INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Appendix III (Rec. G.726)
Appendix II (Rec. G.727)

(05/94)

GENERAL ASPECTS OF DIGITAL TRANSMISSION SYSTEMS

COMPARISON OF ADPCM ALGORITHMS

Appendix III to ITU-T Recommendation G.726
Appendix II to ITU-T Recommendation G.727

(Previously "CCITT Recommendation")

FOREWORD

The ITU-T (Telecommunication Standardization Sector) is a permanent organ of the International Telecommunication Union (ITU). The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1 (Helsinki, March 1-12, 1993).

Appendix III to ITU-T Recommendation G.726 and Appendix II to Recommendation ITU-T G.727 were prepared by ITU-T Study Group 15 (1993-1996) and were approved on the 16 of May 1994.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

CONTENTS

Appendix II to ITU-T Recommendation G.726 and Appendix II to ITU-T Recommendation G.727

COMPARISON OF ADPCM ALGORITHMS

(Geneva, 1994)

1 Background

During the period 1982-1990, the CCITT adopted several adaptive differential pulse code modulation (ADPCM) algorithms. First, the 32 kbit/s (ADPCM) algorithm described in Recommendation G.721 [26; 6] was approved. Later on, Recommendation G.721 was extended with Recommendation G.723 to 40 kbit/s to support voice band data modems at the rate of 9.6 kbit/s, and to 24 kbit/s to allow reduction of the bit rate in cases of network congestion [27]. Prior to the definition of Recommendation G.723, other ADPCM algorithms of performance similar to the 40 kbit/s algorithm had been incorporated in DCME designs and used in telecommunications networks. These algorithms, which may be considered by bilateral agreement, are described in COM XVIII-101 and COM XVIII-102 of the 1984-1988 study period.[1)] Finally, in July 1990, the CCITT combined Recommendations G.721 and G.723 and added operation at 16 kbit/s for overload situations. The combination resulted in a new Recommendation G.726. The CCITT also approved the embedded ADPCM algorithms of Recommendation G.727, which are extensions of the fixed rate ADPCM algorithms defined in Recommendation G.726.

This appendix presents a unified introduction to all these algorithms, their main features and their performance. Clause 2 gives an overview of all ADPCM algorithms that the CCITT has considered. Clause 3 reviews the principles of the algorithms of Recommendations G.726 and G.727 and COM XVIII-102. The principles of the algorithm of COM XVIII-101 are described in clause 4. The remaining clauses outline the main subjective and objective results for the performance of the various algorithms.

2 Overview of ADPCM algorithms

Figures 1 and 2 show a simplified block diagram of a G.726 encoder and decoder, respectively. Figures 3 and 4 show a simplified block diagram of a G.727 encoder and a decoder, respectively. In each set, the coder consists of a logarithmic-to-linear PCM converter, an adaptive quantizer, an inverse adaptive quantizer, and an adaptive predictor.

The PCM converter converts the A-law or μ-law PCM input signal $s(k)$ to a uniform PCM signal (k is the sampling index for a sampling period of 125 μs). The predicted estimate of the input signal $s_e(k)$ is subtracted from the uniform PCM signal, $s_l(k)$, to yield a difference signal $d(k)$:

$$d(k) = s_l(k) - s_e(k) \quad (1)$$

The difference signal is then transformed into a logarithmic presentation with the base 2 and scaled by a scale factor $y(k)$ that is computed as described below.

In Recommendation G.726, the quantizer used is a 31-, 15-, 7- or 4-level non-uniform adaptive quantizer that stops adapting in the presence of a stationary input. This enhances the performance for voiceband data signals. In Recommendation G.727, the adaptive quantizer has 32, 16, 8 or 4 levels. Either quantizer codes the signal $d(k)$ into $I(k)$, a code word of 5, 4, 3 or 2 bits respectively, with one bit always for the sign.

1) The US patent that describes the final algorithm has some differences from the algorithm described in COM XVIII-102 [53].

