SECTION 8I: TECHNICAL AND OPERATING CHARACTERISTICS OF MOBILE SATELLITE SERVICES

REPORT 509-5

# MODULATION AND CODING TECHNIQUE FOR MOBILE SATELLITE SERVICE

(Question 87/8)

(1970-1974-1978-1982-1986-1990)

## 1. General

This Report describes modulation and coding techniques which will be suitable for mobile satellite communications services. Possible techniques are briefly described and their fundamental performance characteristics are surveyed.

## 2. Modulation technique

### 2.1 *Analogue voice modulation*

#### 2.1.1 *General*

Among analogue voice modulation techniques, frequency modulation (FM) and single-sideband amplitude modulation (SSB) have the potential for providing a speech quality which would be tolerable for the public mobile telephone service at reasonable expenditures of satellite power, frequency band, and mobile-borne equipment complexity and cost.

For systems where the maximum number of channels is essentially limited by the satellite power available, the more power efficient FM system with its relatively wider bandwidth requirements is superior to SSB. Where no such power limits are present, SSB could provide higher quality.

#### 2.1.2 *Comparison of SSB and FM in terms of articulation index (AI)*

The relation between articulation index and signal-to-noise density at the receiver is shown in Fig. 1 for typical SSB and FM modulation techniques. For example, for the condition of a peak power limited satellite and a signal-to-noise density of 46 dBHz, it is noted that single-sideband yields an articulation index of 0.43 and that frequency modulation using bandwidths of 8 kHz and 18 kHz yields articulation indices of 0.53 and 0.57, respectively. This $C/N_0$ is 1 dB above the threshold value of an 8 kHz system and is 3 dB below threshold point of an 18 kHz system. A signal-to-noise density of 49 dBHz yields articulation indices of 0.52 for single-sideband, 0.60 for 8 kHz frequency modulation and 0.68 for 18 kHz frequency-modulation, and is above threshold in both of these frequency-modulation systems.

The results shown are based upon a theoretical analysis of data generated by tests with the English language.

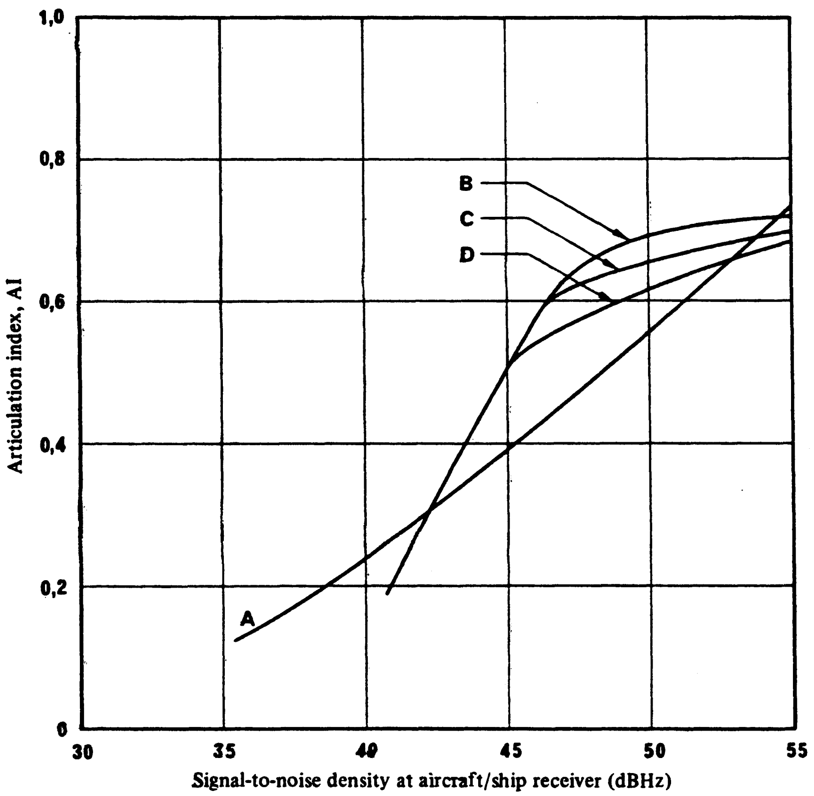


**FIGURE 1** – *Comparison of analogue voice modulation techniques*

A: 5 kHz single sideband-peak radio-frequency power

B: frequency modulation (18 kHz)

C: frequency modulation (10 kHz)

D: frequency modulation (8 kHz)

*Note.* – The articulation index (AI) of a speech circuit indicates the effective proportion of an ideal voice channel (with respect to intelligibility) which that circuit will provide under the given signal and noise conditions [Beranek, 1947]. The audio band is unequally divided into 20 segments, each of which contributes equally to the intelligibility of speech. This contribution is independent of what happens in other segments. To achieve an articulation index of 1.0, a signal-to-noise ratio of 30 dB must be provided in each of the segments, i.e. each contributes 5% of the articulation index when the speech spectrum is not masked by noise and is sufficiently loud to be above the threshold of audibility. Noise density may or may not be constant over the audio baseband so the effect in each segment will vary.

2.1.3 *INMARSAT system*

For the INMARSAT system which is essentially power-limited at present, narrow-band frequency modulation (NBFM) is used for telephony. Table I shows the major parameters of the NBFM used in the INMARSAT system. This NBFM system offers good speech quality with a link $C/N_0$ of 52 dBHz and gives tolerable speech quality with a link $C/N_0$ of 48.6 dBHz when a demodulator employing threshold extension technique, is used.

TABLE I – *Major parameters of NBFM used in the INMARSAT Standard-A system*

| | |
|---|---|
| Baseband | 300-3000 Hz |
| Emphasis | None |
| Companding | 2 : 1 syllabic<br>As specified in CCITT Recommendation G.162 with unaffected level equal to 0 dBm0 |
| Peak clipping level | At peaks of an 800 Hz test tone of 0 dBm0 |
| Peak frequency deviation | 12 kHz at 0 dBm0 baseband level in the absence of peak clipping |
| Mean speech level | – 14 dBm0 |
| Standard deviation | 6 dB |

Figure 2 shows the equivalent speech-to-noise ratio of an NBFM system, with parameters very similar to those shown in Table I, as a function of link $C/N_0$ against talker level (dBr) (r.m.s. power of talker to r.m.s. power for full deviation).

2.1.4 *Other factors*

In addition to voice quality and satellite radiated power requirements, other factors must be considered before selecting the modulation technique. Some of these factors are: spectrum requirements, effects of oscillator instabilities and Doppler shifts, system reliability and costs.

2.2 *Digital modulation*

The modulation methods to be used in the transmission of digital signals in a mobile satellite system should meet the following requirements:

- good bit error ratio (BER) as a function of $E_b/N_0$ performance;
- efficient bandwidth utilization;
- inexpensive and simple implementation.

In addition, it is preferable if the modulation method adopted results in a constant amplitude signal to make it possible to use a class C power amplifier for the mobile transmitter. If amplitude variation exists in the modulated signal waveform, higher power-spectrum side lobes of the signal will be produced in the process of amplification by a non-linear transmitter, i.e. a class C amplifier., The following modulation methods can be considered as candidates for the mobile satellite service application:

- 2-PSK (BPSK);
- unfiltered 4-PSK;
- filtered 4-PSK in which baseband rectangular signals are filtered or shaped before being applied to the 4-PSK modulator;
- offset 4-PSK in which baseband digital signals in one channel are offset by the duration of a half symbol interval relative to those in the orthogonal channel prior to the filtered 4-PSK modulation;
- minimum shift keying (MSK) [de Buda, 1972] often referred to as fast frequency shift keying (FFSK), i.e. FSK with a deviation ratio of 0.5;
- tamed frequency modulation (TFM) [de Jaeger and Dekker, 1978; Muilwijk and Noordanus, 1980] which is a class of modulation methods known as "correlative phase shift keying";
- Gaussian filtered MSK (GMSK) [Hirade and Murota, 1979] which is a modified MSK using a premodulation Gaussian low-pass filter.

