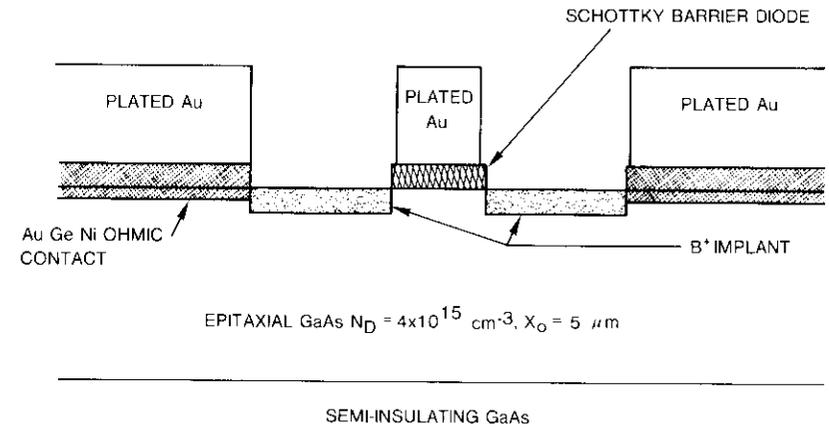


Figure 1. Cross Section of Self-Aligned Boron Implanted Guard Ring in the SAMP Device

and ohmic contact metallizations to mask against the ion beam. The thickness of the metallizations was chosen to be greater than 3,000 \AA , which is adequate to prevent penetration of the boron ions into the GaAs. The boron implanted regions are self-aligned to both metallizations by this procedure. Low-energy $^{11}\text{B}^+$ was implanted at 40 keV, creating a high-resistivity region that extended from the Schottky metal periphery to the AuGeNi contacts. The net effect was to increase the radius of curvature at the Schottky metal edge from zero to infinity and to achieve plane breakdown characteristics in a planar structure. The implantation conditions selected were such that the depth of the boron implant was the minimum necessary to achieve the desired operating voltage. The operating voltage required full depletion of the epitaxial layer. It was found that 40 keV $^{11}\text{B}^+$ was adequate for the carrier concentration range of 1×10^{15} to $1 \times 10^{16} \text{ cm}^{-3}$ and layer thickness of 3 to 5 μm . To determine boron depth, a plot of range vs implant energy based on the LSS theory was generated and is shown in Figure 2. The curve R_p is the projected range corresponding to the peak of the boron distribution, and σ_p corresponds to the standard deviation in projected range for the assumed Gaussian distribution. An implant energy of 40 keV corresponds to a depth of 0.25 μm with $R_p + 3\sigma_p$ as the total penetration for the implanted $^{11}\text{B}^+$.

Experimental results

The relation that describes breakdown voltage for a plane junction where $r_j = \infty$ is given by