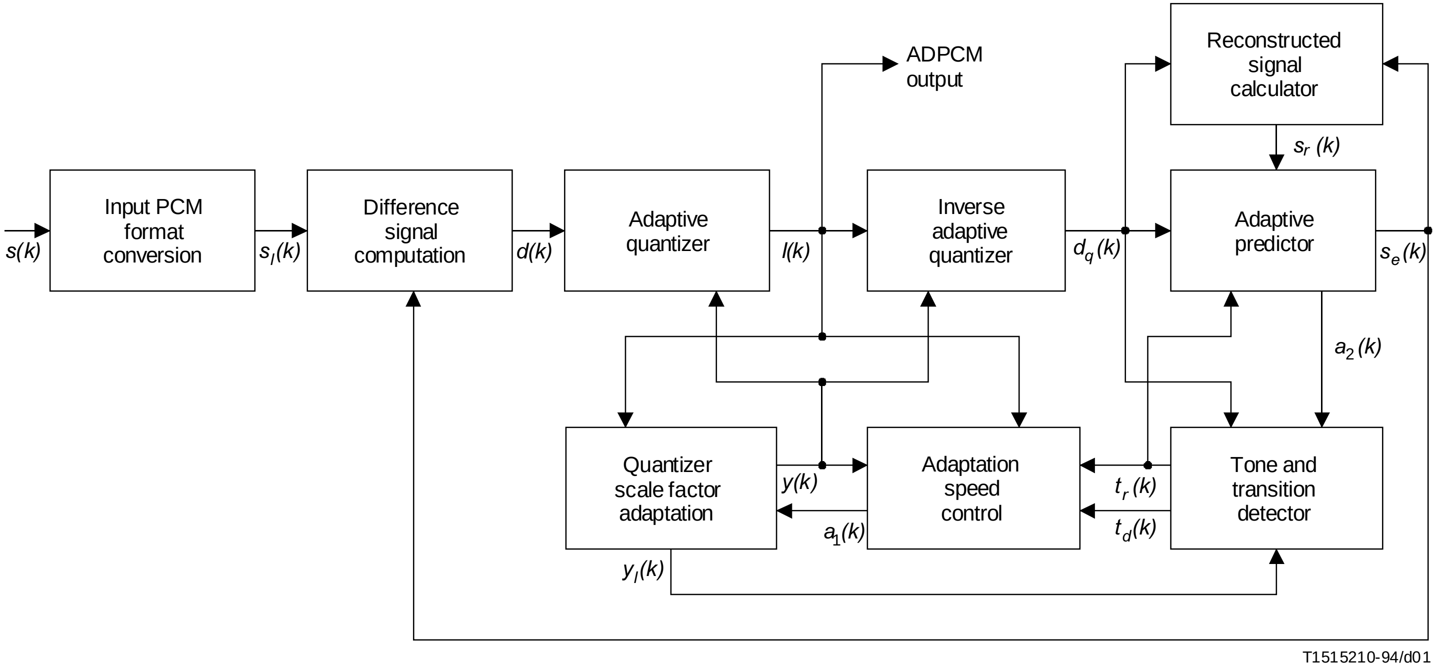


FIGURE 1

G.726 Encoder Block Schematic

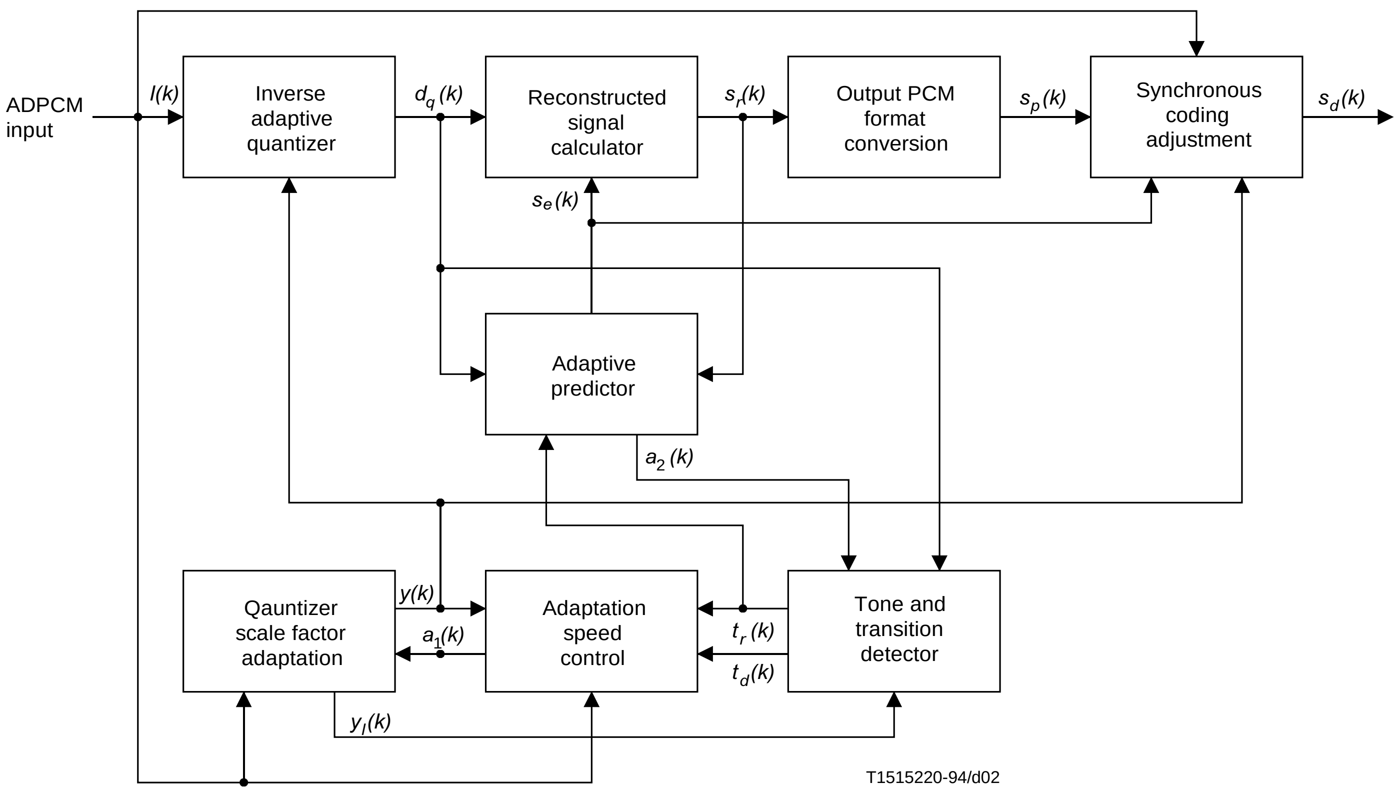


FIGURE 2

G.726 Decoder Block Schematic

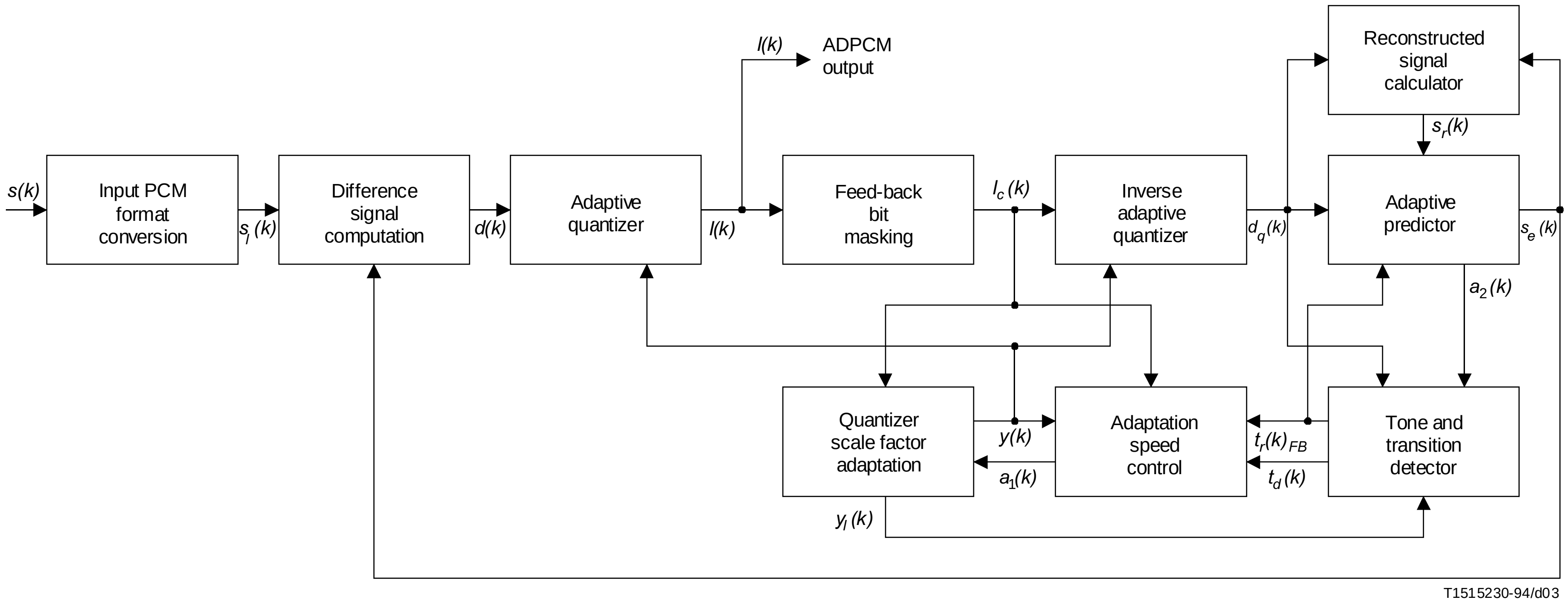


FIGURE 3

G.727 Encoder Block Schematic

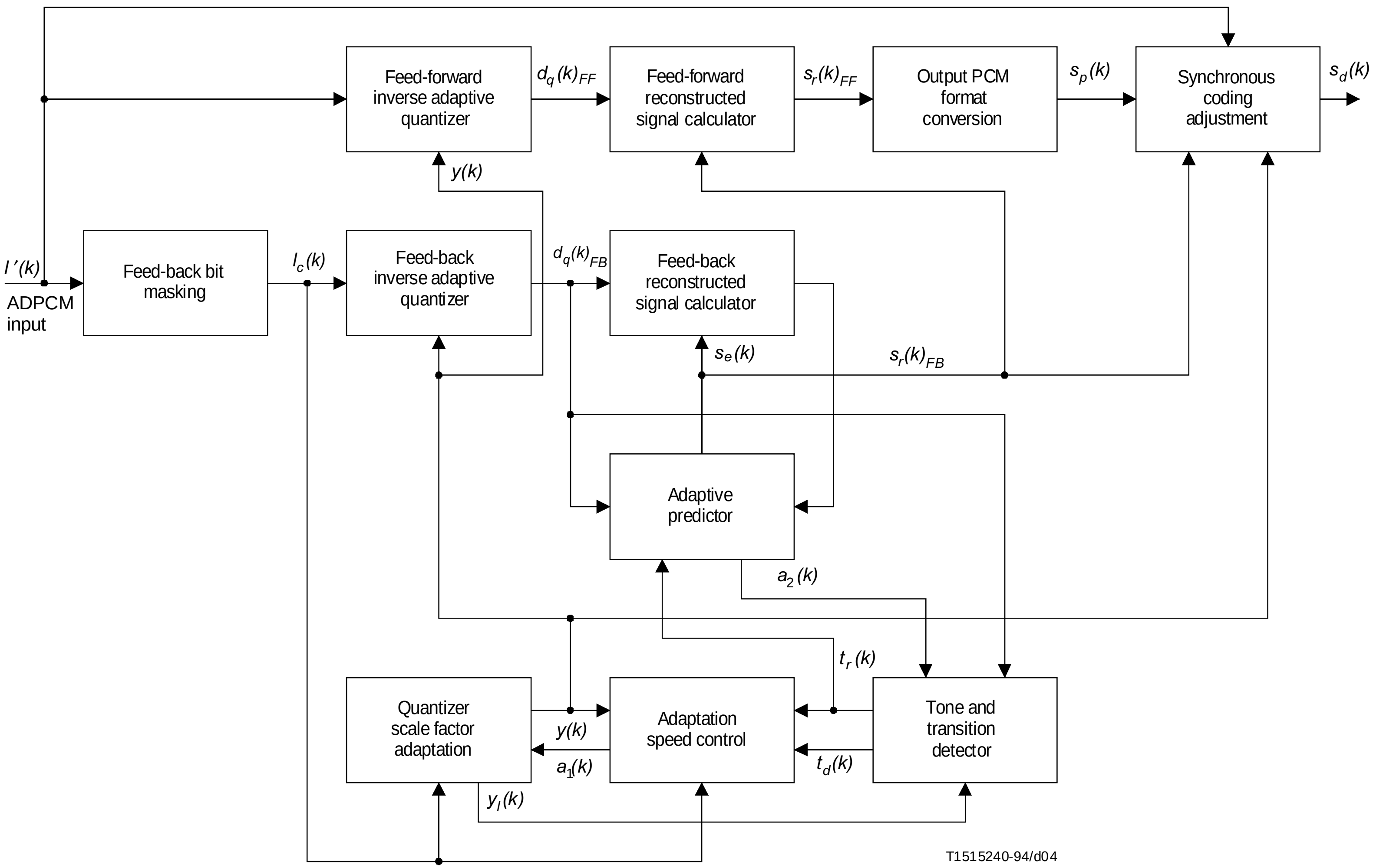


FIGURE 4

G.727 Decoder Block Schematic

The main difference between the fixed ADPCM algorithms of Recommendation G.726, COM XVIII-101 and COM XVIII-102 on the one side, and the embedded algorithms of Recommendation G.727 on the other side is as follows. In the embedded algorithms of Recommendation G.727, the difference between the input and the estimated signal is quantized into code words consisting of enhancement bits and core bits. The core bits are used for prediction, both in the encoder and the decoder, while the enhancement bits are used to reduce the quantization noise in the reconstructed signal. Thus, the core bits must reach the decoder to avoid mistracking, but the enhancement bits can be discarded, if such bit-dropping can alleviate congestion.

The embedded ADPCM algorithms can operate with 5-, 4-, 3- and 2-bits per sample in their feedback path (i.e. at rates of 40, 32, 24 and 16 kbit/s) with conversion to and from 64 kbit/s A-law or μ-law PCM channels. These G.727 algorithms are referenced as **(x,y)** pairs where **x** refers to the feed-forward ADPCM bits and y refers to the feedback ADPCM bits. For example, **(5,2)** represents 40 kbit/s embedded algorithm with two core bits, i.e. with a minimum bit rate is 16 kbit/s.

The impetus for developing Recommendation G.727 was to provide a flexible way to alleviate congestion at any point in a packet network without the need for exchanging control messages between the various nodes in the backward path of the connection (i.e. towards the transmitter). This avoids the “freeze-out” associated with fixed rate ADPCM coding, when transmission capacity is not available and the leading edge of speech bursts are clipped (i.e. the beginnings of words are chopped). This is important when the end-to-end path includes multiple nodes.

The difference between the various fixed rate ADPCM algorithms resides in the way that they accommodate 9.6 kbit/s modem signals. In Recommendation G.726, 40 kbit/s ADPCM is used for voiceband data, while 32 kbit/s ADPCM is used for speech. Accordingly, a bypass arrangement is needed so that upon detection of voiceband data, the appropriate coding is applied without affecting the coding for speech [9].

COM XVIII-101 uses the same 32 kbit/s ADPCM algorithm for speech as well as voiceband data [37]. The predictive structure, which is different from that of all other ADPCM algorithms, is composed of a 10th order adaptive zero predictor, a 4th order adaptive-pole predictor, a 16th-order fixed-pole predictor and an offset predictor. The adaptive pole filter is reserved for highly correlated signals such as speech, while the fixed pole filter is for voiceband data. The relative contribution of each filter is regulated by a set of adaptive gain coefficients. By controlling three different filters within the same structure, the algorithm treats speech signals and voiceband data modem signals up to 9.6 kbit/s with 32 kbit/s ADPCM; it does not require a change-over between speech and voiceband data at 9.6 kbit/s; the price is additional complexity. The adaptive quantizer of this algorithm operates in a 4-bit quantization mode and does not use a tone and transition detector. Figures 5 and 6 give the block diagrams of the encoder and decoder of the COM XVIII-101 algorithm.

COM XVIII-102 uses a special 32 kbit/s ADPCM algorithm that uses 5 bits/sample and is optimized for voiceband data [33; 53]. Following detection of a 2100 Hz tone, the linear PCM bit stream is down-sampled from 8 kHz to 6.4 kHz through a 100-tap symmetric finite impulse response interpolating filter. This interpolating filter introduces a flat delay of 6 ms equally distributed between the encoder and the decoder. To maintain the overall line rate of 32 kbit/s, the ADPCM coding uses 5 bits. To avoid aliasing, the input’s bandwidth must be limited to 3.2 kHz. Also, a realignment from a 6.4 kHz × 5 structure to an 8 kHz × 4 structure is required. The corresponding encoder and decoder block diagrams are shown in Figures 7 and 8, respectively. In these figures, the tone detector block is assumed, because it is not described in the available documents from the algorithm developers [33; 53].

The adaptive predictor relies on the whole codeword $I(k)$ for Recommendation G.726 and the fixed rate ADPCM algorithms, and on the core codeword $I_c(k)$ for Recommendation G.727. The inverse quantizer uses the whole codeword $I_c(k)$ for fixed rate ADPCM (i.e. for Recommendation G.726 and COM XVIII-101, COM XVIII-102), and $I_c(k)$ for embedded ADPCM. Inverse quantization yields the quantized difference signal, $d_q(k)$. The signal estimate from adaptive prediction, $s_e(k)$, is added to this quantized difference signal to yield the reconstructed version, $s_r(k)$, of the input signal.