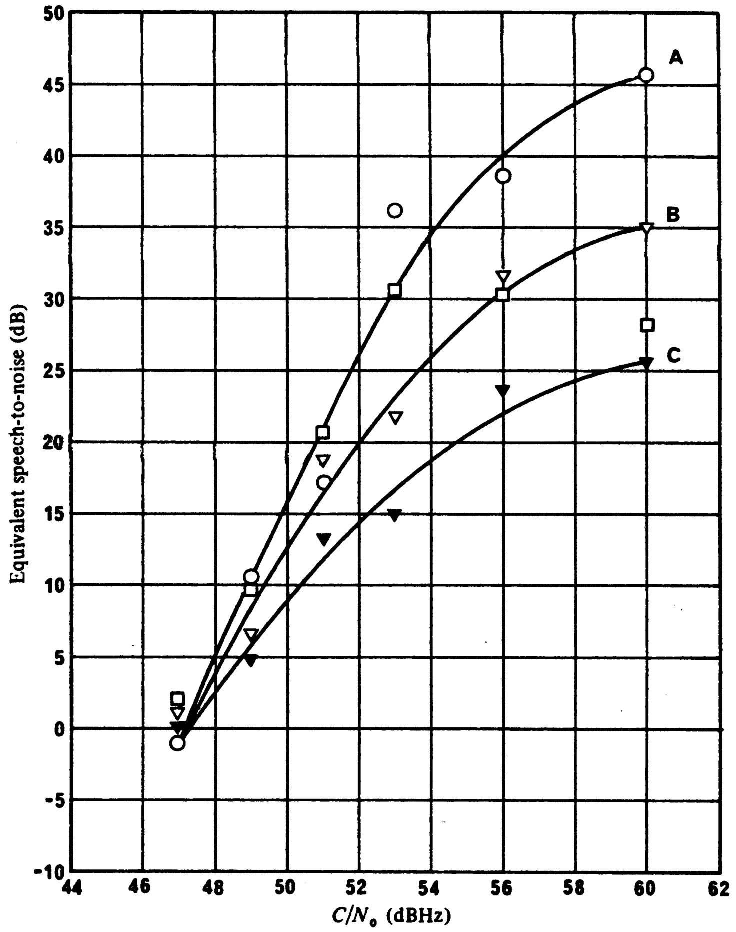


FIGURE 2 – *Results of subjective assessments in terms of equivalent noise of a –15 dBm0 talker as a function of $C/N_0$*

Talker level A, B and C (dBr) (r.m.s. power of talker relative to r.m.s. power for full deviation).

A: ○: -15 dBr

B: ▽: -25 dBr

C: ▼: -35 dBr

r.m.s. test-tone deviation: 8.5 kHz at 0 dBm0

Demodulator: threshold-extended type of regenerative-feedback IF limiter

Unfiltered 4-PSK provides a constant amplitude property, but has a high-level side lobe in its power spectrum. In the case of filtered 4-PSK the level of side lobes in the power spectrum can be adjusted to suit the coding method. An optimum exists between achieving minimum bit error rate and best spectrum utilization. Since filtered 4-PSK does not have a constant amplitude property non-linear amplification (such as class C) cannot be used, due to the spectrum regrowth which this would cause. However, offset 4-PSK can reduce this problem as it has less amplitude variation.

In MSK, the baseband rectangular signal in each channel is sinusoidally shaped in the time domain prior to the offset 4-PSK modulation process. In this sense, MSK can be recognized as a modification of offset 4-PSK. MSK has a constant envelope property with more concentrated power spectum than 4-PSK.

TFM and GMSK can substantially suppress the out-of-band spectrum at the cost of a moderate BER degradation, while their constant amplitude properties are almost maintained.

Figure 3 shows the calculated power spectra of signals modulated by the methods listed above. Figure 3 demonstrates that the spectral density distribution of TFM and GMSK is superior to the other modulation methods particularly for frequencies well separated from the carrier. For filtered 4-PSK and offset 4-PSK, a square-root Nyquist filter such as the square-root raised-cosine shaped filter can be used for premodulation, when a transmitter with a good linearity is used on board the mobile. In this case, filtered 4-PSK and offset 4-PSK can achieve a spectrum compactness comparable to or superior to TFM and GMSK without BER performance degradation.

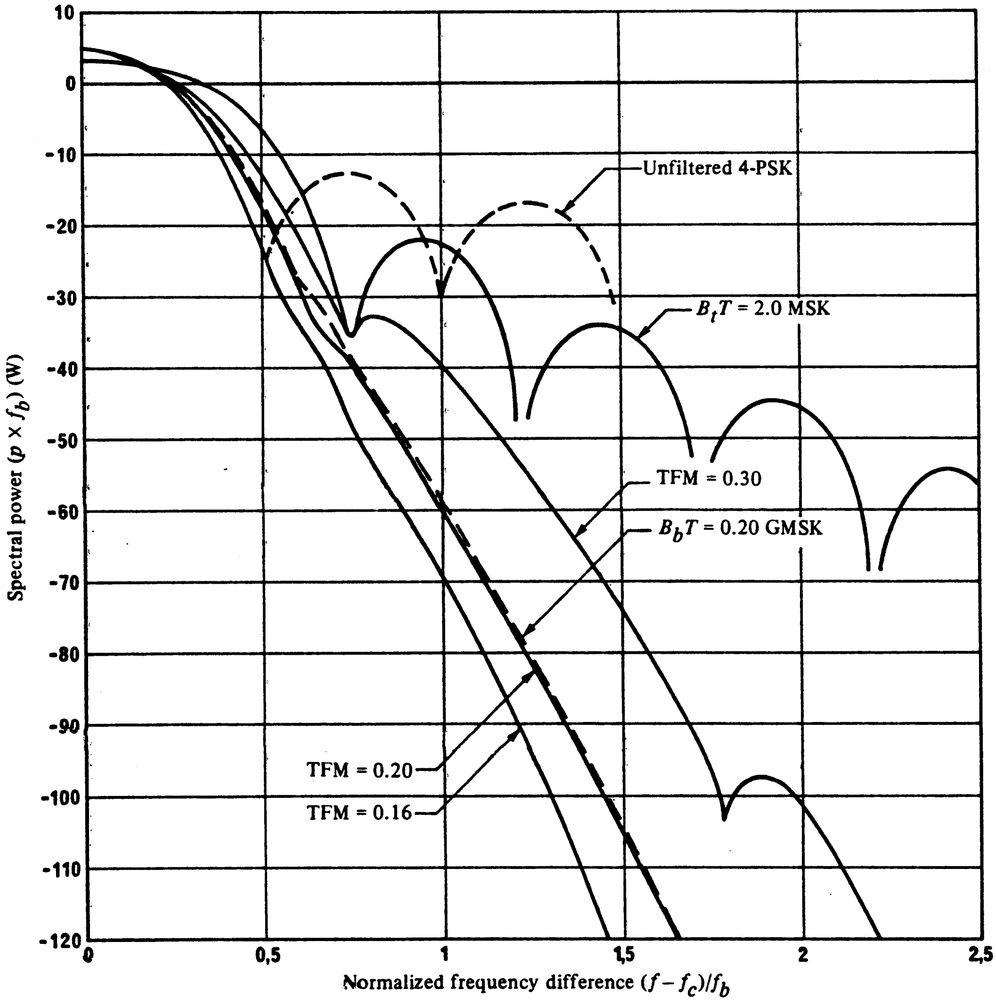


FIGURE 3 – *Power spectral density of typical modulation methods*

$B_bT$: normalized bandwidth of premodulation Gaussian bandpass filter (GMSK)

$B_tT$: normalized bandwidth of transmitting Gaussian bandpass filter for MSK

$f_c$ : carrier frequency (Hz)

$f_b$ : bandwidth (Hz)

$P$ : spectral density (W/Hz)

Figure 4 shows BER as a function of $E_b/N_0$ for typical modulation methods, and shows that theoretical 4-PSK performance can be used as a reference for the judgement of robustness of a modulation method against noise. Filtered 4-PSK can achieve the same BER performance as that for theoretical 4-PSK, when the transmission channel is linear and a square root Nyquist filter is used for premodulation and pre-detection filters, respectively.