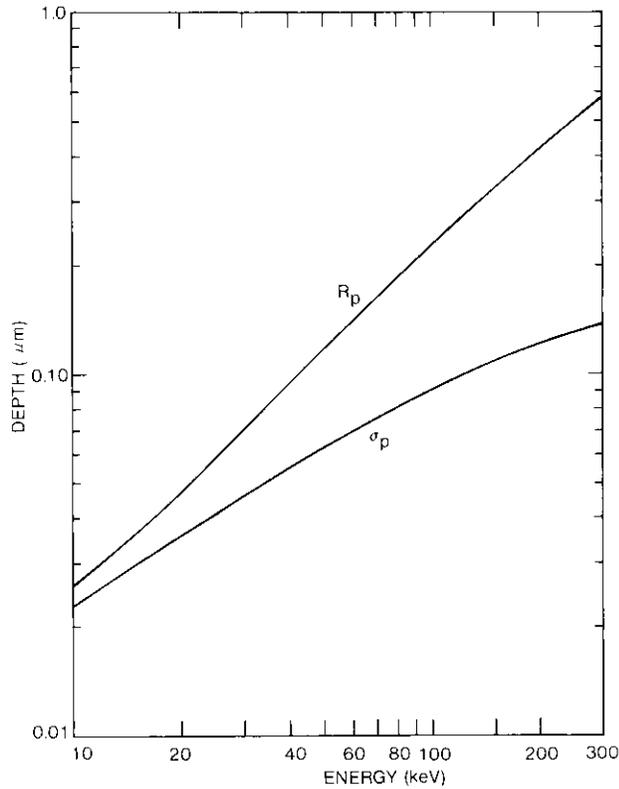


Figure 2. Projected Range (R_p) and Standard Deviation (σ_p) vs Implant Energy for $^{11}\text{B}^+$ in GaAs (assuming Gaussian distribution of LSS theory)

$$V_{BR} = \frac{\epsilon_s \epsilon_o E_{crit}^2}{2q N_D}$$

where

- E_{crit} = critical electric field
- N_D = impurity concentration
- $\epsilon_s \epsilon_o$ = permittivity
- q = electronic charge.

and is plotted in Figure 3 as a function of N_D . The figure also shows depletion layer width for different voltages as a function of carrier concentration. From this figure, the minimum breakdown voltage for a specific epitaxial layer (carrier concentration and thickness) necessary to achieve full depletion can

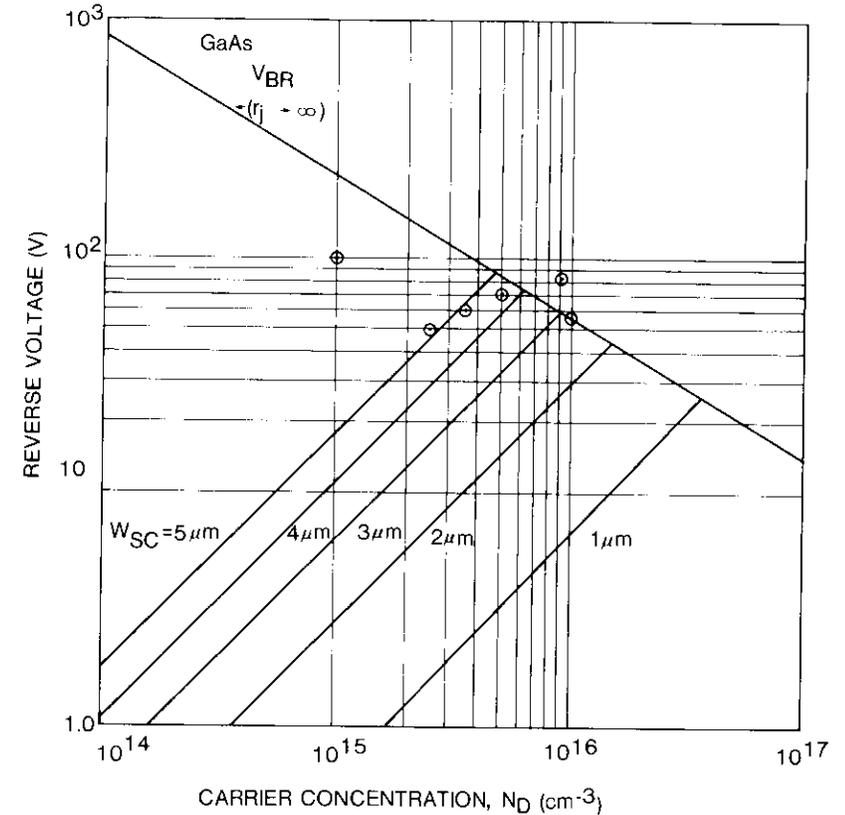


Figure 3. Reverse Breakdown Voltage (V_{BR}) and Depletion Width (W_{sc}) vs Carrier Concentration for GaAs

be ascertained. The maximum voltage value achievable is the V_{BR} curve. The data points in the figure are breakdown voltage values obtained by using the $^{11}\text{B}^+$ implanted guard ring. The starting epitaxial wafers were measured prior to device fabrication by using capacitance-voltage to derive carrier concentration profiles. Possible uncertainties in this measurement can account for the data point corresponding to the $9 \times 10^{15} \text{ cm}^{-3}$ concentration which lies above the V_{BR} curve. However, the data show a clustering near the V_{BR} curve in the mid- 10^{15} to $1 \times 10^{16} \text{ cm}^{-3}$ concentration range.