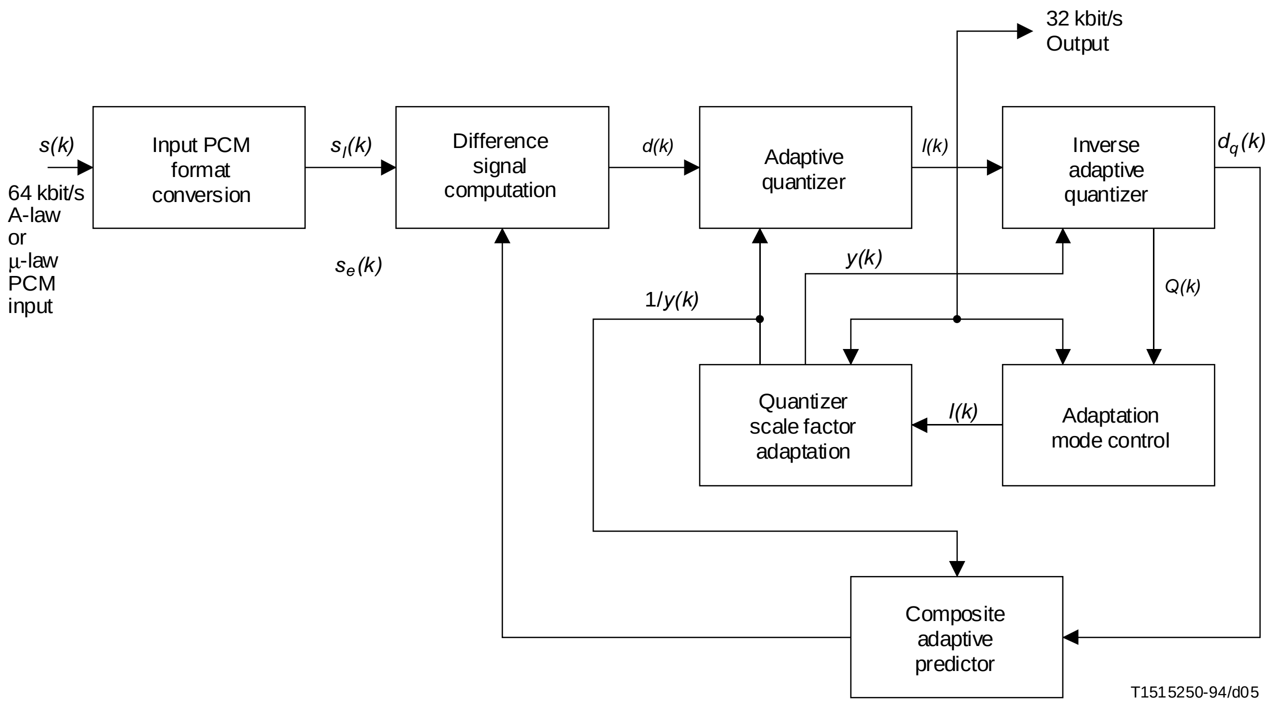


FIGURE 5

COM XVIII-101 Encoder Block Schematic

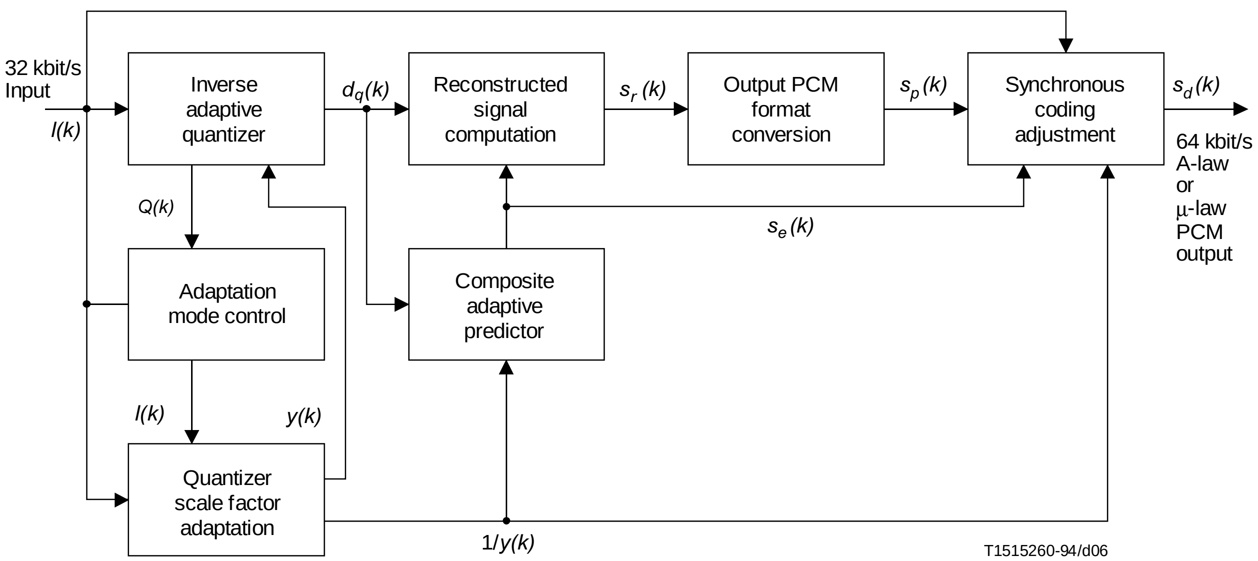
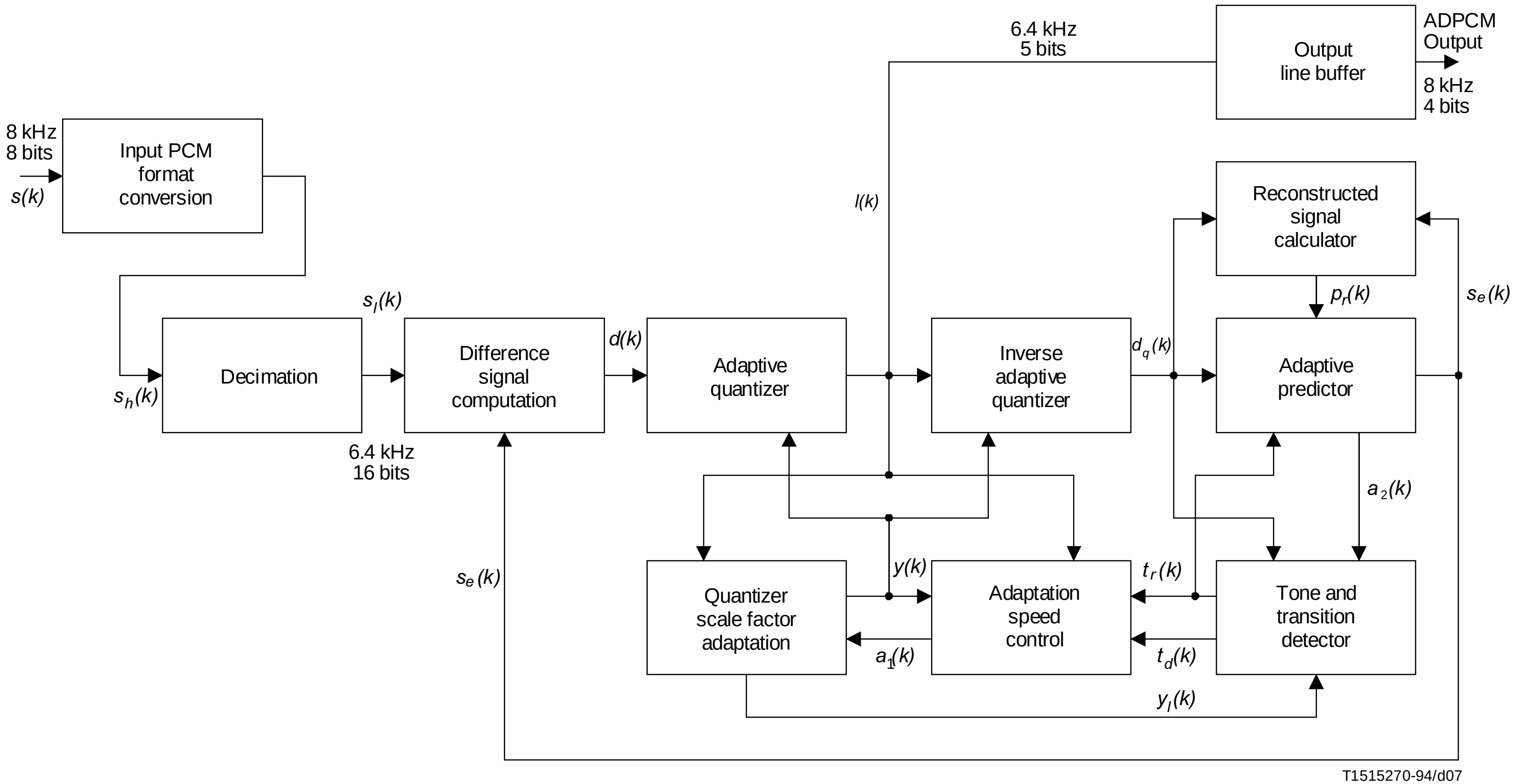


FIGURE 6

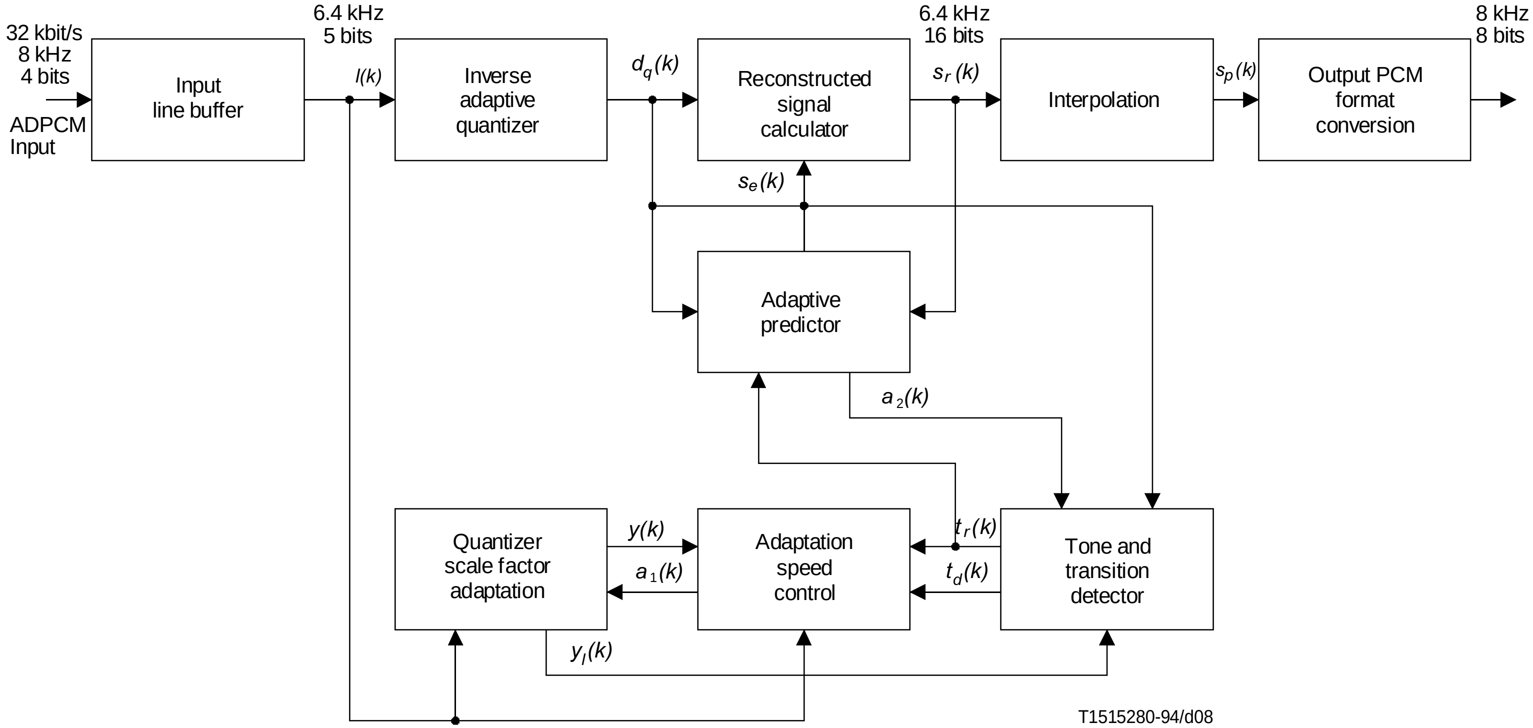
COM XVIII-101 Decoder Block Schematic



NOTE – The tone detector block is assumed; it is neither described in [33] nor in [53].

FIGURE 7

COM XVIII-102 Encoder Block Schematic



NOTE – The tone detector block is assumed; it is neither described in [33] nor in [53].

FIGURE 8

COM XVIII-102 Decoder Block Schematic

The adaptation of both the quantizer and inverse quantizer depends on all the bits of $I(k)$ in the fixed rate algorithms, and on the core bits, $I_c(k)$ in Recommendation G.727. The quantizer output $I(k)$ takes all the corresponding non-zero values, while the adaptation operates on all possible values since the associated all-zeros codeword may result from transmission errors.

The decoder at both the receiver and the transmitter includes a structure identical to the feedback portion of the encoder. In addition, the receiver for both Recommendations G.726 and G.727 transcodes the bit stream from uniform PCM back to A-law or μ-law PCM for synchronous coding adjustment. This operation is based on all the received bits and prevents cumulative distortion on synchronous tandem codings (ADPCM-PCM-ADPCM) under certain conditions. This is achieved by selecting the PCM output codes that eliminate the quantizing distortion in the next ADPCM encoding stage [38].

The algorithm of COM XVIII-102 does not have a synchronous coding adjustment block as in Recommendations G.726, G.727 and COM XVIII-101; therefore, it does not exhibit the synchronous tandem property described in 5.2.

Clause 5 recapitulates the encoder principles and explains the differences among the various algorithms of Recommendations G.726 and G.727 and COM XVIII-102. Discussion of the algorithm of COM XVIII-101, whose structure is different from the other algorithms, is the subject of clause 6.

3 Principles of Recommendations G.726 and G.727 and COM XVIII-102

3.1 Adaptive Prediction and Reconstruction of the Signal

The primary function of the adaptive predictor is to compute the signal estimate $s_e(k)$ from the quantized difference signal $d_q(k)$ and past values of the reconstructed signal $s_e(k)$. The predictor has the form of an autoregressive moving average (ARMA) filter whose frequency spectrum fits a wide range of input voice-band signals.

The signal estimate is computed from the reconstructed signal $s_r(k)$ *as:*

$$s_e(k) = \sum_{i=1}^{2} a_i^{k-1}\, s_r(k-i) + \sum_{i=1}^{6} b_i^{k-1}\, d_q(k-i) \tag{2}$$

with

$$s_r(k) = s_e(k) + d_q(k) \tag{3}$$

and

$$d_q(k) = d(k) + e(k) = Q\left[\frac{d(k)}{y(k)}\right] y(k) \tag{4}$$

where

a_1^k, a_2^k	are the autoregressive coefficients for sample k;
$b_i^k, i = 1, ..., 6$	the moving average coefficients for sample k;
$d(k)$	is the *difference signal at sample* k;
$d_q(k)$	is the quantized difference signal at sample k;
$e(k)$	is the quantization error at sample k;
$Q[x]$	is the normalized quantizer output for the input x;
	$= \lvert x \rvert - y(k)$ in the logarithmic domain;
$y(k)$	is the the scale factor error at sample k.

The starting values are: $d(0) = s_o(0) = s_r(0) = 0$ and $d_q(k) = 0$ for $k < 0$. Equations (2)-(4) are similar for both Recommendations G.726 and G.727 with the exception that the quantization error $e(k)$ in Recommendation G.727 includes the effects of using $I_c(k)$ instead of $I(k)$. As explained above, $d_q(k)$ was obtained from $I_c(k)$ through inverse quantization.