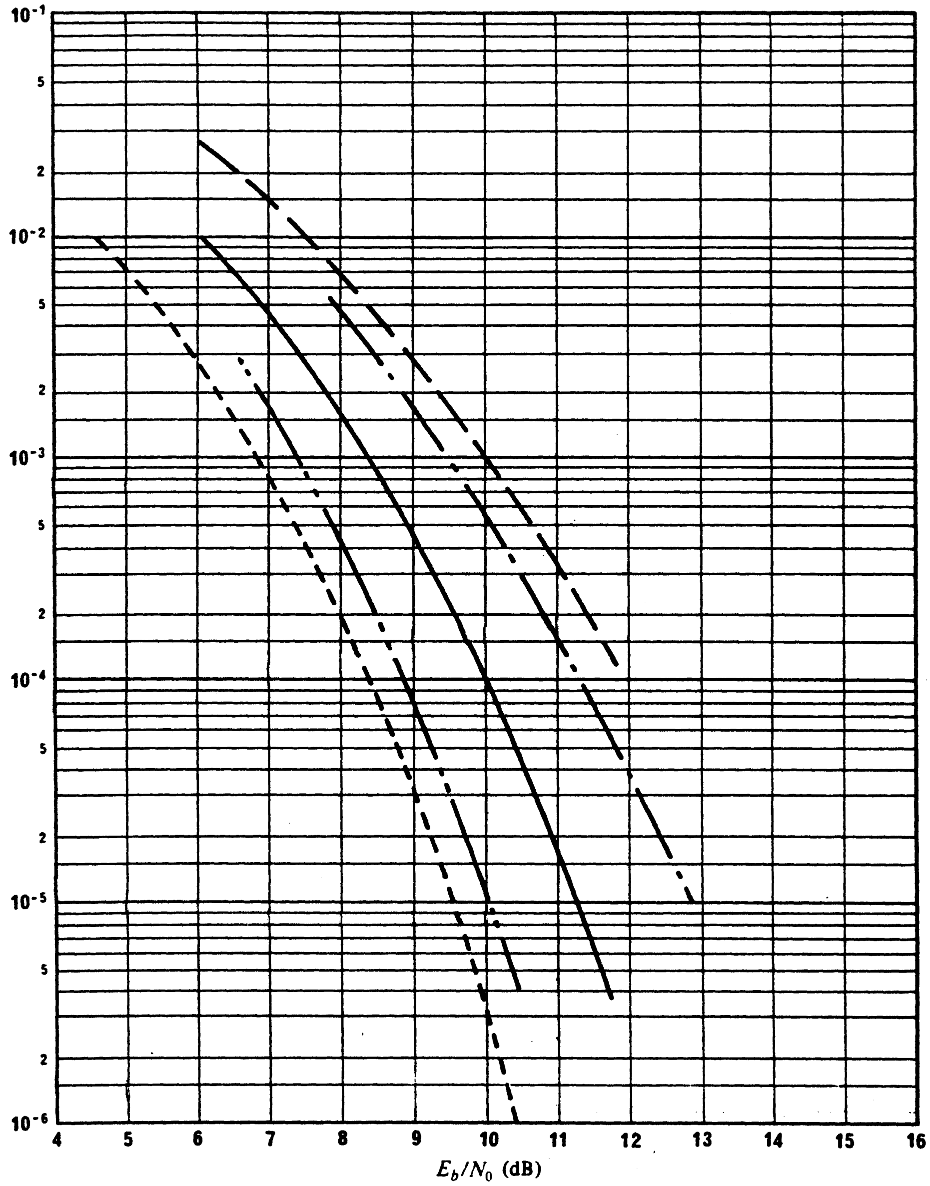


FIGURE 4 – *Bit error ratio as a function of normalized signal-to-noise ratio*

——  —— TFM (measured) [Muilwijk, 1979]

——— · —— GMSK $B_bT$ = 0.20 (measured)

—————— GMSK $B_bT$ = 0.25 (measured) and TFM (theoretical)

—— · · —— MSK $B_tT$ = 2.0 (measured) [Murota and Hirade, 1981]

– – – – – theoretical 2-PSK, 4-PSK and MSK

$B_tT$: normalized bandwidth of transmitting Gaussian bandpass filter for MSK

$B_bT$: normalized bandwidth of transmitting Gaussian low-pass filter in GMSK modulator

The BER performance in the actual satellite link can be degraded by a number of factors. They are modem circuit imperfections, band-limiting effect of transmitting and receiving filters, adjacent-channel interference, intermodulation interference and other miscellaneous interference. These factors are closely related to each other. For instance, narrower transmit filtering results in lower adjacent-channel interference level, but in general, results in a poorer BER performance. Another factor affecting BER performance is the susceptibility of the receiver to interference from other users of the same channel frequency. In this respect, it has been demonstrated that both 2-PSK and MSK have approximately the same sensitivity to co-channel interference and that they are both superior to 4-PSK.

The BER performance of digital modulation methods can be improved by the use of forward error-correcting (FEC) techniques, which is covered in Report 708. The final selection of modulation method should be made taking into account the parameters and requirements of each system such as channel spacing between carriers, required BER values, advantage to be gained from FEC and the effect of interference as listed above, taking into account expected level differences and the variation of carriers due to fading etc.

Table II briefly compares the relative merits and demerits of a number of modulation methods using the same parameters as those in Fig. 4.

TABLE II – *Relative merits and demerits among typical modulation methods*

| Modulation method | BER performance | Bandwidth utilization efficiency | Amplitude |
|---|---|---|---|
| Unfiltered 4-PSK | Excellent | Poor | Constant |
| Filtered 4-PSK | Excellent | Good | Varying |
| Offset 4-PSK | Excellent | Good | Varying |
| MSK | Excellent | Fair | Constant |
| TFM | Fair | Good | Constant |
| GMSK | Fair | Good | Constant |

*Note.* – Typically, 24 kbit/s filtered 4-PSK occupies 20 kHz bandwidth and the same bit rate MSK occupies 30 kHz bandwidth.

In the unique environment presented by low bit rates, e.g. 2.4 kbit/s, over a hard limited aircraft/satellite link, Symmetric BPSK (also known as Aviation Phase Shift Keying; A-BPSK) is a good candidate for this class of service. Tests with simulated modulation characteristics were conducted over a hard limited satellite link with a direct to multipath ratio of 10 dB and a fading rate of one sixth of the bit rate to compare the performance of Differentially Encoded Coherent PSK (DECPSK) chosen for its robust performance in a fading environment, to that of Symmetric BPSK. The results showed that the BER versus Eb/No performance of DECPSK was identical to that of Symmetric BPSK. However, Symmetric BPSK was far more spectrally efficient. Both modulation methods were superior, by 3 dB in Eb/No performance at a BER of $10^{-2}$, compared with filtered* Offset 4-PSK. Figure 5 shows the spectral sidelobe regrowth for Symmetric BPSK, DECPSK and Offset 4-PSK. Symmetric BPSK is more spectrally efficient than DECPSK while Offset 4-PSK is the most spectrally efficient but requires more power under the faded environment. [Lodge et al. 1987].

* Filtered with square root 50 percent raised-cosine filtering prior to hardlimiting.