The effect of an increasing implanted $^{11}\text{B}^+$ dose on breakdown voltage is shown in Table 1, where the data represent a wafer containing 24 devices with $N_D \cong 5 \times 10^{15} \text{ cm}^{-3}$ and epitaxial layer thickness equal to $5 \mu\text{m}$. The initial breakdown voltage measured at $4 \mu\text{A}$ is shown, and improvement is

TABLE 1. SCHOTTKY BARRIER DIODE BREAKDOWN VOLTAGE IMPROVEMENT WITH $^{11}\text{B}^+$ IMPLANT (K-21), $N_D = 5 \times 10^{15}$, $X_{epi} = 5 \mu\text{m}$

DEVICE No.	INITIAL V_B	$8 \times 10^{10} \text{ B}^+/\text{cm}^2$	$2.4 \times 10^{11} \text{ B}^+/\text{cm}^2$	$5.6 \times 10^{11} \text{ B}^+/\text{cm}^2$
	AT $4\mu\text{A}$ (V)	AT $4\mu\text{A}$ (V)	AT $4\mu\text{A}$ (V)	AT $4\mu\text{A}$ (V)
1	36	38	48	50
2	40	50	50	65
3	45	50	50	65
4	15	38	55	65
5	32	45	48	60
6	48	48	60	70
7	4	8	14	14
8	47	48	58	68
9	46	53	60	70
10	4	18	20	22
11	46	52	52	68
12	46	55	60	68
13	34	35	40	40
14	48	44	52	70
15	47	53	60	66
16	48	48	60	70
17	24	25	Short	Short
18	45	50	60	70
19	42	45	48	70
20	36	40	40	40
21	45	53	Probe Damaged	Probe Damaged
22	22	20	20	30
23	45	50	62	70
24	45	50	50	55

observed as the $^{11}\text{B}^+$ dose is increased. The voltage increase for this wafer reached a maximum at 70 V for a dose of $5.6 \times 10^{11} \text{ cm}^{-2}$. In several devices, the breakdown voltage remained at low values, which are indicative of premature breakdown in the central area of the Schottky diode. The source of premature breakdown does not become evident until the $^{11}\text{B}^+$ guard ring fabrication is completed at the $5.6 \times 10^{11} \text{ cm}^{-2}$ dose level. Since contributions to the leakage current from the Schottky edge are suppressed, diodes with V_{BR} less than approximately 70 V have breakdown characteristics controlled by metal/substrate imperfections such as material defects and interfacial contamination. The guard ring process can be used as a technique for Schottky barrier development in GaAs where it is desirable to isolate fabrication variables related to metal deposition and wafer preparation.

To accommodate variations in material quality and Schottky barrier uniformity, the final $^{11}\text{B}^+$ implant conditions were 40 keV and $1 \times 10^{12} \text{ cm}^{-2}$

for all wafers fabricated with this process. The combination of low energy and low dose permitted the implantation step to be done at the end of wafer processing. Self-alignment is inherent in this process, since both metallizations act as implant masks to the ion beam. Elevated temperature annealing at 500°C was performed on several device wafers that used Ta-Au Schottky barrier metallization. The boron implants showed no change in reverse breakdown characteristics at this temperature. The implanted regions remain at high resistivity, which suggests that the boron interacts chemically with the GaAs, since significant implant damage should anneal at 500°C . This interaction is further discussed in the next section.

Boron implantation in GaAs MESFET fabrication

Direction ion implantation of semi-insulating GaAs liquid-encapsulated Czochralski (LEC) material for MESFET n -layer formation is strongly influenced by the quality of the substrate. It has been shown that substrate stoichiometry considerations are of prime importance in obtaining well-behaved implant layers [9]–[11]. Ingot selection procedures based on electrical measurements before and after heat treatment of sample wafers or device fabrication of sample wafers are in general use [12]. Experimental work has been done using pre-damage by heavy ions (Ar^+) to improve the implanted layer transport properties and reduce substrate effects [13]. In the present work, high-energy boron implantation was used to alter the n -type implanted profile in the region of the semi-insulating substrate interface. This resulted in improved low-noise MESFET performance in both noise figure and associated gain.