With $B =$ delay operator such that $Bd_q(k) \triangleq d_q(k-1)$, we can combine (2) and (3):

$$s_e(k) = \frac{\left(a_1^{k-1} + b_1^{k-1}\right) B_+ \left(a_2^{k-1} + b_2^{k-1}\right) B_+^2 \dots + b_6 B^6}{1 - a_1^{k-1} B - a_2^{k-1} B^2} \cdot d_q(k) \tag{5}$$

From (1), (3) and (5) we can write:

$$\begin{aligned}\left(1 + \sum_{i=1}^{6} b_i^{k-1} B^i\right) d(k) &= \left(1 - a_1^{k-1} B - a_2^{k-1} B^2\right) s_l(k) \\ &+ \left[\left(a_1^{k-1} + b_1^{k-1}\right) B_+ \left(a_2^{k-1} + b_2^{k-1}\right) B_+^2 \dots + b_6 B^6\right] e(k)\end{aligned} \tag{6}$$

The update equations for the autoregressive and moving average coefficients of the ARMA (2,6) predictor follow the same simplified gradient algorithm used in Recommendation G.721 [39; 45].

$$a_1^k = (1 - 2^{-8})\, a_1^{k-1} + \left(3 \cdot 2^{-8}\right) \operatorname{sgn}[p(k)] \operatorname{sgn}[p(k-1)] \tag{7}$$

$$\begin{aligned} a_2^k = (1 - 2^{-7})\, a_2^{k-1} + 2^{-7} \Big(&\operatorname{sgn}[p(k)] \operatorname{sgn}[p(k-2)] \\ &- f\left(a_1^{k-1}\right) \operatorname{sgn}[p(k)] \operatorname{sgn}[p(k-1)] \Big)\end{aligned} \tag{8}$$

where $p(k)$ is an MA (6) process defined as:

$$\begin{aligned} p(k) &= d_q(k) + s_{ez}(k) = \\ &= \left(1 + b_1^{k-1} B_+ \dots + b_6^{k-1} B^6\right) d_q(k)\end{aligned} \tag{9}$$

and

$$f(a_1) = \begin{cases} 4a_1, & |a_1| \le 2^{-1} \\ 2 \operatorname{sgn}(a_1), & |a_1| > 2^{-1} \end{cases}$$

From equation (5), the frequency response of the predictor is given by:

$$H(w) = \frac{\left(a_1^{k-1} + b_1^{k-1}\right) e^{-j\omega T} + \left(a_2^{k-1} + b_2^{k-1}\right) e^{-j2\omega T} + \dots + b_6^{k-1} e^{-j6\omega T}}{1 - a_1^{k-1} e^{-j\omega T} - a_2^{k-1} e^{-j2\omega T}} \tag{10}$$

$$j = \sqrt{-1},\ 0 \le \omega \le \frac{\pi}{T} \quad \text{and} \quad T = 125\ \mu\text{s}$$

The frequency responses for the 40 kbit/s algorithms of Recommendation G.726, and the algorithms of COM XVIII-101 and COM XVIII-102 are available in the literature [23; 24].

For stationarity, the autoregressive (AR) parameters must fall in the triangular region defined by [11]:

$$|a_1^k| + a_2^k < 1$$

$$|a_2^k| < 1$$

This condition is satisfied with the stability constraints:

and
$$|a_1|^k \le 1 - 2^{-4} - a_2^k$$
$$|a_2|^k \le 0.75$$

3.1.1 Recommendations G.726 and G.727

For Recommendations G.726 and G.727, the update equations for the moving average (MA) coefficients b_i are [39; 45]:

$$b_i^k = (1 - 2^{-8})\, b_i^{k-1} + 2^{-7} \operatorname{sgn}[d_q(k)] \operatorname{sgn}[d_q(k - i)] \tag{11}$$

for $i = 1,2, \ldots, 6$.

Note that $|\, b_i(k)\, | \le 2$. As will be seen later, when a transition is detected the predictor coefficients are reset, i.e. $a_1^k = a_2^k = 0$, and $b_i^k = 0$ for $i = 1, 2, ..., 6$.

As above, sgn [0] = 1, and sgn $[d_q(k)] = 0$, for $k < 0$

For 40 kbit/s coding, the adaptive predictor is changed to decrease the leak factor used for zeroes coefficient operation. In this case, equation (11) becomes:

$$b_i(k) = (1 - 2^{-9})\, b_i(k - 1) + 2^{-7} \operatorname{sgn}[d_q(k)] \operatorname{sgn}[d_q(k - i)] \tag{11´}$$

Note that $b_i(k)$ is implicitly limited to ± 2.

3.1.2 COM XVIII-102

In the 32 kbit/s algorithm of COM XVIII-102 [33], the coefficient update equations are of the same form but the leak factor is lower for both the poles and the zeros. The corresponding equations are:

$$a_1^k = (1 - 2^{-11})\, a_1^{k-1} + \left(2^{-9}\right) \operatorname{sgn}[p(k)] \operatorname{sgn}[p(k - 1)] \tag{7´}$$

$$\begin{aligned} a_2^k = {} & (1 - 2^{-10})\, a_2^{k-1} + 2^{-9} \big(\operatorname{sgn}[p(k)] \operatorname{sgn}[p(k - 2)] \\ & - f\left(a_1^{k-1}\right) \operatorname{sgn}[p(k)] \operatorname{sgn}[p(k - 1)]\big) \end{aligned} \tag{8´}$$

and

$$b_i^k = (1 - 2^{-10})\, b_i^{k-1} + 2^{-9} \operatorname{sgn}[d_q(k)] \operatorname{sgn}[d_q(k - i)] \tag{9´}$$

3.2 Adaptive Quantizer

All ADPCM algorithms have non-uniform midrise adaptive quantizers that are based on the minimum-mean-squared error Lloyd-Max quantizer at 32 kbit/s [35; pp. 131-134]. They operate in a bimodal fashion (slow and fast) with an adaptive scale factor, $y(k)$, to accommodate both speech and voice-band data signals [47]. In Recommendation G.726, the quantizer is a 31-, 15-, 7- or 4-level non-uniform adaptive quantizer for operation at 40, 32, 24 or 16 kbit/s, respectively. Each rate has it own separate quantizer. In Recommendation G.727, a 32-, 16-, 8- or 4-level non-uniform adaptive quantizer used to quantize the difference signal, $d(k)$. The various quantizer tables are embedded within each other so that the decision levels are forcibly aligned to ensure that the decision levels for the 32, 24 and 16 kbit/s quantizers are subsets of those for the 40 kbit/s quantizer. This contrasts with the algorithms of Recommendation G.726 where the decision levels are not aligned which makes them unsuitable for embedded applications, but with a slightly improved signal to quantization noise ratio as shown 8.3.1.

Tables 1, 2 and 3 give the input/output normalized characteristics of the G.726 quantizer for operation at 40, 32 and 24 kbit/s. Table 4 corresponds to the algorithm of COM XVIII-102 [33].

TABLE 1

Quantizer Normalized Input/Output Characteristic for 40 kbit/s Operation of Recommendation G.726

Normalized Quantizer Input Range $\log_2 \mid d(k) \mid - y(k)$	$\mid I(k) \mid$	Normalized Quantizer Output $\log_2 \mid d_q(k) \mid - y(k)$
[4.31, + ∞)	15	4.42
[4.12, 4.31)	14	4.21
[3.91, 4.12)	13	4.02
[3.70, 3.91)	12	3.81
[3.47, 3.70)	11	3.59
[3.22, 3.47)	10	3.35
[2.95, 3.22)	9	3.09
[2.64, 2.95)	8	2.80
[2.32, 2.64)	7	2.48
[1.95, 2.32)	6	2.14
[1.54, 1.95)	5	1.75
[1.08, 1.54)	4	1.32
[0.52, 1.08)	3	0.81
[–0,13, 0.52)	2	0.22
[–0,96, –0.13)	1	–0.52
(– ∞, –0.96)	0	– ∞

Tables 5 and 6 give respectively the normalized input and output characteristic (infinite precision values) of the quantizer for the G.727 algorithms with 5, 4, 3 and 2 core bits. Both tables show the alignment of the decision levels for the various bit rates. In all these tables, the most significant bit is the sign bit and the remaining bits represent the magnitude. The 5-, 4-, 3- or 2-bit quantizer output, $I(k)$, of Table 6 forms the 40, 32, 24 or 16 kbit/s output signal that comprises both the enhancement and core bits. The 16 kbit/s algorithm is the same as the (2,2) algorithm of Recommendation G.727. It can be seen from comparing Tables 1 and 4 that the normalized quantizer output is less for the algorithm of COM XVIII-102 than for the 40 kbit/s algorithm of Recommendation G.726 for the same codeword.

Bit masking is another difference between the fixed rate ADPCM algorithms of Recommendation G.726, COM XVIII-101, COM XVIII-102 and the embedded ADPCM algorithms of Recommendation G.727. Through this process, the enhancement bits are discarded by a logical right shift of $I(k)$ by a number equal to the number of enhancement bits. The core bits, $I_c(k)$, are then used by the inverse adaptive quantizer in the feedback path, the quantizer scale factor adaptation, and the adaptation speed control blocks. The inverse adaptive quantizer generates $d_q(k)$ (the quantized difference signal) by adding $y(k)$ to the value that corresponds to $I_c(k)$ in the normalized quantizing characteristic of the respective table.

TABLE 2

Quantizer Normalized Input/Output Characteristic for 32 kbit/s Operation of Recommendation G.726

Normalized Quantizer Input Range $\log_2 \mid d(k) \mid - y(k)$	$\mid I(k) \mid$	Normalized Quantizer Output $\log_2 \mid d_q(k) \mid - y(k)$
[3.12, + ∞)	7	3.32
[2.72, 3.12)	6	2.91
[2.34, 2.72)	5	2.52
[1.91, 2.34)	4	2.13
[1.38, 1.91)	3	1.66
[0.62, 1.38)	2	1.05
[–0.98, 0.62)	1	0.031
(– ∞, –0.98)	0	– ∞

TABLE 3

Quantizer Normalized Input/Output Characteristic for 24 kbit/s Operation of Recommendation G.726

Normalized Quantizer Input Range $\log_2 \mid d(k) \mid - y(k)$	$\mid I(k) \mid$	Normalized Quantizer Output $\log_2 \mid d_q(k) \mid - y(k)$
[2.58, + ∞)	3	2.91
[1.70, 2.58)	2	2.13
[0.06, 1.70)	1	1.05
(– ∞ , 0.06)	0	– ∞

TABLE 4

Quantizer Normalized Input/Output Characteristic for COM XVIII-102

Normalized Quantizer Input Range $\log_2 \mid d(k) \mid - y(k)$	$\mid I(k) \mid$	Normalized Quantizer Output $\log_2 \mid d_q(k) \mid - y(k)$
[3.45, + ∞)	15	3.59
[3.20, 3.45)	14	3.31
[2.99, 3.20)	13	3.09
[2.80, 2.99)	12	2.89
[2.61, 2.80)	11	2.71
[2.42, 2.61)	10	2.52
[2.23, 2.42)	9	2.33
[2.02, 2.23)	8	2.12
[1.79, 2.02)	7	1.90
[1.52, 1.79)	6	1.66
[1.22, 1.52)	5	1.38
[0.84, 1.22)	4	1.04
[0.35, 0.84)	3	0.61
[–0.40, 0.35)	2	0.00
[–1.98, –0.40)	1	–0.98
(– ∞, –1.98)	0	– ∞

TABLE 5

Quantizer Normalized Input Characteristics
for Embedded Operation (Recommendation G.727)

Normalized Quantizer Input Magnitude Range $\log_2 \lvert d(k) \rvert - y(k)$	$\lvert I(k) \rvert$ or $\lvert I_c(k) \rvert$			
	Number of Bits			
	5			
	4			
	3			
	2			
(– ∞, –1.05)	0	0	0	0
[–1.05, –0.05)	0	0	0	1
[–0.05, 0.54)	0	0	1	0
[0.54, 0.96)	0	0	1	1
[0.96, 1.30)	0	1	0	0
[1.30, 1.58)	0	1	0	1
[1.58, 1.82)	0	1	1	0
[1.82, 2.04)	0	1	1	1
[2.04, 2.23)	1	0	0	0
[2.23, 2.42)	1	0	0	1
[2.42, 2.60)	1	0	1	0
[2.60, 2.78)	1	0	1	1
[2.78, 2.97)	1	1	0	0
[2.97, 3.16)	1	1	0	1
[3.16, 3.43)	1	1	1	0
[3.43, ∞)	1	1	1	1

TABLE 6

Quantizer Normalized Output Characteristic
for Embedded Operation (Recommendation G.727)

$\lvert I(k) \rvert$ or $\lvert I_c(k) \rvert$				Normalized Quantizer Output Magnitude $\log_2 \lvert d_q(k) \rvert - y(k)$			
Number of bits				Number of bits			
5				5			
4				4			
3				3			
2				2			
0	0	0	0	0.91	–0.09	–1.06	–2.06
0	0	0	1	"	"	"	–0.48
0	0	1	0	"	"	0.53	0.27
0	0	1	1	"	"	"	0.76
0	1	0	0	"	1.55	1.29	1.13
0	1	0	1	"	"	"	1.44
0	1	1	0	"	"	1.81	1.70
0	1	1	1	"	"	"	1.92
1	0	0	0	2.85	2.4	2.23	2.13
1	0	0	1	"	"	"	2.33
1	0	1	0	"	"	2.59	2.51
1	0	1	1	"	"	"	2.69
1	1	0	0	"	3.09	2.95	2.87
1	1	0	1	"	"	"	3.05
1	1	1	0	"	"	3.34	3.27
1	1	1	1	"	"	"	3.56

Note that the (5,5) algorithm has not been standardized because no embedded APCM algorithms with more than 5 bits in the forward loop have been considered.