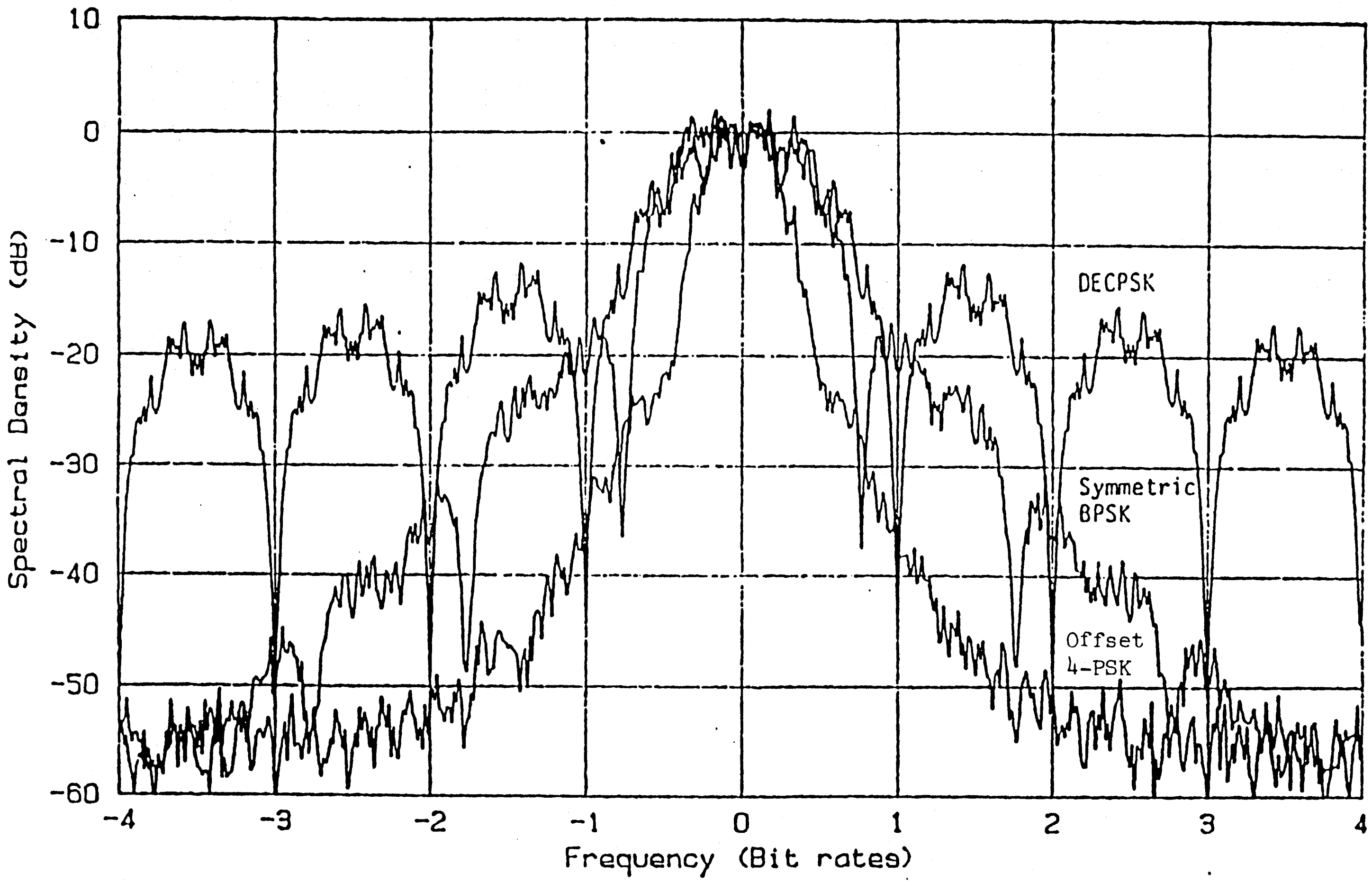


**FIGURE 5**

SPECTRAL SIDELOBE REGROWTH OF DECPSK, SYMMETRIC BPSK AND OFFSET 4-PSK AFTER HARDLIMITING AMPLIFIER

## 3. Voice coding technique

### 3.1 *General*

When a digital modulated transmission channel is used for telephone signal transmission, the analogue voice signal has to be coded into a digital data stream prior to modulation. 64 kbit/s pulse coded modulation (PCM), where the analogue voice signal is sampled at 8 kHz and every sample is quantized with 8 bits, is generally used for the public telephone services. However, the coding bit rate of 64 kbit/s is too high to be applied to the mobile satellite communications service, because the transmitting power from the satellite is severely limited and satisfactory BER performance cannot be achieved for such a high bit rate channel with reasonable satellite power. 32 kbit/s voice coding based on Adaptive Differential PCM (ADPCM) has been standardised in the CCITT (Recommendation G.721) and this provides a voice quality equivalent to conventional 64 kbit/s PCM but even this bit rate is undesirably demanding in satellite power and bandwidth and a lower bit rate is preferable. Furthermore, the voice coding technique used has to be robust against the channel bit errors, because the bit error ratio achievable is generally high in the mobile satellite services. It should also be noted that the final selection of coding technique will be affected by coding delay and relative equipment complexity.

*Note.* – Report 921 also provides information on this topic specifically for ship earth stations.

### 3.2 *Medium bit rate voice coding*

Possible candidates of waveform voice coding technique for mobile satellite services which may provide acceptable quality, are listed below:

- adaptive differential PCM (ADPCM);
- adaptive predictive coding (APC);
- adaptive transform coding (ATC);
- adaptive delta modulation (ADM);
- sub-band coding with ADPCM (SBC-ADPCM);
- sub-band coding with APC (SBC-APC);
- adaptive predictive coding with adaptive bit allocation (APC-AB) which is an advanced version of SBC-APC, but which requires more complicated hardware than conventional SBC-APC;
- multi-pulse excited coding (MPEC).

Mean opinion scores (MOS) can be used to evaluate the speech quality of voice coding techniques. The MOS is obtained by averaging the scores which are voted through subjective assessment tests having a 5-point scale with categories, i.e. "excellent", "good", "fair", "poor" and "bad". A comparison of the results of speech quality in MOS among typical coding methods under an error-free environment as a function of coding bit rate has been reported [Daumer, 1982]. In the literature, as an example of an actual circuit environment, a simulated subscriber loop of approximately 2 km with bandpass filtering of 200-3400 Hz, which simulates the bandpass filter in the channel bank, prior to coding, and a subscriber loop of approximately 2 km after decoding were considered. It was concluded that APC provided the highest MOS value at a coding bit rate of 16 kbit/s.

Another example comparing the speech quality of typical 16 kbit/s coding methods is shown in Fig. 6 [Yatsuzuka *et al.*, 1983]. In Fig. 6, the MOSs under a BER of $10^{-3}$ as well as those under an error-free environment are shown. MOS for NBFM as used in the INMARSAT Standard-A system and for standard PCM with various bit rates are also indicated. In Fig. 6, a 0.5 mm diameter 2 km simulated subscriber loop with circuit noise of 10 000 pW0p was taken into account, and bandpass filtering of 300-3400 Hz was also used. As shown in Fig. 6, 16 kbit/s APC and 16 kbit/s SBC-APC can achieve, under an error-free environment, a subjective quality comparable to NBFM with $C/N_0$ of 53 dBHz and, under a BER of $10^{-3}$, better quality than NBFM with $C/N_0$ of 48.6 dBHz which is the threshold link level $C/N_0$ for the faded condition in the INMARSAT Standard-A system design.

Annex I to Report 921 also discusses this specific topic.

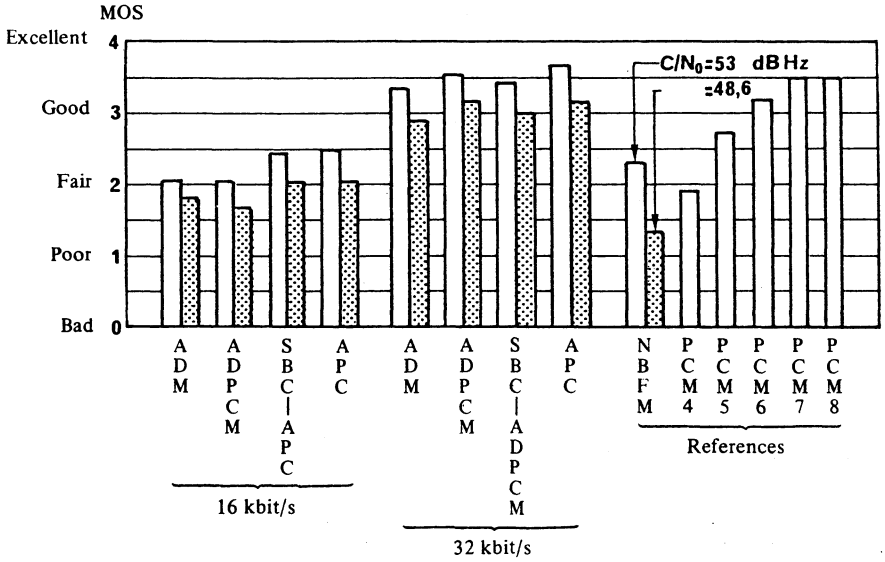


FIGURE 6 – *MOSs for typical* ———— *16 kbit/s voice coding methods* [Yatsuzuka *et al.*, 1983]

Error free

Random error (BER = $10^{-3}$)

PCM4: 4 bit/sample PCM
PCM5: 5 bit/sample PCM
PCM6: 6 bit/sample PCM
PCM7: 7 bit/sample PCM
PCM8: 8 bit/sample PCM

3.3 *Low bit rate voice coding technique*

At coding bit rates below 16 kbit/s and specifically the bit rate range 9.6 to 4.8 kbit/s, the waveform coding method can provide reproduced speech which is highly intelligible, but which has noticeable quality degradation with some detectable distortion, and perhaps reduced talker recognition. At coding bit rates in the range 4.8 kbit/s and below, source coding (vocoder) methods can provide synthetic quality, although the signal usually has lost substantial naturalness, typically sounding reedy [Flanagan *et al.*, 1979].

Possible candidates for voice coding methods in this coding bit rate range are listed below [Flanagan *et al.*, 1979]:

- APC,
- ATC,
- MPEC,
- voice excited linear predictive coding (VELP),
- residual excited linear predictive (RELP) coding,
- channel vocoder (CV),
- formant vocoder,
- linear predictive coding (LPC).

The first three are forms of waveform coding and the last three are forms of vocoder. VELP and RELP are categorized as hybrid coding method.

In the digital ship earth station systems with medium-gain antenna and the aeronautical satellite communication systems with airborne antenna of limited gain, severe power limitations suggest that voice coding techniques at bit rates of 9.6 kbit/s or lower would be necessary. The provision of acceptable voice quality is a requirement for the public telephone service and waveform coding methods such as APC /_Yatsuzuka et al., 1986_/ and MPEC /_Taguchi et al., 1984_/ are possible candidates for these systems.