Experimental procedures

Adjacent LEC GaAs wafers from a single ingot with resistivity $\geq 10^7 \Omega\text{-cm}$ and mobility of $3,700 \text{ cm}^2\text{v}^{-1}\text{s}^{-1}$ were used. Post-implantation capping and annealing procedures consisted of plasma-enhanced CVD deposition of 900 Å-thick Si_3N_4 at 400°C and furnace annealing in forming gas (10 percent H_2) at 850°C for 20 minutes. The photolithographic mask set consisted of characterization test patterns and several low-noise FET designs. Conventional post-ion implant processing was used to complete the wafer fabrication, including AuGe/Ni ohmic contacts, gate recessing and Cr/Au Schottky barrier metallization.

Two LEC GaAs wafers were ion implanted with ^{29}Si and ^{11}B , with the ^{11}B buried below the ^{29}Si distribution. The ^{11}B distribution was assumed to be Gaussian and the range/energy curves of Figure 2 were used to select implant energy and dose. Since the ^{29}Si implant was performed at 200 keV, ^{29}Si

range/energy data [10] were also used to select the ^{11}B implant parameters so that the peak concentration would occur below the ^{29}Si peak and would have its major effect on the tail of the ^{29}Si distribution. It is this region of the ^{29}Si profile that is undepleted and provides channel conduction under low-noise FET bias conditions. The transport properties in this portion of the channel have a major effect on FET performance and can be strongly influenced by the substrate quality. Wafer CA116 was ^{11}B implanted at 270 keV, $7 \times 10^{12} \text{ cm}^{-2}$ and then implanted with ^{29}Si at 200 keV, $4.225 \times 10^{12} \text{ cm}^{-2}$. This wafer was capped and annealed as described above. Wafer CA113 was first implanted with ^{29}Si at 200 keV, $4.225 \times 10^{12} \text{ cm}^{-2}$ and nitride capped at 400°C . The ^{11}B was then implanted at 270 keV, $7 \times 10^{12} \text{ cm}^{-2}$ through the Si_3N_4 film and annealed.

Experimental results

Figure 4 shows the concentration profiles for the implants performed. The curves B1 and B2 are LSS theory profiles for 270 keV $^{11}\text{B}^+$. The B2 profile is displaced toward the surface to account for the 40 keV energy loss in the Si_3N_4 cap. Profiles B1 and B2 correspond to the boron implant conditions in CA116 and CA113, respectively. The dashed curve is the LSS theory profile for 200 keV, $4.225 \times 10^{12} \text{ cm}^{-2}$ ^{29}Si implant conditions, which is the same for both wafers. The profiles labeled CA116, CA113, and CA110 are measured carrier distributions that use the capacitance-voltage technique. Wafer CA110 was not implanted with ^{11}B and is shown in the figure for comparison. Wafer CA110 had ^{29}Si implant conditions of 200 keV, $3.74 \times 10^{12} \text{ cm}^{-2}$ and 75 keV, $1 \times 10^{12} \text{ cm}^{-2}$. This wafer was processed prior to wafers CA116 and CA113, at which time a change in implant parameters to a single implant energy became standard. Since all wafers were implanted with ^{29}Si at 200 keV and the ^{11}B implant effect occurs at the substrate interface, the observed results are not affected by the double implant of CA110.