3.3 Quantizer Scale Factor Adaptation

To enhance the robustness of the adaptation, the quantizer and the inverse quantizer rely on the composite scaling factor, $y(k)$, that conibines a slow (locked) scale factor and a fast (unlocked) scale factor $y_f(k)$. $y_l(k)$ is used for transition detection while $y(k)$ is used for speed adaptation [47].

3.3.1 Adaptation for Recommendations G.726 and G.727

Fast adaptation is for signals (such as speech) whose first difference time series fluctuates rapidly. Slow adaptation is for signals (such as voiceband data or tones) whose first difference exhibits small deviations. The effective speed of adaptation combines both scale factors.

The fast (unlocked) scale factor, $y_u(k)$, is recursively computed in the base 2 logarithmic domain from the resultant logarithmic scale factor, $y(k)$:

$$\begin{aligned} y_u(k) &= (1 - 2^{-5})\, y(k) + 2^{-5}\ W[I(k)], \quad k = 1, 2, \ldots \\ y(0) &= y_z(0) = 1.06 \end{aligned} \tag{12}$$

where $y_u(k)$ is limited by $1.06 \leq y_u(k) \leq 10.00$ and $W[I(k)]$ is a discrete function that defines a quantizer multiplier whose value depends on the code word I for Recommendation G.726 and on the core bits I_c, for Recommendation G.727 as shown in Appendix I. The leakage factor $(1 - 2^{-5})$ gives the adaptive process a finite memory to allow recovery after transmission errors.

For both Recommendations G.726 and G.727, the slow (locked) scale factor $y_l(k)$ is derived from $y_u(k)$ with a low pass ARMA (1,1) filter:

with

$$\begin{aligned} y_l(k) &= (1 - 2^{-6})\, y_l(k - 1) + 2^{-6}\, y_z(k), \quad k = 1, 2, \ldots \\ y_u(0) &= 1.06, \quad \text{and} \quad y_l(0) = 1.06 \end{aligned} \tag{13}$$

Equation (13) can be written for Recommendation G.726 as:

$$y_l(k) = \frac{2^{-6}(1 - 2^{-5})\ y(k) + 2^{-11}\ W[I(k)]}{1 - (1 - 2^{-6})\, B} \tag{13.a}$$

and for Recommendation G.727 as:

$$y_l(k) = \frac{2^{-6}(1 - 2^{-5})\ y(k) + 2^{-11}\ W[I_c(k)]}{1 - (1 - 2^{-6})\, B} \tag{13.b}$$

The pole of this filter is positive and close to the unit circle; therefore its autocorrelation function slowly decays exponentially towards zero. This means that the low frequencies dominate the frequency spectrum.

The steps used in the adaptation of the scaling factor $y(k)$ are the same ones in both Recommendations G.726 and G.727 except that, in Recommendation G.727, the discrete function $W[I_c(k)]$ is defined for 2, 3 and 4 core bits instead of being defined for the whole code word as in the function $W[I(k)]$ in Recommendation G.726. The values of both functions for 2-bit core operation are the same for Recommendation G.726 at 16 kbit/s [19]. The function $W[I(k)]$ is defined in Appendix I.

In the following development, we will use exclusively the variable $I_c(k)$ for all algorithms, noting that the core code word is the whole code word for non-embedded algorithms.

The fast and slow scale factors are then combined to form the resultant scale factor:

$$y(k) = a_l(k)\, y_u(k-1) + [1 - a_l(k)]\, y_l(k-1)$$

Therefore

$$\begin{aligned} y(k) = {} & \{a_l(k)(1 - 2^{-5}) + [1 - a_l(k)]2^{-6}(1 - 2^{-5})\}\, y(k-1) \\ & + \{[1 - a_l(k)](1 - 2^{-6})\}\, y_l(k-1) \\ & + \{2^{-5} a_l(k) + [1 - a_l(k)]2^{-11}\}\, W[I_c(k-1)], \qquad k = 1, 2, \ldots \end{aligned} \tag{14.a}$$

where

$$0 \le a_l(k) \le 1 \quad \text{and} \quad a_l(0) = 0$$

or

$$y(k) = \frac{[1 - a_l(k)][1 - 2^{-6}]\, B\, y_l(k) + \left\{2^{-5} a_l(k) + [1 - a_l(k)]2^{-11}\right\} W[I_c(k-1)]}{1 - \left[2^{-6} + (1 - 2^{-6})\, a_l(k)\right](1 - 2^{-5})\, B} \tag{14.b}$$
$$k = 1, 2, \cdots$$

When $a_l(k) = 1$, the quantizer is unlocked and $y(k)$ is the same as $y_u(k)$, i.e.

$$y(k) = \frac{2^{-5}\ W[I_c(k-1)]}{1 - (1 - 2^{-5})\, B}, \quad k = 1, 2, \ldots \tag{15.a}$$

In this case, $y(k)$ does not include a long term average and is suitable for rapidly varying signals such as speech. When $a_l(k) = 0$, the quantizer is locked and $y(k)$ is given by:

$$\begin{aligned} y(k) = {} & 2^{-6}(1 - 2^{-5})\, y(k-1) + (1 - 2^{-6})\, y_l(k-1) \\ & + 2^{-11}\ W[I_c(k-1)] \end{aligned} \tag{15.b}$$

$$y(k) = \frac{\left(1 - 2^{-6}\right) B y_l(k) + 2^{-11}\ W[I_c(k-1)]}{1 - 2^{-6}(1 - 2^{-5})\, B}, \quad k = 1, 2, \ldots \tag{15.c}$$

Here, $y(k)$ depends on the long term average and is less dependent on its past values (the autoregressive coefficient is 64 times smaller) and the quantizer adapts more slowly. This case applies to stationary signals, such as voiceband data and tones, where the quantization must be small to reduce the quantization error.

The controlling parameter, $a_l(k)$, is derived on the basis of the rate of change of the quantized first difference of the input time series, such that its values fall in the range [0,1]. This is shown in the following subclause.

3.3.2 Adaptation of COM XVIII-102

The adaptation of the scale factor in the algorithm of COM XVIII-102 is given by the following equation:

$$y(k) = (1 - a(k))\, y(k-1) + a(k)\, W[I(k)], \ k = 1, 2, \ldots \tag{16}$$

The controlling factor $a(k)$ can assume values in the range 2^{-9}, 2^{-6}. The lower value is for the low speed mode (voiceband data signals and tones), while the higher value is for silence periods. Initially, the value is 2^{-6}. If three consecutive samples are observed such that $y(k) > 3.5$, $a(k)$ is allowed to decay according to the following AR model:

$$a(k) = (1 - 2^{-3})\, a(k - 1) \quad k = 1, 2, \ldots$$

As soon as an energy drop is observed, $a(k)$ is changed back to 2^{-6} to operate in the high speed mode.

The values of $W[I]$ for COM XVIII-102 are also defined in Appendix I. It should be noted that these values are smaller than the values used for Recommendation G.726.

3.4 Adaptation Speed Control

There are two measures for the average magnitude of $I(k)$ *or* $I_c(k)$: 1) a short term measure $d_{ms}(k)$ and, 2) a long term measure $d_{ml}(k)$. The calculation of both measures is exactly the same as in both Recommendations G.726 and G.727:

$$d_{ms}(k) = (1 - 2^{-5})\, d_{ms}(k - 1) + 2^{-5} F[I(k - 1)] \tag{16.a}$$

$$d_{ms}(k) = (1 - 2^{-5})\, d_{ms}(k - 1) + 2^{-5} F[I_c(k - 1)] \tag{16.b}$$

and

$$d_{ml}(k) = (1 - 2^{-7})\, d_{ml}(k - 1) + 2^{-7} F[I(k - 1)] \tag{17.a}$$

$$d_{ml}(k) = (1 - 2^{-7})\, d_{ml}(k - 1) + 2^{-7} F[I_c(k - 1)] \tag{17.b}$$

with $d_{ms}(0) = d_{ml}(0) = 0$, and $F[I(k - 1)]$ and $F[I_c(k - 1)]$ are the corresponding step functions. $F[I_c(k - 1)]$ is defined for 2, 3 and 4 core bits instead of being defined for the whole code word as for $F[I(k - 1)]$. The values of both functions for 2-bit core operation are the same for Recommendations G.726 at 16 kbit/s [19]. These functions are defined in Appendix I.

If $\dfrac{|\, d_{ms}(k) - d_{ml}(k)\, |}{d_{ml}(k)} \geq 2^{-3}$, the quantizer is unlocked by having:

$$a_p(k) = (1 - 2^{-4})\, a_p(k - 1) + 2^{-3} \quad \text{for } k > 0 \quad \text{and} \quad a_p(0) = 0$$

This allows the quantizer to track the changes in $I(k)$ in Recommendation G.726 or $I_c(k)$ in Recommendation G.727.

When the characteristics of the input signal change abruptly while the predictor gain is high, i.e. its prediction error is low, problems could occur, as in the case of the 1984 version of the algorithm. When the predictor adapts to extended tones (such as in the case of FSK modems operating in the character or asynchronous mode), it will track the input signal closely. Because the residual error is small, the quantizer will be in the locked mode. If the frequency of the tone changes abruptly, the predictor and the quantizer will take some time to adapt to the new signal, because, as seen from equation (15.a), the quantizer scale factor will depend for some time on its past values, so that it will also require a period to unlock and increase its scale factor [7; 8]. This slow adaptation causes some instabilities for the Bell 202 modems and the CCITT V.23 series of modems, where the mark and space tones are wide apart, when they operate in the character mode.

Modifications were introduced to force the quantizer in the fast adaptation mode, and to reset the predictor coefficient, when an FSK signal is present. In addition, 15-level quantization is adopted instead of the original 16-level quantization to allow the use in US Networks that do not provide bit sequence independence. Thus, the predictor coefficients are reset to ‘0’ and the quantizer is unlocked, i.e. ($a_p = 1$), following a transition between two tones [7; 34]. A tone is detected if $a_2(k) < -0.71775$ so that if the binary variable $t_d(k)$ signals the presence of a tone:

$$t_d(k) = \begin{cases} 1, & a_2(k) < -0.71875 \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

In the presence of a tone (or a partial band signal), a_p is set to 1 to unlock the quantizer and the residual errors are observed. If the residual error $d_q(k)$ exceeds a certain threshold, then it is assumed that a tone transition has occurred. The predictor coefficients are then reset to zero while the quantizer is in the fast mode of adaptation. The binary variable $t_r(k)$ indicates the presence of a tone, then we have:

$$t_r(k) = \begin{cases} 1, & a_2(k) < -0.71875 \text{ and } |d_q(k)| > 24 \cdot 2^{yl(k)} \\ 0, & \text{otherwise.} \end{cases} \tag{19}$$

To summarize, the intermediate variable $a_p(k)$ is defined as follows:

$$a_p(k) = \begin{cases} (1 - 2^{-4})\, a_p(k-1) + 2^{-3}, & \text{if } |d_{ms}(k) - d_{ml}(k)| \geq 2^{-3}\, d_{ml}(k) \\ & \text{or } y(k) < 3 \\ & \text{or } t_d(k) = 1 \\ 1 \qquad , & t_{r(k)} = 1 \\ (1 - 2^{-4})\, a_p(k-1) \qquad , & \text{otherwise,} \end{cases} \tag{20}$$

and

$$a_p(k) = a_p \text{ for } k \leq 0$$

Thus, $a_p(k) \rightarrow 2$ whenever one of the following conditions is true:

1) $\dfrac{|d_{ms}(k) - d_{ml}(k)|}{d_{ml}(k)}$ is large because the average magnitude of $I_c(k)$ is changing.

2) There is an idle signal (as indicated by $y(k) < 3$).

3) A tone is detected (as indicated by $t_d(k) = 1$).

When $a_p(k)$ increases, the quantizer unlocks so that the scale factor changes rapidly and the quantizer step size can adapt to the input signal.

In contrast, $a_p(k) \rightarrow 0$ if the difference is small (average magnitude of $I_c(k)$ is relatively constant).