With a vocoder, only essential speech parameters are extracted and transmitted. Vocoders consist of a speech analyzer part, which extracts speech parameters, and a speech synthesizer which reconstitutes the original speech on the basis of the received parameters. The speech analyzer of a pitch excited vocoder tracks the continuous variation of the speech pitch frequency, indicates the presence of voiced or unvoiced sounds and provides information on the short-term speech spectrum. Depending on the way in which the short-term speech spectrum is determined, vocoders are classified as belonging to various categories, e.g. channel-, formant- or LPC (linear predictive coding) vocoder [Flanagan *et al.*, 1979]. A channel vocoder uses filter banks to extract rectified smoothed amplitude information for each band. A formant vocoder also uses filter banks to detect the instantaneous frequency and the average amplitude for each band. An LPC vocoder extracts the so-called vocal tract parameter by minimizing mean square error between the input signal and signal values predicted by a linear weighted summation of past values of the signal.

Two particular points of interests are:

- there is a strong relation between the voice quality of a vocoder and the accuracy of pitch extraction; and
- the extraction of the pitch frequency from a speech signal is one of the most critical and difficult tasks in speech analysis.

In this latter regard, a fairly recent major achievement reported in literature is a new type of pitch extractor which forms the basis of a high-quality type of pitch-excited vocoder at low bit rate (Sluyter *et al.*, 1980].

Rather than looking at pitch extraction as a purely technical problem, this particular technique is based on the understanding of human pitch perception. The indicated concept includes spectral analysis and harmonic pattern recognition. Using a well-chosen algorithm, the pitch parameter can be calculated in agreement with perception and in real-time.

An example comparison result of speech quality in MOS among typical coding methods can be seen in Fig. 7 [Yato *et al.*, 1983]. Also shown are comparative MOS for narrow-band FM at particular $C/N_0$ values and for a 4-bit PCM system.

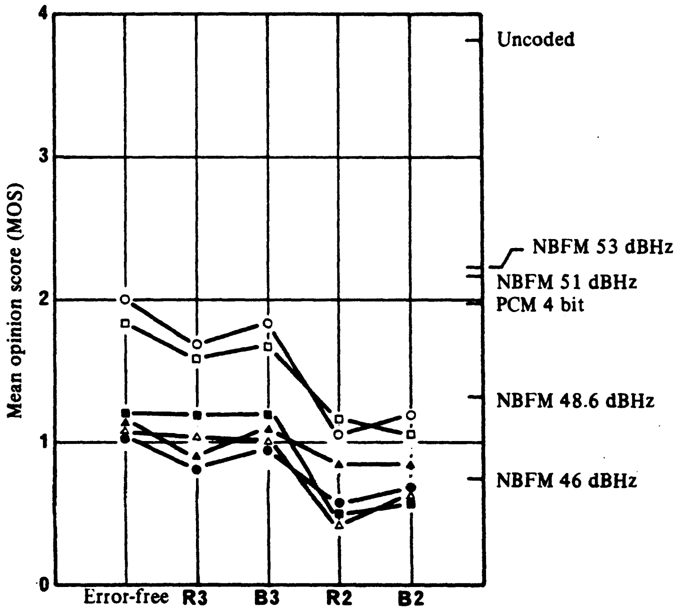


FIGURE 7 – *MOSs for typical low bit rate voice coding methods*
[Yato *et al.*, 1983]

○ RELP 9.6 (kbit/s)

□ RELP 7.2

▲ RELP 4.8

■ LPC 4.8

△ LPC 2.4

● CV 2.4

R3: random errors of BER = $10^{-3}$ were added

B3: burst errors of BER = $10^{-3}$ were added

R2: random errors of BER = $10^{-2}$ were added

B2: burst errors of BER = $10^{-2}$ were added

Regarding the speech quality of LPC vocoders, some intelligibility tests have been carried out with several types of LPC vocoders developed in Japan. For instance, it is reported that one type of 2.4 kbit/s LPC vocoder achieves a syllable articulation score of about 68% in Japanese [Kurematsu *et al.*, 1979] and another type of LPC vocoder PARCOR [Itakura and Saito, 1968] obtains AEN values (see Note) (CCITT Recommendation P.41-45) of 6 dB and 3 dB at 2.4 kbit/s and at 4.8 kbit/s, respectively.

Regarding the performance of channel vocoders, intelligibility tests have been carried out on 2.4 kbit/s channel vocoders (1.8 kbit/s vocoder with 0.6 kbit/s FEC) developed in the United Kingdom using phonetically balanced word lists [Lehiste and Peterson, 1959] which resulted in percentage intelligibility approaching 80% even for BER up to 5%. It is generally accepted that a minimum acceptable score for such tests is 55-60%. Tests have been carried out in practical environments and in noise chamber simulations on the suitability of channel vocoders for use in vehicles with very high levels of acoustic noise, considerably worse than would be expected in a ship, and they were found usable with high acoustic noise combined with a BER of 2.5% (with FEC) [Kingsbury and Amos, 1980]. In these tests on channel vocoders the microphones used were practical types in common use; however carbon microphones were not used. The FEC used was designed specifically for use with the vocoder, and only applies correction to the most significant bits i.e. those that affect the speech most.

It can be expected that high-quality vocoder systems will be available in the foreseeable future. However, with regard to the applicability of vocoders in the mobile-satellite services, further study is required on several aspects, e.g.:

- influence of a limited-quality input speech to the vocoder, as will be experienced especially in the shore-to-ship direction (band-limitation, range of talker levels, impulse noise, carbon microphones, etc.);
- sensitivity to transmission errors (BER-performance);
- influence of accoustic background noise.

It is noted that subjective tests will be necessary to evaluate the overall both-way conversational quality of the systems under study.

*Note.* – AEN (Affaiblissement équivalent pour la netteté; Articulation reference equivalent) value is the difference between the attenuation level of the reference circuit and that of a circuit under test, each of which is adjusted so that a sound articulation score of 80% is obtained.

## 4. Data coding techniques

Very low $G/T$ mobile earth stations (e.g. $-24$ dB($K^{-1}$)) may suffer from bad propagation conditions (multipath at low elevation, shadowing effects even at higher elevation). A communication system which is designed to cope with all possible link conditions including those occurring at low elevation angles, may lead to protocol and coding overheads which are more powerful than necessary for much of the time. One way to improve the overall system efficiency is to utilize adaptive coding. The description and evaluation of such a coding method as applied to low $G/T$ maritime, aeronautical and land mobile earth stations is presented in Annex II. The concept under consideration relies on the use of short Reed-Solomon block codes with a bi-dimensional arrangement and the implementation of an ARQ scheme. As can be seen in Annex II, the proposed adaptive coding scheme leads to a greater efficiency in normal operation than convolutional encoding.

## REFERENCES

BERANEK, L. L. [September, 1947] The design of speech communication systems. *Proc. IRE*, Vol. 35, 880-890.

DAUMER, W. R. [April, 1982] Subjective evaluation of several efficient speech coders. *IEEE Trans. Comm.*, Vol. COM-30, **4**, 655-662.

DE BUDA, R. [June, 1972] Coherent demodulation of frequency-shift keying with low deviation ratio. *IEEE Trans. Comm.*, Vol. COM-20, **3**, 429-434.

DE JAEGER, F. and DEKKER, C. B. [May, 1978] Tamed FM, a novel method to achieve spectrum economy in digital transmission. *IEEE Trans. Comm.*, Vol. COM.26, **5**, 534-542.

FLANAGAN, J. L., SCHROEDER, M. R., ATAL, B. S., CROCHIERE, R. E., JAYANT, N. S. and TRIBOLET, J. M. [April, 1979] Speech coding. *IEEE Trans. Comm.*, Vol. COM-27, **4**, 710-737.

HIRADE, K. and MUROTA, K. [March, 1979] A study of modulation for digital mobile technology. 29th Annual Conference of the IEEE Vehicular Technology Society, Arlington Heights, IL, USA, Conf. Record, Session 1A.

ITAKURA, F. and SAITO S. [1968] Analysis synthesis telephony based on the maximum likelihood method. Ed. Y. Kohashi. Reports of 6th International Congress on Acoustics, Tokyo, C-5-5.

KINGSBURY, N. G. and AMOS, W. A. [9 April, 1980] A robust channel vocoder for adverse environments. Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '80), 9-11 April, Denver, CO, USA, Vol. 1, Session S1: Narrowband speech, S1.4, 19-22.

KUREMATSU, A., ISHIGAMI, H., KITAYAMA, S., YATO, F. and TAMURA, J. [2 April, 1979] A linear predictive vocoder with new pitch extraction and exciting source. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '79), 2-4 April, Washington DC, USA, Conf. Record, Session 2: Narrowband speech communication, 2.8 69-76.