The carrier profile of CA116 shows good agreement with the theoretical ^{29}Si distribution below $\sim 7 \times 10^{16} \text{ cm}^{-3}$, where the boron-to-silicon ratio is greater than 0.3. Deviation from theory increases as this ratio becomes smaller. The boron appears to enhance the silicon activation by interacting with arsenic vacancies in a manner which reduces the amphoteric doping nature of silicon in GaAs. As the boron concentration falls off toward the surface, the effect on the silicon profile is reduced. Since the peak boron concentration lies below the implanted silicon, the boron could also be acting as a buried diffusion barrier to out-diffusing arsenic vacancies from the substrate, thus enhancing electrical activation in the profile tail. The maximum effect from a diffusion barrier should occur at the interface, since the boron

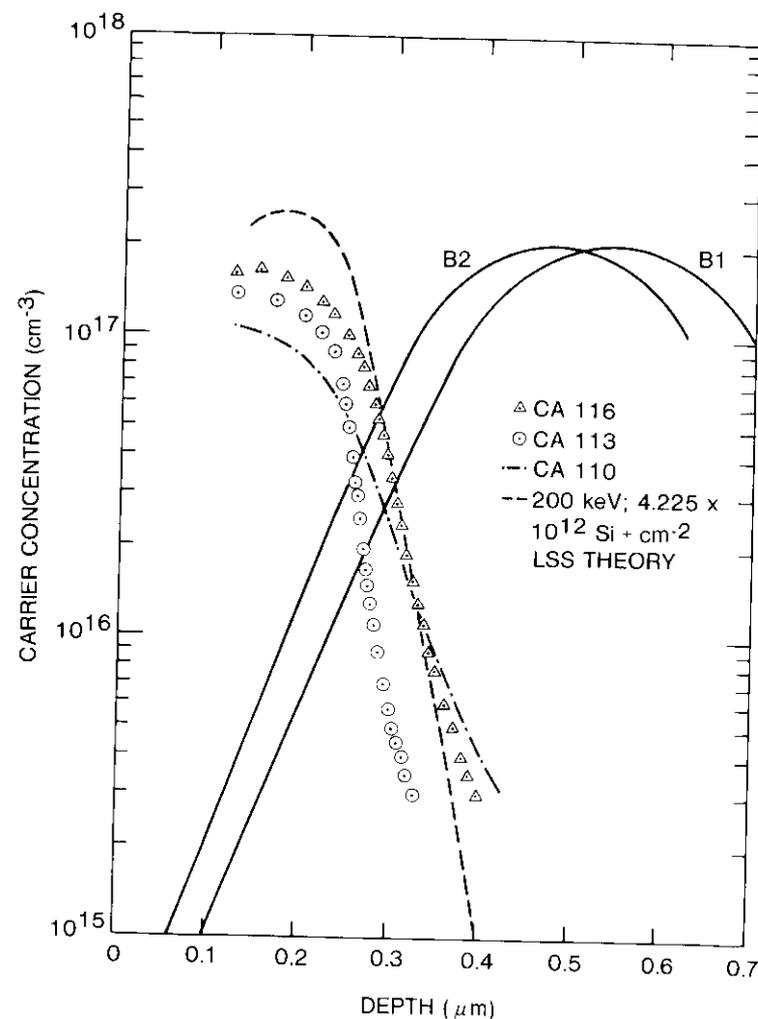


Figure 4. Ion Implanted ^{29}Si and ^{11}B Distributions and Electron Density Profiles in LEC GaAs

concentration increases with depth into the substrate. This is in agreement with the discussion in Reference 11 concerning stoichiometric defect diffusion in LEC material. This reference notes that, in material grown in an arsenic-rich condition, low arsenic vacancy concentrations contribute to good thermal stability and high-quality ion implanted layers.

The profile of CA113 shows a different behavior of the implanted boron. In this case, the silicon implanted region underwent a partial anneal during the Si₃N₄ deposition at 400°C (~20 minutes). The measured profile is reduced in thickness by approximately the difference in penetration depth of the two boron distributions. The peak carrier concentration is also reduced. Boron, in this case, affects the impurity profile in a different way than in wafer CA116. From the observed hysteresis in capacitance-voltage behavior for $V_R \geq 3$ V, corresponding to $n \leq 7 \times 10^{16} \text{ cm}^{-3}$, boron implanted into a partially annealed implanted layer produces traps which reduce free carriers in the tail of the profile. The boron-induced traps are stable after subsequent high-temperature anneal (850°C). Figure 5 shows capacitance-voltage curves and the hysteresis for CA113. The slopes of the two carrier profiles in Figure 4 (CA113 and CA116) are nearly the same and show an improvement in abruptness and peak electron concentration when compared to CA110 without boron. The total silicon concentration in CA110 was slightly higher than that used for the boron implanted wafers.