Similar to the case of Recommendations G.721 and G.723, $a_p(k-1)$ is then limited to yield $a_1(k)$ used in equation (10) above:

$$a_l(k) = \min\{a_p(k-1), 1\} \tag{21}$$

so that

$$a_l(0) = 0 \text{ since } a_p(-1) = 0.$$

The asymmetrical limiting reduces premature transitions for pulsed input signals such as switched carrier voiceband data, where the modem switches the carrier on to transmit the message and switches it off after the message is sent. This is because it postpones a fast to slow state transition until the absolute value of $I_c(k)$ has remained constant for some time. Note that $a_p(k)$ is set to a clipping level of 1 after the detection of a partial band signal transition [indicated by $t_r(k) = 1$].

4 Principles of COM XVIII-101

The description of this algorithm is based on [15; 31; 36] and on personal correspondence with Mr. Atsushi Shimbo from OKI Electric.

4.1 Prediction

The transfer function $H(B)$ of the composite adaptive predictor of COM XVIII-101 is given by:

$$H(B) = \frac{1 + \Theta(B)}{(1 - g_{ap}\Phi_{ap}(B))(1 - g_{fp}\Phi_{fp}(B))} \tag{22}$$

where $\Theta(B)$ is the transfer function of the 10th-order adaptive moving average, $\Phi_{ap}(B)$ is the transfer function of the 4th-order adaptive autoregressive predictor, $\Phi_{fp}(B)$ is the transfer function of the 16th-order fixed autoregressive predictor, and g_{ap} and g_{fp} are the adaptive gains for the adaptive and fixed predictors.

Comparing equation (22) with equation (10), we see that the predictor of COM XVIII-101 is composed of the following components:

1) a 10th degree adaptive moving average (MA) predictor;

2) a 4th degree adaptive autoregressive (AR) predictor;

3) a 16th degree fixed autoregressive (AR) predictor; and

4) an offset.

As explained earlier, the adaptive AR(4) filter is for speech signals while the fixed AR(16) filter is for voice band signals. The order of the fixed predictor is so chosen because the prediction gain tends to saturate above the 16th order [31]. The adaptive gain g_{ap} is chosen such that it tends to 1 for speech signals and to 0 for voiceband data. Inversely, the fixed gain g_{fp} is chosen such that it tends to 0 for speech signals and to 1 for voiceband data.

Thus, the signal estimate s_e, is computed from four components $e_i(k)$ as follows:

$$s_e(k) = \sum_{i=1}^{4} e_i(k) \tag{23}$$

In this equation, $e_1(k)$ is the output of the MA(10) predictor whose coefficients for the kth sample are b_i^k, $i = (1, ..., 10)$, $e_2(k)$ is the output of the adaptive AR(4) predictor with the coefficients at k φ_i^k, $i = (1, ..., 4)$, $e_3(k)$ is the output of the fixed AR(10) filter whose fixed coefficients are denoted as c_i, $i = (1, ..., 16)$, and the offset $e_4(k)$ is a filtered version of the quantized difference $d_q(k)$.

Therefore, the estimate $s_e(k)$ is given by:

$$\begin{aligned} s_e(k) = {} & \sum_{i=1}^{10} b_i^{k-1} d_q(k-1) + g_{ap}(k)\left\{-\frac{1}{2}\sum_{i-1}^{4} \varphi_i^{(k-1)} d_2(k-1)\right\} \\ & + g_{fp}(k)\left[\sum_{i=1}^{16} c_i d_3(k-i)\right] + e_4(k) \end{aligned} \tag{24}$$

The prediction error $e(k)$ is normalized by the scale factor $y(k)$ that is calculated from equation (50) below. The resultant signal is then quantized into the quantized difference $d_q(k)$ by a 15-level step size quantizer whose step sizes are given in Table 7.

TABLE 7

Quantizer Normalized Input/Output Characteristic for COM XVIII-101

Normalized Quantizer Input Range $\log_2 \lvert d(k) \rvert - y(k)$	$\lvert I(k) \rvert$	Normalized Quantizer Output $\log_2 \lvert d_q(k) \rvert - y(k)$
[2.344, + ∞)	7	2.681
[1.776, 2.344)	6	2.007
[1.361, 1.776)	5	1.546
[1.013, 1.361)	4	1.175
[0.7030, 1.013)	3	0.8512
[0.4143, 0.7030)	2	0.5548
[0.1369, 0.4143)	1	0.2739
	0	0.0
[–0.1369, 0.1369)	8	0.0
[–0.4143, –0.1369)	9	–0.2739
[–0.7030, –0.4143)	10	–2.681
[–1.013, –0.7030)	11	–0.8512
[–1.361, –1.013)	12	–1.175
[–1.776, –1.361)	13	–1.546
[–2.334, –1.776)	14	–2.007
(– ∞, –2.344)	15	–2.681

In the above equations, the variables $d_l(k)$, $d_2(k)$, $d_3(k)$ and $e_4(k)$ are filtered versions of $d_q(k)$, and are obtained as follows:

$$\begin{aligned} d_l(k) &= d_q(k) + e_1(k) \\ &= d_q(k) + \sum_{i=1}^{10} b_i^{k-1}\, d_q(k-1) \\ &= \left[1 + \sum_{i=1}^{10} b_i^{k-1} B^i\right] d_q(k) \\ &= \left[1 + \Theta(B)\right] d_q(k) \end{aligned} \tag{25}$$

Now, we have

$$\begin{aligned} d_2 &= d_1(k) + e_2(k) \\ &= \left[1 + \sum_{i=1}^{10} b_i^{k-1} B^i\right] d_q + g_{ap}(k)\left[-\frac{1}{2}\sum_{i=1}^{4} \varphi_i^{k-1}\, d_2(k-i)\right] \end{aligned}$$

Therefore

$$d_2(k) = \frac{\left[1 + \sum_{i=1}^{10} b_i^{k-1} B^i\right]}{\left[1 + \frac{g_{ap}(k)}{2} \sum_{i=1}^{4} \varphi_i^{k-1} B^i\right]}\, d_q(k)$$

$$= \frac{\left[1 + \Theta(B)\right]}{\left[1 + g_{ap}\frac{(k)}{2}\, \Phi_{ap}(B)\right]}\, d_q(k) \tag{26}$$

$$d_3(k) = d_2(k) + e_3(k)$$

$$= \frac{\left[1 + \sum_{i=1}^{10} b_i^{k-1} B^i\right]}{\left[1 + \frac{g_{ap}(k)}{2} \sum_{i=1}^{4} \varphi_i^{k-1} B^i\right]} d_q(k) + g_{fp}(k) \sum_{i=1}^{16} c_i d_3(k - i)$$

$$= \frac{\left[1 + \sum_{i=1}^{10} b_i^{k-1} B^i\right]}{\left[1 + \frac{g_{ap}(k)}{2} \sum_{i=1}^{4} \varphi_i^{k-1} B^i\right]\left[1 - g_{fp}(k) \sum_{i=1}^{16} c_i B^i\right]} d_q(k)$$

$$= \frac{[1 + \Theta(B)]}{\left[1 + g_{ap}\frac{(k)}{2} \Phi_{ap}(B)\right]\left[1 - g_{fp}(k)\,\Phi_{fp}(B)\right]} d_q(k) \tag{27}$$

$$e_4(k) = \begin{cases} (1 - 2^{-7})\, e_4(k - 1) + 2^{-14} \operatorname{tsgn}\{d_q(k - 1)\}, & \text{if } 1/y(k - 1) \geq 256 \\ (1 - 2^{-7})\, e_4(k - 1) + 2^{-11} \operatorname{tsgn}\{d_q(k - t)\}, & \text{otherwise} \end{cases}$$

i.e.

$$e_4(k) = \begin{cases} \dfrac{2^{-14} \operatorname{tsgn}[d_q(k - 1)]}{1 - (1 - 2^{-7})B}, & \text{if } 1/y(k - 1) \geq 256 \\ \dfrac{2^{-11} \operatorname{tsgn}[d_q(k - 1)]}{1 - (1 - 2^{-7})B}, & \text{otherwise} \end{cases} \tag{28}$$

In equation (28), $1/y(k)$ is the linear scale factor of the adaptive quantizer obtained from the logarithmic scale factor $v(k)$ as shown in equation (50).

Therefore equation (24) can be written as:

$$s_e(k) = \left[\sum_{i=1}^{10} b_i^{k-1} B^i\right.$$

$$\left. + g_{ap}(k) \left\{ -\frac{1}{2} \sum_{i=1}^{4} \varphi_i^{k-1} \left\{ \frac{[1 + \sum_{j=1}^{10} b_j^{k-i-1} B^j]}{[1 + \frac{g_{ap}(k - i)}{2} \sum_{l=1}^{4} \varphi_l^{k-i-1} B^3]} B^i \right\} \right\} \right]$$

$$+ g_{fp}(k) \sum_{i=1}^{16} c_i \frac{[1 + \sum_{j=1}^{10} b_j^{k-i-1} B^j]}{[1 + \frac{g_{ap}(k - i)}{2} \sum_{l=1}^{4} \varphi_l^{k-i-1} B^l] \; [1 - g_{fp}(k - 1) \sum_{m=1}^{16} c_m B^m]} \Bigg] d_q(k - i)$$

$$+ \frac{\gamma \, \mathrm{tsgn}\,[d_q(k - 1)]}{1 - (1 - 2^{-7})\, B} \tag{29}$$

with

$$\gamma = \begin{cases} 2^{-14}, & \text{if } 1/y(k - 1) \geq 256 \\ 2^{-11}, & \text{otherwise.} \end{cases} \tag{30}$$

4.2 The Fixed AR Filter

The coefficients c_i, $i = 1, ..., 19$ of the fixed AR filter are optimized for an averaged spectrum of V.29 modem signals. The corresponding values are as follows:

i	1	2	3	4	5	6	7
c_i	1.0034	–1.7419	1.4170	–1.8785	1.0615	–1.2004	–0.2230
i	8	9	10	11	12	13	14
c_i	–0.3077	–0.4208	0.1938	–0.5917	0.2434	–0.3826	0.0840
i	15	16	17	18	19		
c_i	–0.1217	–0.0214	0.5017	–0.36755	–0.5343		

The gain $g_{fp}(k)$ is adapted according to the following equation:

$$g_{fp}(k) = \begin{cases} 0 & , \text{ if } g_4(k - 1) \leq 0 \\ 1{,}1 & , \text{ if } g_4(k - 1) \geq 1{,}1 \\ g_4(k - 1), & \text{otherwise} \end{cases} \tag{31}$$

where

$$g_4(k) = \begin{cases} (1 - 2^{-10})\, [g_4(k - 1) - 1] + 1 - 2^{-8} & , \text{ if } 1/y(k) \geq 750 \\ (1 - 2^{-10})\, [g_4(k - 1) - 1] + 1 + 2^{-8}\, \mathrm{tsgn}\,[d_2(k)]\, \mathrm{tsgn}\,[e_6(k)], & \text{otherwise} \end{cases} \tag{32}$$

with

$$g_4(0) = 0 \text{ and } g_4(k) \text{ limited to } -0.125 \leq g_4 \leq 1.25$$

Theoretically, the gain factor $g_4(k)$ for the fixed AR predictor should have been updated using the correlation between $d_2(k)$ and the output $e_3(k)$ of the fixed AR predictor. A moving average of $d_2(k)$, denoted as $e_6(k)$ is used instead to protect against possible divergence due to bit errors in the transmission path. The value of $e_6(k)$ is obtained from:

$$e_6(k) = \sum_{i=1}^{3} c_{i+16}\, d_2(k - i) = \sum_{i=1}^{3} c_{i+16} \frac{[1 + \sum_{j=1}^{10} b_j^{k-1} B^j]\, B^i}{[1 + \frac{g_{ap}(k)}{2} \sum_{l=1}^{4} \varphi_b^{k-1} B^l]}\, d_q(k) \tag{33}$$

4.3 Adaptive AR Filter

In a practical implementation, it is not easy to monitor the locations of the roots of an AR(n) adaptive filter, when $n > 2$. Therefore, to ensure stability, the adaptive AR filter is chosen in the form of polynomials that satisfy the "strictly positive real conditions" at all times [28]. In the algorithm of COM XVIII-101, the AR filter is expressed as a pair of Chebyshev polynomials (i.e. mirror and antimirror image polynomials) in the form:

$$\begin{aligned} \Phi_{ap}(B) &= (\sum_{i=1}^{4} \varphi_i\, B^i) \\ &= 1 - \frac{1}{2}\left((1 + B \prod_{i=1}^{2} (B^2 - 2\cos\xi_i B + 1)\right) \\ &\quad - \frac{1}{2}\left((1 + B \prod_{i=1}^{2} (B^2 - 2\cos\varsigma_i B + 1)\right) \end{aligned} \tag{34}$$

Here φ_i are the coefficients of the autoregressive filter and ξ_i and ζ_i are the roots of each pair of Chebyshev polynomial equations. According to the Hurwitz stability theory, these roots must be on the unit circle in the complex B plane and they must be simple and interleaved such that $0 < \xi_1 < \zeta_1 < \xi_2 < \zeta_2 < \pi$. These roots correspond to frequencies according to the following relation

$$\frac{2\pi f}{T}$$

where

$$T = 125\ \mu\text{s}$$

In the following discussion, and to facilitate notation, the following equivalences, will be used: $z_l \equiv \xi_1$, $z_2 \equiv \zeta_1$, $z_3 \equiv \xi_2$, $z_4 \equiv \zeta_2$, and $r_1 \equiv 2\cos\xi_1$, $r_2 \equiv 2\cos\zeta_1$, $r_3 \equiv 2\cos\xi_2$, $r_4 \equiv 2\cos\zeta_2$.