LEHISTE, I. and PETERSON, G. E. [March, 1959] Linguistic considerations in the study of speech intelligibility. *J. Acous. Soc. Amer.* Vol. 31, **3**, 280-286.

LODGE, J.L., MOHER, M.L., GROZIER, S.W. - A comparison of data modulation techniques for land mobile satellite channels. IEEE Transactions on vehicle technology. Vol. VT-36 No.1 February 1987 Pages 28-35.

MUILWIJK, D. [March, 1979] Tamed frequency modulation – a bandwidth-saving digital modulation method suited for mobile radio. *Philips Telecomm. Rev.*, Vol. 37, **1**, 35-49.

MUILWIJK, D. and NOORDANUS, J. [1980] Digital phase modulation methods giving a band-limited spectrum for satellite communications. Presented at the 8th AIAA Communications Satellite System Conference, Orlando, FL, USA.

MUROTA, K. and HIRADE, K. [October, 1981] Transmission performance of GMSK modulation (in Japanese). *Trans. Inst. Electron. Comm. Engrs. Japan*, Vol. J64-B, **10**, 1123-1130.

SLUYTER, R. J., KOTMANS, H. J. and LEEUWAARDEN, A. V. [6 March, 1980] The harmonic-sieve method for pitch extraction from speech and a hardware model applicable to vocoder systems. Proc. IEEE International Zurich Seminar on Digital Communications, 4-6 March, Switzerland, E2.1-E2.6.

TAGUCHI, T. and KOBAYASHI, M. [October, 1984] Real-time implementation of multipulse coding systems. Trans. of Committee on Speech Research. The Acoustical Society of Japan, S84-49, pp.381.

YATO, F., KITAYAMA, S., YATSUZUKA, Y. and KUREMATSU, A. [June, 1983] Performance evaluation of voice coding schemes applicable to INMARSAT Standard-B system. 3rd IEE International Conference on Satellite Systems for Mobile Communications and Navigations, London, UK, Conf. Record, 162-166.

YATSUZUKA, Y., YATO, F. and KUREMATSU, A. [September, 1983] Application of 32 and 16 kbit/s speech encoding techniques to digital satellite communications. 6th International Conference on Digital Satellite Communications, Phoenix, AZ, USA, Conf. Record, VII-B.

YATSUZUKA, Y., IIZUKA, S. and YAMAZAKI, T. [April 1986] A variable rate coding by APC with maximum likelihood quantization from 4.8 kbit/s to 16 kbit/s. Int. Conf. Acoustics, Speech and Signal Processing, Tokyo, pp.56.12.1 - 56.12.4.

# ANNEX I

## RESULTS OF SUBJECTIVE TESTS ON VOICE CODING TECHNIQUES

### 1. Introduction

This Annex presents the results of a subjective test programme on voice coding techniques in the range 7.2-16 kbit/s, undertaken by INMARSAT in conjunction with ten of its signatories. This programme was part of an overall study on candidate digital voice coding techniques for potential future INMARSAT ship earth-station applications (see also Report 921).

The objectives of the tests were to assess the subjective voice quality of the techniques and determine possible language sensitivities of the different countries.

### 2. Voice coding techniques

The five digital voice coding techniques investigated and the associated channel test conditions are summarized in Table III. These techniques were implemented either in the form of hardware codecs or by means of computer simulation; in the case of sub-band coding (SBC), the algorithm used in the simulation differed from that in the hardware SBC coder. Also included in the tests were reference conditions of speech correlated noise (modulated noise reference unit: MNRU) at specific values of signal-to-noise ratio ($S/N$).

### 3. Test procedures

The ten participating INMARSAT signatories were from France, the Federal Republic of Germany, Greece, Italy, Japan, the Netherlands, Norway, Sweden, the UK and the USA.

Each participant prepared a national language source tape which was then processed with the voice coding techniques and test conditions shown in Table I. Processed tapes were then returned to each participant for subjective evaluation by a number of listener subjects (between 12 and 56), on a mean opinion score (MOS) basis similar to that described in § 3.2 of Report 751, but according to the following scale:

| *Rating* | *Score* |
|---|---|
| Excellent | 4 |
| Good | 3 |
| Fair | 2 |
| Poor | 1 |
| Bad | 0 |

TABLE III – *Voice coding techniques and test conditions*

| Technique | Bit rate (kbit/s) | BER or $S/N$ (dB) | Code |
|---|---|---|---|
| Adaptive predictive coding (simulation) | 16 | BER = 0<br>$10^{-3}$<br>$10^{-2}$ | APC 0<br>APC 3<br>APC 2 |
| Sub-band coding (simulation: S) | 16 | BER = 0<br>$10^{-3}$<br>$10^{-2}$ | SBC 16 0(S)<br>SBC 16 3(S)<br>SBC 16 2(S) |
| Sub-band coding (hardware: H) | 16 | BER = 0<br>$10^{-3}$<br>$10^{-2}$ | SBC 16 0(H)<br>SBC 16 3(H)<br>SBC 16 2(H) |
| Sub-band coding (hardware) | 9.6 | BER = 0<br>$10^{-3}$<br>$10^{-2}$ | SBC 9 0<br>SBC 9 3<br>SBC 9 2 |
| Residual excited linear predictive coding (simulation) | 7.2 | BER = 0<br>$10^{-3}$<br>$10^{-2}$ | RELP 0<br>RELP 3<br>RELP 2 |
| Reference conditions of speech correlated noise | (MNRU) | $S/N$ = 5<br>10<br>15<br>20<br>25<br>30 | MNR 05<br>MNR 10<br>MNR 15<br>MNR 20<br>MNR 25<br>MNR 30 |

## 4. Test results

Table IV shows the average MOS for each voice coding technique and test condition for the participant and the average for all participants. The following observations are made on these results (see also Fig. 8):

- the adaptive predictive coder (APC) at 16 kbit/s is clearly the best coding scheme of those tested. The influence of transmission errors on quality up to a BER of $10^{-3}$ appears to be very small. The speech quality remains "fair" at a BER of $10^{-2}$ which is encouraging for the future possible implementation of Standard-B ship earth-station codecs;
- the performance of the hardware sub-band coder at 16 kbit/s appears to have been comparable to the simulated version under the error-free condition, but less resistant to transmission errors; however, the algorithm used in each implementation was different, because they were provided from different sources;
- the particular 9.6 kbit/s sub-band coder used in these tests would appear to exhibit poor performance;
- the residual excited linear predictive coder (RELP) at 7.2 kbit/s tolerates transmission errors well, but its overall performance is rather poor.

Table V shows the country-to-country correlations for MOS results on the MNRU test conditions. These correlations are generally high, as were those for the voice coding techniques (not shown in this Annex). Using the MNRU conditions, it would thus be possible to reliably predict results in one country based on measurements performed in another country.

It was not possible to isolate language or cultural factors from other aspects of the test situation specific to the country providing test results (such as recording and listening equipment or interpretation of opinion scales). Accordingly, it is not possible to say that a particular type of coder performs better with one language than with another. The overall results, however, suggest that the implementation of 16 kbit/s voice coding, particularly with APC, could provide acceptable speech quality for the languages tested.