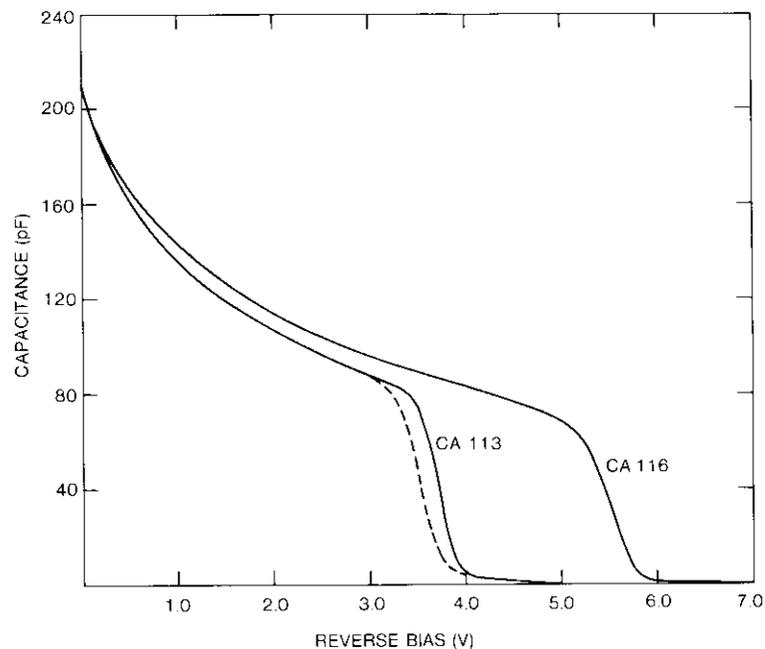


Figure 5. Capacitance-Voltage Characteristics for the ¹¹B Implanted Wafers CA113 and CA116

GaAs MESFET performance

GaAs MESFETs with nominal $1 \times 500 \mu\text{m}$ gates were evaluated for microwave performance of noise figure and associated gain at 4 and 12 GHz and I-V characteristics. Figure 6 shows typical I-V characteristics for devices from the above wafers. The transconductance of CA113 and CA116 is

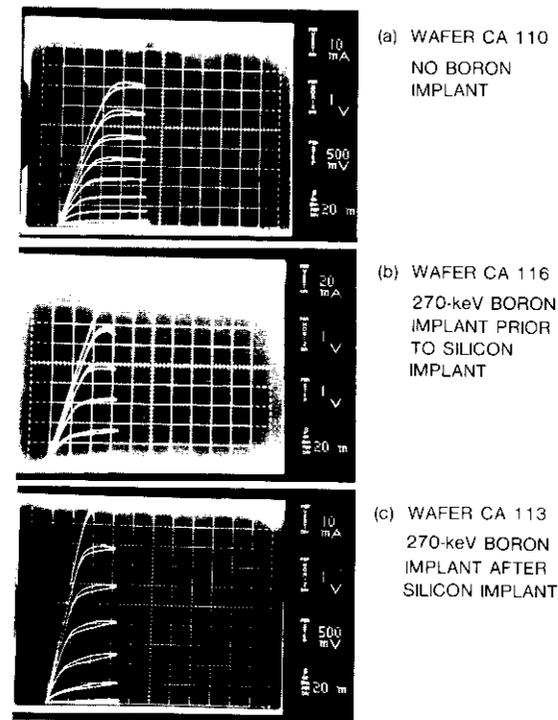


Figure 6. Comparison of Current-Voltage Characteristics of GaAs MESFETs Prepared With and Without High-Energy Boron Implantation

42 mS at I_{DSS} , which decreases to 24 mS at the low-noise bias current of 10 mA. In comparison, the CA110 device shows a transconductance of 28 mS at I_{DSS} , decreasing to 10 mS at the low-noise bias (10 mA). Improvement in transconductance is due to the boron implant effects on the silicon profile. Table 2 contains microwave performance data comparing the different devices. The boron implanted devices exhibit improved noise figure and gain over the CA110 device.