Equation (34) can be expanded as:

$$\Phi_{ap}(B) = \left(\sum_{i=1}^{4} \varphi_i B^i\right)$$

$$= 1 - \frac{1}{2}(1 + B)\left[B^4 - (r_1 + r_3) B_3 - (2 - r_1 r_3) B^2 - (r_1 + r_3) B + 1\right]$$

$$- \frac{1}{2}(1 - B)\left[B^4 - (r_1 + r_2) B_3 - (2 - r_1 r_4) B^2 - (r_2 + r_4) B + 1\right]$$

$$= 1 - \frac{1}{2}\left[B^5 + (1 - r_1 - r_3) B_4 + (-2 - r_1 - r_3 + r_1 r_3) B^3\right.$$

$$+ (-2 - r_1 - r_3 + r_1 r_3) B^2 + (1 - r_1 - r_3) B + 1 \Big]$$

$$- \frac{1}{2} \Big[- B^5 + (1 + r_2 + r_4) B_4 + (2 - r_2 - r_4 - r_2 r_4) B^3$$

$$+ (-2 + r_2 + r_4 + r_2 r_4) B^2 + (-1 - r_2 - r_4) B + 1 \Big]$$

$$= \quad - \frac{1}{2} \Big[(2 - r_1 - r_2 + r_3 + r_4) B_4 + (- r_1 - r_2 - r_3 - r_4 + r_1 r_3 - r_2 r_4) B^3$$

$$+ (-4 - r_1 + r_2 - r_3 + r_4 + r_1 r_3 + r_2 r_4) B^2 + (- r_1 - r_2 - r_3 - r_4) B \Big] \qquad (35)$$

By comparing both sides of equation (35), we obtain the following relations of the coefficients φ_i:

$$\left. \begin{aligned} \varphi_1 &= -\frac{1}{2}(- r_1 - r_2 - r_3 - r_4) = -\frac{1}{2}\left(\sum_{j=1}^{4} r_j\right) \\ \varphi_2 &= -\frac{1}{2}(4 - r_1 + r_2 - r_3 + r_4 + r_1 r_3 + r_2 r_4) \\ \varphi_3 &= -\frac{1}{2}(- r_1 - r_2 - r_3 - r_4 + r_1 r_3 - r_2 r_4) \\ &= \varphi_1 - \frac{1}{2}(r_1 r_3 - r_2 r_4); \\ \varphi_4 &= -\frac{1}{2}(2 - r_1 + r_2 - r_3 + r_4) \end{aligned} \right\} \qquad (36)$$

4.4 Update Equations

4.4.1 The Moving Average Filter

The MA coefficients b_i^k are updated according to the following recursion:

$$b_i^k = (1 - 2^{-7})\, b_i^{k-1} + 2^{-6} \operatorname{tsgn}[d_q(k)] \operatorname{tsgn}[d_q(k-1)] \quad \text{for} \quad i = 1, 2, ..., 10 \qquad (37)$$

with

$$\operatorname{tsgn}(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases} \qquad (38)$$

Notice that the leakage factors here are twice the leakage factors for the algorithms of Recommendations G.726 and G.727 as shown in equation (11).

4.4.2 Adaptive Filter

4.4.2.1 The Gain

The adaptive gain g_{ap} ( ) for the adaptive AR filter is calculated as follows:

$$g_{ap}(k) = \begin{cases} 0 & , \text{ if } g_1(k-1) \le 0 \\ 1 & , \text{ if } g_1(k-1) \ge 1 \\ g_1(k-1), & \text{otherwise} \end{cases} \tag{39}$$

where

$$g_1(k) = (1 - 2^{-10})\, g_1(k-1) + 2^{-8} \operatorname{tsgn}[d_1(k)] \operatorname{tsgn}[e_5(k)] \tag{40}$$

with the initial condition: $g_l(0) = 0$. Thus, the update of $g_1(k)$ depends on $d_1(k)$. Notice that the leakage factor is the same as for the gain g_{fp} of the fixed AR filter.

The variable $e_5(k)$ is a filtered estimate of $d_1(k)$ calculated from:

$$e_5(k) = -\sum_{i-1}^{4} \left\{ \varphi_i^{k-1} \left[d_1(k-1) + \frac{1}{2} e_5(k-i) \right] \right\} \tag{41}$$

The above equation can be rewritten in the form:

$$\left[1 + \sum_{i=1}^{4} \frac{\varphi_i^{k-1}}{2} B^i \right] e_5(k) = \left(- \sum_{i=1}^{4} \varphi_i^{k-1} B^i \right) d_1(k)$$

Therefore

$$e_5(k) = \frac{\left(- \sum_{i=1}^{4} \varphi_i^{k-1} B^i \right)}{1 + \sum_{i=1}^{4} \frac{\varphi_i^{k-i}}{2} B^i} \times \left[1 + \sum_{i=1}^{10} b_i^{k-1} B^i \right] d_q(k)$$

$$= \frac{-\Phi_{ap}(B)\ [1 + \Theta(B)]}{1 + \Phi_{ap}(B)} d_q(k) \tag{42}$$

Equation (39) is constructed such that $g_{ap}(k) \rightarrow 1$ for voice and $\rightarrow 0$ for voiceband data.

4.4.2.2 AR Coefficients

The coefficients of the adaptive pole predictor φ_i^k are updated through the following steps.

First, the interleaved roots of the Chebyshev polynomials z_i^k (i = 1, ..., 4) are updated as specified in the following equations:

$$z_i^k = \begin{cases} (1 - 2^{-7})\,[z_i^{k-1} - Z_i] + Z_i - 2^{-8} & \\ \quad \times \operatorname{tsgn}[d_q(k)] \operatorname{tsgn}[p_i(k-1)], & \text{if } z_i^{k-1} - z_{i-1}^{k-1} > Z_i \text{ and } z_{i+1}^{k-1} - z_i^{k-1} > Z_i + 1 \\ (1 - 2^{-7}) \left[z_i^{k-1} - Z_i \right] + Z_i, & \text{otherwise,} \end{cases} \tag{43}$$

where

$$Z_0^k = 0, Z_5^k = \pi)$$

The initial values and the corresponding frequencies are given in the following tables:

Root Location		Frequency (Hz)	
Variable	Value	Variable	Value
z_0^0	0.3501	f_0^0	0
z_1^0	0.3501	f_1^0	445.8
z_2^0	0.5520	f_2^0	702.8
z_3^0	1.5010	f_3^0	1911.1
z_4^0	2.2160	f_4^0	2821.5
z_5^0	3.1416 (π)	f_5^0	4000

Root Location		Frequency (Hz)
Variable	Value	Value
z_1	0.1570	200
z_2	0.0313	40
z_3	0.0313	40
z_4	0.0313	40
z_5	0.3928	500

Thus, the conditions for updating the location of the roots can be expressed as shown in the following table:

Root No.	Condition No. 1	Condition No. 2
1	$f_1^{k-1} - f_0^{k-1} > 200$	$f_2^{k-1} - f_1^{k-1} > 40$
2	$f_2^{k-1} - f_1^{k-1} > 40$	$f_3^{k-1} - f_2^{k-1} > 40$
3	$f_3^{k-1} - f_2^{k-1} > 40$	$f_4^{k-1} - f_3^{k-1} > 40$
4	$f_4^{k-1} - f_3^{k-1} > 40$	$f_5^{k-1} - f_4^{k-1} > 500$

Because, f_0 is always 0 Hz and f_5 is always 4000 Hz, the above conditions show that the first update equation is used when one or more of the frequencies corresponding to the roots of the Chebyshev polynomials is in the range (200, 3500) Hz and is separated from its neighbour by at least 40 Hz.

If these conditions are not satisfied, then the locations of the roots are updated according to the leakage operation which converges the initial values.

Let, as before, $r_i^k = 2 \cos z_i^k$. The gradients p_i, $i = 1, ..., 4$ are calculated as

$$\frac{\partial \varphi_{ap}(B)}{\partial z_i}, \; i = 1; ...; 4.$$

Therefore

$$\frac{\partial \varphi_{ap}(B)}{\partial \, z_i} = \partial \left(\sum_{i=1}^{4} \varphi_i B^i \right)$$

$$= - \sin z_1 \left[B^4 + (B^3 + B^2)(1 - r_3) + B \right] \tag{44.a}$$

$$\frac{\partial \varphi_{ap}(B)}{\partial z_2} = - \sin z_2 \left[- B^4 + (B^3 - B^2)(1 + r_4) + B \right] \tag{44.b}$$

$$\frac{\partial \varphi_{ap}(B)}{\partial z_3} = - \sin z_3 \left[B^4 + (B^3 + B^2)(1 - r_1) + B \right] \tag{44.c}$$

$$\frac{\partial \varphi_{ap}(B)}{\partial z_d} = - \sin z_4 \left[- B^4 + (B^3 - B^2)(1 + r_2) + B \right] \tag{44.d}$$

Because by construction $0 < z_i < \pi$ so that $\sin z_i > 0$ we have:

$$\operatorname{tsgn}[p_1(k)] = - \operatorname{tsgn}[d_q(k-1) + (1 - r_3^k)\{d_q(k-2) + d_q(k-3)\} + d_q(k-4)] \tag{45.a}$$

$$\operatorname{tsgn}[p_2(k)] = - \operatorname{tsgn}[d_q(k-1) - (1 + r_4^k)\{d_q(k-2) - d_q(k-3)\} - d_q(k-4)] \tag{45.b}$$

$$\operatorname{tsgn}[p_3(k)] = - \operatorname{tsgn}[d_q(k-1) + (1 - r_1^k)\{d_q(k-2) + d_q(k-3)\} + d_q(k-4)] \tag{45.c}$$

$$\operatorname{tsgn}[p_4(k)] = - \operatorname{tsgn}[d_q(k-1) - (1 + r_2^k)\{d_q(k-2) - d_q(k-3)\} - d_q(k-4)] \tag{45.d}$$

After the new root locations have been calculated, then the update equations of φ_i are:

$$\varphi_1^{(k+1)} = - \frac{1}{2} (- r_1^k - r_2^k - r_3^k - r_4^k) = - \frac{1}{2} \left(\sum_{j=1}^{4} r_j^k \right)$$

$$\varphi_2^{(k+1)} = - \frac{1}{2} \left(- 4 - r_1^k + r_2^k - r_3^k + r_4^k + r_1^k r_3^k + r_2^k r_4^k \right)$$

$$\varphi_3^{(k+1)} = -\frac{1}{2}\left(-r_1^k - r_2^k - r_3^k - r_4^k + r_1^k r_3^k - r_2^k r_4^k\right) \tag{46}$$

$$= \varphi_i^k - \frac{1}{2}\left(,r_1^k r_3^k - r_2^k r_4^k\right)$$

$$\varphi_4^{(k+1)} = -\frac{1}{2}\left(2 - r_1^k + r_2^k - r_3^k + r_4^k\right)$$

4.5 Quantizer Adaptation

The adaptation of the adaptive quantizer employs a scale factor with a dual adaptation speed. The adaptation of the scale factor and of the speed is described in the following two subclauses.

4.5.1 Quantizer Scale Factor Adaptation

The quantizer scale factor is a weighted average of a fast scale factor, $v_1(k)$, and a slow scale factor, $v_2(k)$.

The fast scale quantizer scale factor is computed recursively as follows:

$$v_1(k) = (1 - 2^{-6})\, v(k) + W_s[I(k)] + l(k)\, W_d[I(k)] \tag{47}$$

and the slow scale factor is given by:

$$v_2(k) = (1 - 2^{-3})\, v_2(k - 1) + 2^{-3}\, v_1(k) \tag{48}$$

with $v_1(0) = -11.2876$ and $v_2(0) = 0$. $W_s[I(k)]$ and $W_d[I(k)]$ are step functions optimized for speech and voiceband data respectively. These functions are defined in Appendix I.