TABLE IV – *Average MOS results*

| Condition | Code | USA | Norway | Japan | Sweden | UK | Netherlands | Greece | France | Germany (Federal Republic of) | Italy | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 16 kbit/s adaptative predictive coding | APC 0 | 3.17 | 2.46 | 2.82 | 2.88 | 3.04 | 2.58 | 2.9 | 3.11 | 3.05 | 2.70 | 2.87 |
| | APC 3 | 2.93 | 2.46 | 2.52 | 2.79 | 2.92 | 2.48 | 2.8 | 2.95 | 2.81 | 2.63 | 2.72 |
| | APC 2 | 1.93 | 1.97 | 1.39 | 2.13 | 1.98 | 1.10 | 1.8 | 1.90 | 2.05 | 1.43 | 1.75 |
| 16 kbit/s sub-band coding (simulation: S) | SBC 0(S) | 2.39 | 2.44 | 1.98 | 2.56 | 2.55 | 2.01 | 2.1 | 2.91 | 2.51 | 2.44 | 2.39 |
| | SBC 3(S) | 2.25 | 1.79 | 1.61 | 2.34 | 2.17 | 1.65 | 2.2 | 2.38 | 2.56 | 2.07 | 2.10 |
| | SBC 2(S) | 1.32 | 1.12 | 0.70 | 1.42 | 1.22 | 0.91 | 1.1 | 1.27 | 1.14 | 0.91 | 1.11 |
| 16 kbit/s sub-band coding (hardware: H) | SBC 16 0(H) | 2.44 | | 2.09 | 2.44 | 2.70 | 2.19 | 2.6 | 2.73 | 2.71 | 2.28 | 2.46 |
| | SBC 16 3(H) | 1.61 | | 1.36 | 1.61 | 1.78 | 1.38 | 1.0 | 2.08 | 1.03 | 1.30 | 1.46 |
| | SBC 16 2(H) | 0.27 | | 0.04 | 0.65 | 0.17 | 0.08 | 0.1 | 0.16 | 0.05 | 0.12 | 0.18 |
| 9,6 kbit/s sub-band coding | SBC 9 0 | 0.93 | | 0.74 | 0.95 | 0.72 | 0.74 | 1.2 | 0.88 | 0.91 | 0.56 | 0.85 |
| | SBC 9 3 | 0.21 | | 0.24 | 0.34 | 0.52 | 0.21 | 0.2 | 0.26 | 0.22 | 0.33 | 0.28 |
| | SBC 9 2 | 0.03 | | 0.02 | 0.0 | 0.0 | 0.0 | 0.0 | 0.02 | 0.01 | 0.01 | 0.01 |
| 7,2 kbit/s RELP (vocoder) | RELP 0 | 1.61 | 1.66 | 1.30 | 1.70 | 1.94 | 1.19 | 1.8 | 1.95 | 2.10 | 1.43 | 1.66 |
| | RELP 3 | 1.57 | 1.44 | 1.09 | 1.68 | 1.97 | 1.10 | 1.5 | 1.95 | 1.86 | 1.07 | 1.52 |
| | RELP 2 | 0.32 | 0.12 | 0.08 | 1.32 | 0.44 | 0.23 | 0.2 | 0.24 | 0.21 | 0.15 | 0.33 |
| Modulated noise reference unit | MNR 05 | 0.72 | 0.42 | 0.63 | 0.79 | 0.92 | 0.43 | 0.6 | 0.63 | 0.43 | 0.72 | 0.63 |
| | MNR 10 | 1.44 | 0.77 | 1.25 | 1.19 | 1.53 | 0.83 | 1.3 | 1.42 | 1.20 | 0.95 | 1.19 |
| | MNR 15 | 2.24 | 1.12 | 2.07 | 1.59 | 1.97 | 1.59 | 1.8 | 2.16 | 1.46 | 2.00 | 1.80 |
| | MNR 20 | 3.14 | 2.39 | 2.72 | 2.56 | 2.68 | 2.39 | 3.0 | 3.20 | 2.82 | 2.66 | 2.76 |
| | MNR 25 | 3.38 | 2.48 | 3.07 | 3.17 | 3.10 | 2.58 | 2.9 | 3.37 | 3.15 | 3.15 | 3.03 |
| | MNR 30 | 3.18 | 3.25 | 3.19 | 3.46 | 3.02 | 2.83 | 3.0 | 3.27 | 3.03 | 2.56 | 3.07 |

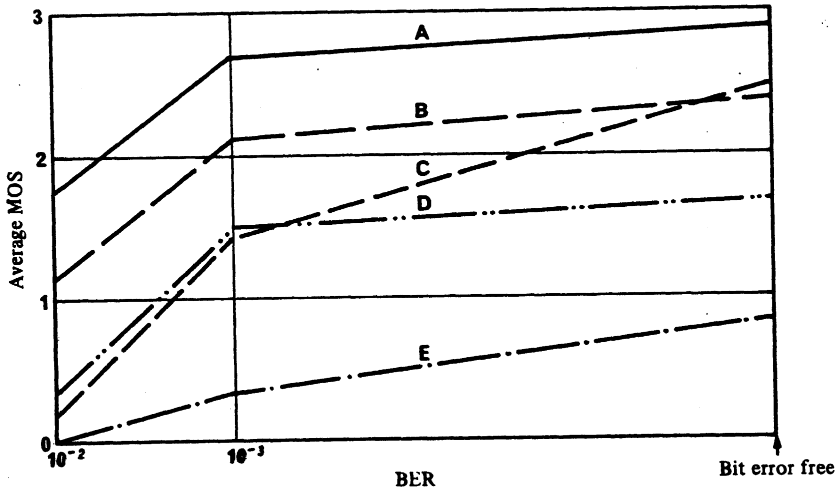


FIGURE 8 – *Average MOS for all participants versus BER*

Curves A: APC 16 kbit/s
B: SBC 16 kbit/s (simulation)
C: SBC 16 kbit/s (hardware)
D: RELP 7.2 kbit/s
E: SBC 9.6 kbit/s

TABLE V – *Country-to-country correlation of mean opinion score for MNRU reference conditions*

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| USA | 1.00 | | | | | | | | | |
| NOR | 0.924 | 1.00 | | | | | | | | |
| JPN | 0.990 | 0.952 | 1.00 | | | | | | | |
| SWE | 0.945 | 0.985 | 0.969 | 1.00 | | | | | | |
| UK | 0.989 | 0.952 | 0.993 | 0.977 | 1.00 | | | | | |
| NET | 0.983 | 0.972 | 0.995 | 0.978 | 0.987 | 1.00 | | | | |
| GRE | 0.989 | 0.952 | 0.981 | 0.951 | 0.982 | 0.984 | 1.00 | | | |
| FRA | 0.999 | 0.956 | 0.991 | 0.951 | 0.991 | 0.986 | 0.994 | 1.00 | | |
| GER | 0.980 | 0.959 | 0.975 | 0.974 | 0.991 | 0.977 | 0.987 | 0.985 | 1.00 | |
| ITA | 0.980 | 0.876 | 0.964 | 0.916 | 0.964 | 0.957 | 0.951 | 0.972 | 0.949 | 1.00 |
| | USA | NOR | JPN | SWE | UK | NET | GRE | FRA | GER | ITA |

# ANNEX II

## AN ADAPTIVE CODING SCHEME FOR DATA ONLY MOBILE TERMINAL

## 1. Introduction

This Annex describes a coding scheme used in the ESA PRODAT data only communication system based on a $G/T$ of $-24$ dB($K^{-1}$) which is appropriate for land, maritime and aeronautical applications.

Several types of services, and/or functions are provided in addition to the usual system organization function requirements (channel assignment, random access etc.). The system provides, on a primary basis, a message handling service (electronic mail box system), but also broadcast/paging mode, polling mode, and possibly dialogue modes.

The PROSAT Phase I [PROSAT, 1985] propagation experiments demonstrated that the channel quality for all the three applications varies from very good conditions (almost non-faded channel), corresponding to open area for a land mobile system or "reasonable" elevation angles for the maritime (see Report 884) and aeronautical channels, to extremely bad situations, characterized by a $C/M$ ratio of 7 dB for the maritime channels, or complete black-out in the case of the land-mobile application.

In order to accommodate satisfactory operation over extreme propagation conditions whilst preserving system efficiency in normal conditions, it is proposed to implement an ARQ scheme which provides for auto-adaptability.

## 2. Description of the selected coding scheme

### 2.1 *Bi-dimensional coding scheme*

Short Reed-Solomon (RS) block codes with a bi-dimensional arrangement have been selected.

The information is organized in "blocks", each block comprises a set of 8 elementary Reed-Solomon code vectors. In addition each block is encoded in the other dimension, using again RS coding technique, as follows (see Fig. 9):

- in one dimension (vertical): one extended RS code forms a vector ($k$ data symbols, $n - k$ redundancy symbols) and is used to correct random errors, and to detect error burst which exceeds the capability of the code;
- in the other dimension (horizontal): another extended RS code ($k'$ primary vectors and $n - k'$ redundancy vectors) is used to correct burst errors, taking advantage of the erasure correction capability of the code;
- as an example, with a forward link having the structure of a TDM frame, each vertical vector could correspond to a slot in the frame (see Annex VI to Report 921); if the vertical code vector transmission time is short with respect to the fading bandwidth, deep fading will correspond to erasure of a complete code vector. The "vertical code" has to be used to correct the random errors. In addition, as it is in any case necessary to use the channel state information to detect that fading situation, which may exceed the error detection capability of the code itself, the channel state information could also be used to extend the capabibilty of the code to its maximum by using the erasure correction technique.

This scheme, like any bi-dimensional code can be considered as an interleaving scheme for the horizontal code, and as a consequence, is inherently capable of coping with error bursts. The method of adapting the scheme of ARQ follows.

### 2.2 *RS bi-dimensional coding + ARQ*

In a "classical" ARQ system, one asks for a retransmission, when a vertical vector erasure situation would have been detected. The transmission delay in the particular case of satellite systems, imposes either the numbering of the blocks, or/and restriction in the selection of frame parameters. In addition, the quality of the return channel is not better than the direct channel, and therefore, the acknowledgement procedure is by itself a source of difficulty. It should therefore be limited as far as possible, or used in a safe way.