TABLE 2. MICROWAVE PERFORMANCE OF GaAs MESFETS WITH NOMINAL $1 \times 500 \mu\text{m}$ GATES

	WAFER NUMBER		
	CA110	CA113	CA116
Noise Figure at 4 GHz (dB)	2.23	1.55	1.68
Associated Gain at 4 GHz (dB)	9.7	11.1	11.3
Noise Figure at 12 GHz (dB)	No Measurement	2.9	2.85
Associated Gain at 12 GHz (dB)	No Measurement	6.0	7.6

Summary

Ion implantation of boron, a nondopant in GaAs, was used to improve Schottky barrier diodes and MESFETS. High reverse breakdown Schottky barrier guard rings were fabricated in a self-aligned structure which exhibited plane breakdown voltage characteristics. The guard rings are stable when annealed at 500°C. The high-resistivity regions formed at room temperature are due to implant-damage-induced defects. Annealing at 500°C removes a significant amount of the damage, and the boron interacts with the GaAs to form a complex that maintains the high-resistivity characteristics of the layer.

Boron implanted into LEC GaAs for MESFET fabrication was shown to improve the characteristics of an ion implanted channel. Carrier concentration profiles are given for different implant conditions. When boron is implanted prior to the silicon implant, the boron behaves like a buried diffusion barrier to arsenic vacancies diffusing into the silicon implanted region. When boron is implanted into a partially annealed silicon implanted region, traps are formed in the tail of the silicon profile which are stable after 850°C anneal. In both instances, the silicon profiles are more abrupt than those produced without the boron implant. Both transconductance and microwave performance show significant improvement for boron implanted devices.

References

- [1] M. Feng et al., "Low Noise GaAs Metal Semiconductor Field Effect Transistor Made by Ion Implantation," *Applied Physics Letters*, Vol. 40, No. 9, May 1, 1982, pp. 802-804.

- [2] B. M. Welch et al., "LSI Processing Technology for Planar GaAs Integrated Circuits," *IEEE Transactions on Electron Devices*, ED-27, No. 6, June 1980, pp. 1116-1124.
- [3] J. P. Donnelly, "The Electrical Characteristics of Ion Implanted Compound Semiconductors," *Proceedings of the 2nd International Conference on Ion Beam Modification of Materials*, R. E. Benenson, ed., New York: North-Holland, 1981, p. 553.
- [4] A. G. Foyt et al., "Isolation of Junction Devices in GaAs Using Proton Bombardment," *Solid State Electronics*, Vol. 12, 1969, pp. 209-214.
- [5] T. Asono et al., "Formation of Thick, Thermally Stable High Resistivity Layers in GaAs by Oxygen Ion Implantation," *Japanese Journal of Applied Physics*, Vol. 20, No. 5, May 1981, pp. 901-907.
- [6] C. O. Bozler et al., "Impurity Gettering in Semi-Insulating Gallium Arsenide Using Ion Implantation Damage," *Applied Physics Letters*, Vol. 29, No. 11, December 1, 1976, pp. 698-699.
- [7] D. E. Davies et al., "Compensation From Implantation in GaAs," *Applied Physics Letters*, Vol. 23, No. 11, December 11, 1973, pp. 615-616.
- [8] P. L. Fleming et al., "GaAs SAMP Device for K_v-Band Switching," *IEEE Transactions on Microwave Theory and Techniques*, MTT-27, December 1979, pp. 1032-1035.
- [9] R. N. Thomas et al., "Growth and Characterization of Large Diameter Undoped Semi-Insulating GaAs for Direct Ion Implanted FET Technology," *Solid State Electronics*, Vol. 24, 1981, pp. 387-399.
- [10] T. Onurra et al., "Range and Standard Deviation of Ion Implanted Si in GaAs," *Journal of Applied Physics*, Vol. 52, No. 10, October 1981, pp. 6128-6132.
- [11] L. B. Ta et al., "Effects of Stoichiometry on Thermal Stability of Undoped, Semi-Insulating GaAs," *Journal of Applied Physics*, Vol. 53, No. 8, August 1982, pp. 5771-5775.
- [12] J. A. Higgins et al., "Low Noise GaAs FET's Prepared by Ion Implantation," *IEEE Transactions on Electron Devices*, ED-25, No. 6, June 1978, pp. 587-596.
- [13] S. G. Liu et al., "²⁸Si Implantation Into ⁴⁰Ar Implant-Pretreated Semi-Insulating GaAs Substrates—Mobility and Activation Efficiency Enhancement," *Applied Physics Letters*, Vol. 41, No. 1, July 1982, pp. 72-75.



Philip J. McNally received a B.S. in Physics from St. Francis College in 1959 and pursued graduate studies at Pennsylvania State University and Northeastern University. Prior to joining COMSAT Laboratories in 1974, he was with Ion Physics Corporation and Honeywell, Inc., where he was an early contributor in the field of ion implantation and its application to semiconductor device technology. These activities included both silicon and compound semiconductor devices, the latter emphasizing photodetector development. He is currently a Research Scientist in the Microelectronics Division at

COMSAT, where he has been involved in research on IMPATT diodes, ion implanted GaAs FETs, and Schottky barriers. He is currently responsible for ion implantation technology development for GaAs integrated circuits. He is a member of the IEEE, Sigma Pi Sigma, and Delta Epsilon Sigma.

Translations of Abstracts

Modulation par déplacement de phase à 8 états avec codage dans les télécommunications en AMRT par satellite

S. A. RHODES, R. J. FANG ET P. Y. CHANG

Sommaire

On a étudié l'utilisation de la modulation par déplacement de phase à 8 états (MDP-8 avec codage) à titre de technique de signalisation applicable aux télécommunications en mode d'accès multiple par répartition dans le temps (AMRT). Le fonctionnement de la MDP-8 avec codage a été examiné en vue de son utilisation avec divers codes de correction d'erreurs sans voie de retour (FEC) ayant des niveaux de complexité différents. On a utilisé des simulations sur ordinateur pour évaluer la probabilité d'erreurs sur les bits, P_b , en fonction du rapport E_b/N_0 pour décoder les symboles MDP-8 détectés par décisions quantifiées, pour divers niveaux de quantification. Grâce au décodeur de trajet à maximum de vraisemblance (Viterbi), la technique MDP-8 avec codage assure un gain de codage par rapport à la modulation par déplacement de phase à 4 états (MDP-4) sans codage. De même, l'utilisation d'un taux de codage de 2/3 permet à chaque symbole MDP-8 d'acheminer deux bits d'information, et donc de fournir le même rendement la largeur de bande que la MDP-4 sans codage.

D'importants gains de codage par rapport à la MDP-4 sans codage ont été obtenus grâce à la technique MDP-8 avec codage tant dans un canal auquel on a ajouté du bruit blanc Gaussien, que dans un canal INTELSAT V simulé. Les simulations sur ordinateur ont révélé que pour un code de correction d'erreurs FEC à 64 états, la MDP-8 avec codage transmise dans un canal avec bruit blanc Gaussien ajouté produit un gain de codage de 3 dB environ lorsque $P_b = 10^{-4}$. Dans le cas où $P_b = 10^{-6}$, le gain de codage estimé est de 4 dB environ dans ce même canal. Les gains de codage correspondants pour la MDP-8 avec codage dans le canal INTELSAT V se sont élevés à 3,2 dB pour $P_b = 10^{-4}$, et 4,5 dB pour $P_b = 10^{-6}$. On a déterminé également que la MDP-8 avec codage était quelque peu plus performante que la MDP-4 sans codage pour une distorsion de temps de propagation (composante parabolique) élevée dans le filtre du multiplexeur d'entrée du répéteur simulé du satellite INTELSAT V.