The basic principle of the scheme is as follows:

- in the vertical direction, no modification;
- in the horizontal direction (time axis) the selected code is a rate ½ code, i.e. a (16 : 8) code, which means that the block is a set of $k' = 8$ vectors of information, plus additional redundancy vectors, as required. The decoding in that direction is performed by considering that a vertical vector is either correct (after having applied the correction algorithm) or erased. With the selected code, it is possible to decode a code vector which has up to 8 erasures, or in other words, we can say that as soon as we have received succesfully 8 vectors, we have enough information for the decoding. We can, then, acknowledge the reception of this set of information vectors and proceed with the next set. The only risk is that the transmitter sends one or two information blocks more than required.

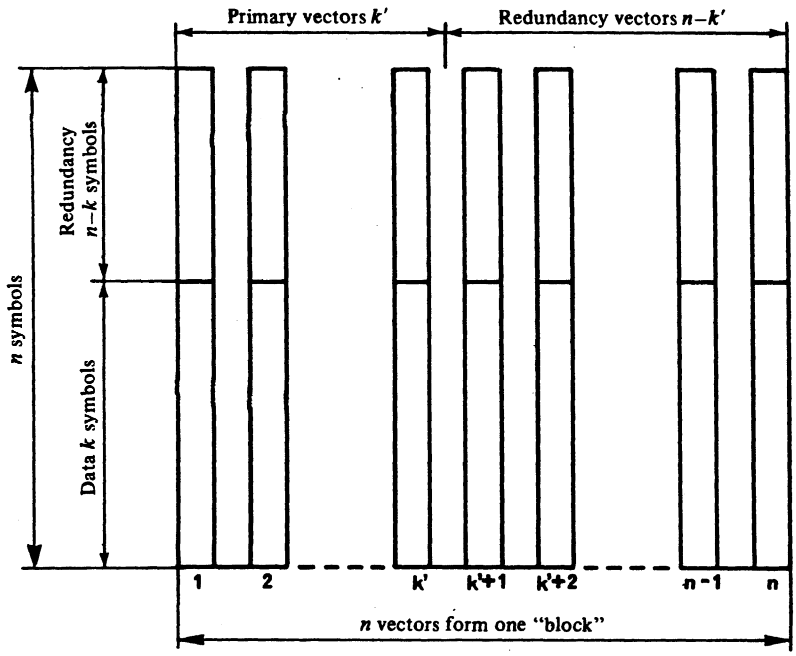


FIGURE 9

Theoretically, the vertical vector throughput must be at least 80% to achieve a block transmission failure rate of $10^{-3}$. As a result a (16 : 12) code for the forward link (which has a TDM structure) was adopted, e.g. a rate as large as possible. For the return link, a lower rate code is required and therefore a (16 : 10) code is used, because the vertical vectors are not interleaved but are transmitted sequentially.

## 3. Comparison against convolutional coding techniques

The main advantage of convolutional coding techniques is the possibility to implement a soft decision decoder which provides some additional coding gain. As such it could be considered as a good reference for the evaluation of the block code scheme, in particular with respect to the vertical vector throughput. However, the convolutional decoding technique is not necessarily well adapted to the particular channels being considered. The classical Viterbi decoder performance is impaired by the error bursts, in general, and differential detection at low $E_b/N_0$s.

In order to reduce this degradation it is necessary to introduce a high degree of interleaving to break up the errors. For the maritime channel this is possible by using interleaving periods of many seconds. However, for the land mobile channel where blockage of the signal can occur for undefined periods, it is impractical. Interleaving also suffers from the problem of spreading errors into good data.

Therefore, coherent detection with sign ambiguity resolution by a unique word at the start of the message, is required. The length of a data block between unique words is governed by the probability of a cycle skipping, leading to phase reversals during a block. In order to evaluate the performance of such a scheme, a block length of 100 bits was selected as a reasonable size. Within each block a unique word of 16 bits is used to resolve the phase ambiguity and synchronize the Viterbi decoding. A pattern of 8 tailing zeros is added to the data to flush out the encoder, and to form an uncoded data pattern 42 bits long. The data is ½ rate convolutionally encoded, using a constraint length of 5, to form a coded data pattern 84 bits long, to which the unique word is added.

A simulation was carried out for the return link, at a $C/M$ value of 7 dB, using soft decision decoding as shown in Fig.10. If we define the threshold as in the case of the RS block code for a $C/M$ of 7 dB, and a block throughput of 80%, the required $E_b/N_0$ is 7 dB for the convolutionally coded return link, a value somewhat lower than for the RS block code link. However, a selective repeat ARQ scheme would require that each encoded block is received correctly, and therefore the block throughput would have to be nearer 90% for an overall efficiency of around 80%. This would require an $E_b/N_0$ of nearer 10 dB which is significantly higher than the RS block code scheme.

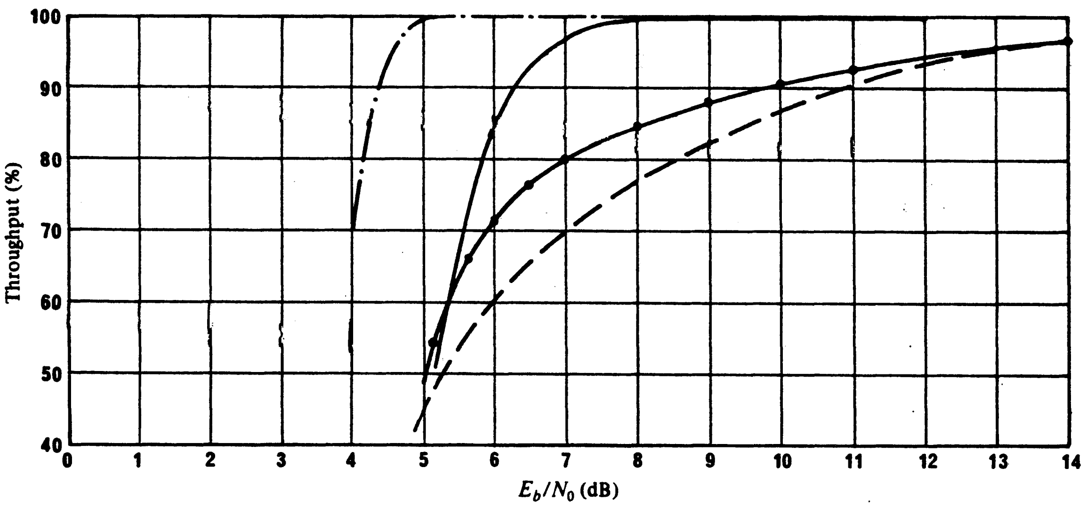


FIGURE 10 – *Throughput as a function of $E_b/N_0$ (300 Bd – fading bandwidth = 3 Hz) (percentage of error-free decoded vectors or blocks)*

Rician fading $C/M$ = 7 dB:

—•—•— Viterbi $k$ = 5, coherent demodulation, 1/2 rate, 100 bit blocks

— — — RS vectors

——— RS blocks (8 successful RS vectors)

Non-faded condition:

·—·—· RS blocks

When considering the overall system efficiency as defined in the case of block code, the value of 50% corresponding to the half rate code is further reduced when considering the overhead which is imposed by the need for a unique word and additional tailing bits. In addition the basic Viterbi decoder does not detect residual errors. The convolutional code has to be concatenated with a short block code to detect these residual errors. The channel state information would also most likely be used.

Thus the complete overall system efficiency for the convolutionally coded return link would be nearer 30%, assuming an error detection code equal to 10% of the data and with the above defined overheads. This is to be compared with the 35% efficiency experimentally found on the Reed-Solomon scheme (16 : 10 code). On the forward link, the convolutional code could never match the derived Reed-Solomon efficiency of 60%. Also the convolutional code has a fixed coding rate, whilst the Reed-Solomon scheme has an adaptive coding rate which leads to a greater system efficiency in normal operation.

It is interesting to note that the curves showing the non-faded and Rician channel ($C/M$ = 7 dB) throughputs for the RS block code are only 2 dB apart, demonstrating the high performance of the scheme.

It is recognized that the convolutional link efficiency could be increased by using a higher rate code, for instance the punctured codes whose performances have also been estimated through experimental simulation. However, this will have an opposite effect on the $E_b/N_0$ threshold value, and a variable rate punctured code scheme would have to be implemented to equal the efficiency of the Reed-Solomon code.

However, with the low data rates used in the defined system, and modified to form the PRODAT system, convolutional codes suffer from such high overheads, as to make the Reed-Solomon scheme seem more attractive.

## REFERENCES

PROSAT, Phase 1 Report [1986] ESA STR-216. European Space Agency, Paris, France.