The parameter $l(k)$, which is the output of the adaptation controller, plays a role analogous to that played by the function $a_l(k)$ for Recommendations G.726 and G.727 as defined in equation (15). Thus, it is derived on the basis that the rate of change of the quantized first difference $d_q(k)$ be in the range [0-1] as explained in the next clause.

From a comparison of equation (47) with equation (12), it is clear that the algorithm of COM XVIII-102 involves two $W[I(k)]$ functions instead of one. Similarly, by comparing equation (48) with equation (11), it is clear that the leakage factors used in COM XVIII-101 are larger.

Equation (48) can be written in the following form:

$$v_2(k) = \frac{2^{-3}\, v_1(k)}{[1 - (1 - 2^{-3})\, B]} \tag{48′}$$

The combined scale factor is given by

$$v(k) = [1 - l(k - 1)]\, v_2(k) + 2^{-3}\, v_1(k) \tag{49.a}$$

Using the expression of $v_2(k)$ in equation (48), $v(k)$ takes the form:

$$v(k) = \left\{\frac{[1 - l(k - 1)] \cdot 2^{-3}}{[1 - (1 - 2^{-3})\, B]} + 2^{-3}\right\} v_1(k) \tag{49.b}$$

$v(k)$ is limited to $-11.2876 \le v(k) \le -1.3218$.

By combining equations (47) and (49.b) we obtain:

$$v(k) = \left\{ \frac{[1 - l(k - 1)] \cdot 2^{-3}}{[1 - (1 - 2^{-3})\, B]} + 2^{-3} \right\} \left\{ (1 - 2^{-6})\, v(k) + W_s[I(k)] + l(k)\, W_d[I(k)] \right\}$$

$$= \frac{1 - 2^{-3}\, [1 - l(k - 1) - (1 - 2^{-3})\, B]}{1 - [1 - 2^{-3} l(k - 1)](1 - 2^{-6}) - (1 - 2^{-3})\, B[1 - 2^{-3}(1 - 2^{-6})]}$$

$$\left\{ W_s[I(k)] + l(k)\, W_d[I(k)] \right\} \tag{49.c}$$

The logarithmic scale factor $v(k)$ corresponds to the linear scale factor $y(k)$ through the relation:

$$y(k) = 2^{v(k)} \tag{50}$$

4.5.2 Adaptation Speed Control

The speed of adaptation depends on the change in the average magnitude of the 4-bit code output, $I(k)$, and on the correlation between the normalized quantizer output $Q(k)$ and a filtered estimate of $Q(k)$, $e_8[Q(k) = d_q \frac{(k)}{y(k)}]$. Variations of the average of $I(k)$ are indicative of the change in the signal power, while the correlation of $Q(k)$ and $e_8(k)$ gives an estimate of the frequency content of the quantized difference signal $d_q(k)$.

The measure for power variation is the normalized variable $d_m(k)$ given by:

$$d_m(k) = (1 - 2^{-7})\, d_m(k - 1) + 2^{-7}\, F[I(k)] \tag{51}$$

with $d_m(0) = 0$, and $F[I(k)]$ is the step function described in Appendix I.

We introduce the intermediate variables $t_1(k)$ and $t(k)$ such that:

$$t_1(k) = \begin{cases} (1 - 2^{-7})\, t_1(k - 1) + 2^{-7}, & \text{if } t(k) \geq 0.8 \\ (1 - 2^{-7})\, t_1(k - 1) & , \text{otherwise} \end{cases} \tag{52}$$

and

$$t(k) = \begin{cases} (1 - 2^{-10})\, [t(k - 1) - 1] + 1 - 2^{-8} & , \text{if } d_m(k) < 2^{-2} \\ (1 - 2^{-10})\, [t(k - 1) - 1] + 1 + 2^{-8}\, \text{tsgn}\, [Q(k)]\, \text{tsgn}\, [e_8(k)], & \text{otherwise} \end{cases} \tag{53}$$

$t(k)$ is limited to $0 \leq t(k) \leq 1$, and $Q(k)$ is the normalized quantizer output.

In the above equations, $t(k) \to 0$ when $d_m(k)$ is less than a specific power level, and $\to 1$ otherwise, which is the case for modem signals. The signal $t_1(k)$ is a filtered version of $t(k)$.

The variable $e_8(k)$ is the output of an ARMA (2,2) filter that acts on the normalized quantizer output $Q(k)$. The parameters of this filter are selected to represent an average V.29 modem signal at 9.6 kbit/s.

Thus, we write:

$$e_8(k) = \frac{[-\ 0.123\, B - 0.787\, B^2]}{1 + \frac{0.123}{2} B + \frac{0.787}{2} B^2}\, Q(k) \tag{54}$$

Finally, the mode controlling parameter $I(k)$ is a low-pass filtered version of $t_1(k)$, which is itself a filtered version of $t(k)$. These two low-pass operations are used to prevent errors.

$$l(k) = \begin{cases} (1 - 2^{-7})\, l(k-1) + 2^{-7}, & \text{if } t_1(k) \geq 0.8 \\ (1 - 2^{-7})\, l(k-1) \quad , & \text{otherwise} \end{cases} \tag{55}$$

The mode controlling parameter $l(k)$ will be 0 for rapidly varying signals such as speech, and will be 1 for slowly varying signals such as voiceband data. [This is the reverse of the parameter $a_1(k)$ for Recommendations G.726 and G.727.]

5 ADPCM Decoder

5.1 General Description

The ADPCM decoders for both fixed-rate and embedded operation have the same basic structure. The main difference is that, in embedded operation, the feed-back path uses the core bits to calculate the signal estimate. In fixed-rate operation, the feed-back path uses the whole code word to calculate the signal estimate. The feed-forward path in both cases uses the whole code word to reconstruct the output PCM code word. Both the decoder and the encoder have similar structure. Furthermore, the transmitting side has a local decoder that operates on the same signal that the receiving decoder would receive if there were no line errors. Each decoder has two inverse adaptive quantizers, one in the feed-forward loop and the other in the feedback loop. The feed-forward path of the decoder begins at the input of the feed-forward inverse adaptive quantizer and ends at the output of the synchronous coding adjustment block. In embedded operation, the feedback path is from the input of the bit masking block to the output of the adaptive predictor.

The feed-forward inverse adaptive quantizer uses $I(k)$, to generate the feed-forward quantized difference $d_q(k)_{FF}$. In constrast, the feedback inverse adaptive quantizer of the embedded decoder uses the core bits, $I_c(k)$, to compute the feedback quantized difference, $d_q(k)_{FB}$. [The feedback inverse adaptive quantizer of the fixed rate decoder uses the whole code word $I(k)$.] Adaptive prediction and signal reconstruction proceeds on the basis of $d_q(k)_{FB}$ as in the encoder. Finally, the output PCM code word is reconstructed by adding $s_e(k)$ to $d_q(k)_{FF} \cdot I(k)$. Note that in Recommendation G.727, as long as $I(k) > I_c(k)$, the feed-forward inverse quantizer will operate at a higher bit rate than of the feedback inverse quantizer.

5.2 Synchronous Coding Adjustment

Synchronous coding adjustment is a method to prevent cumulative distortion occurring on synchronous tandem codings (i.e. of the form PCM-ADPCM-PCM-ADPCM) when the integrity of the bit stream is preserved [38]. Bit integrity means that no digital signal processing devices disturb the bit streams (such as through digital to analog conversions, echo cancellation, digital filtering, digital loss or gain adjustment or robbed-bit signaling). Synchronous adjustment is available for Recommendations G.726 and G.727 and COM. XVIII-101. It is not possible for the algorithm of COM XVIII-102 due to the band limiting effects of the filter used in the encoder and in the decoder.

The basic concept is to ensure that the output PCM will be converted to the same ADPCM code word at all stages, even by modifying the output PCM if necessary [26; 38; 44]. This ensures that all ADPCM code words are the same at every stage and that from the second stage onwards, the code words in every transcoding pair stage are the same.

The inclusion of a synchronous coding adjustment guarantees that the PCM-to-PCM signal to noise ration is preserved, if not improved [40; 41]. An unpublished subjective evaluation by Bowker [1986] showed that the subjective voice quality using 32 kbit/s ADPCM is the same after four synchronous encodings as after a single encoding. If the transmitting encoder and the receiving decoder have different initial conditions, such as after line errors, then the synchronous tandeming property will take time to establish since the encoder/decoder must be tracking. As a rule of thumb, resynchronization may require about 500 ms for tones and about 250 ms for speech.

6 Objective Evaluation of ADPCM

6.1 Theoretical Background

This clause outlines an approach to analyse the performance of ADPCM analytically. It is assumed that there is no distinction between a mid-tread and a mid-rise quantizer. For Recommendation G.727, this assumption is valid when the number of bits in the feedback loop is 4 or 5.

From equation (2), we can write

$$E\left[(1 + \sum_{i=1}^{6} b_i^{k-1} B^j)\, d(k)\right]^2 = E\left[(1 - \sum_{j=1}^{2} a_j^{k-1} B^i)\, s_l(k)\right]^2$$

$$+ E\left[\left(\sum_{i=1}^{6} (a_i^{k-1} + b_i^{k-1})\, B^i\right) d_q(k)\right]^2 \qquad (56)$$

with

$$a_i^{k-1} = 0 \text{ for } i = 3, ..., 6$$

It is assumed that the input signal $s_l(n)$ is uncorrelated with the quantizer noise $e(n)$ and that $e(n)$ is white. The latter assumption is valid for an optimal quantizer with 2 or more quantization levels [35; p. 158].

Equation (56) can be written as:

$$\Theta'\, T\Theta = A'RA + E\,[e^2(n)]\, P^2 \qquad (57)$$

where

$$\Theta' = [1, - b_1^{k-1}, ..., b_6^{k-1}]$$

$$A' = [1, - a_1^{k-1}, - a_2^{k-1}]$$

$$P^2 = \sum_{i=1}^{2} (a_i^{k-1})^2 + \sum_{j=1}^{6} (b_j^{k-1})^2 + 2(a_1^{k-1}\, b_1^{k-1} + a_2^{k-1}\, b_2^{k-1})$$

T is a 7×7 matrix with the (i, j)th element given by $E[d(k - i + 1)\, d(k - j + 1)]$, and R is a 3×3 matrix with the (i, j)th element given by $E[s_l(k - j + 1)\, s_l(k - j + 1)]$

The gain G of ADPCM is given by:

$$G = \frac{E\,[s_l^2(k)]}{E\,[d^2(k)]} = \frac{1}{A\tilde{R}A}\left[\Theta\tilde{T}\Theta - \frac{P^2}{SNR_q}\right] \qquad (58)$$

where the quantized Signal to Noise Ratio SNR_q is given by:

$$SNR_q = \frac{E[d^2(k)]}{E[e^2(k)]}$$

$$\tilde{R} = \frac{R}{E[s_l^2(k)]}, \quad \tilde{T} = \frac{T}{E[d^2(k)]}$$

Thus we can write equation (58) as:

$$G = G_1 G_2$$

with

$$G_1 = \frac{1}{A'\tilde{R}A} \quad \text{and } G_2 = \Theta'\tilde{T}\Theta - \frac{P^2}{SNR_q} \tag{59}$$

G_1 depends on the autocorrelation of the input signal s_l and the poles of the predictor coefficients. Hence, it depends indirectly on the quantized difference signal through the adaptation described in [39].

G_2 represents the degradation of the prediction gain because of:

1) effects of past residuals;

2) the indirect effect of the quantized difference signal on the update of the zeros of the predictor; and

3) the effects of the quantized difference being passed through the adaptive predictor.

Note that for an all-poles predictor, equations (56)-(59) reduce to equations (17)-(22) of Suzuki and Taka [51]. Therefore, equation (59) can be solved to estimate G and SNR_q by a numerical estimation procedure similar to the one they have described.

6.2 Performance of the ADPCM Algorithms for Voiceband Data

The test conditions for evaluating the performance of the ADPCM voiceband data under various network conditions are defined by the so-called “R-28” impairments, their designation referring to an annex of a December 1983 CCITT SG XVIII report of the same number [13]. An outline of the tested modems is available [1] and results for the modified G.721 algorithm approved in 1986 have been presented to the CCITT [1; 2].

The results in these contributions show that the modifications allow good performance with character mode FSK voice band data and eliminates the decoder oscillatory behavior for idle code inputs. The modifications did not significantly alter the performance for other speech and non-speech signals except for single tones. In this case, the degradation was about 1.5 dB in S/N ratio.

6.2.1 Performance with 9.6 Kbit/s voiceband data

The objective tests for the extension of 32 kbit/s ADPCM to 40 kbit/s were conducted between October 1986 and May 1987 by COMSAT Laboratories. The test program included conditions of single, two and three asynchronous encodings for the following V.29 modems: CODEX 2640, OKI VLSI96, CODEX LSI/V.29 (at 9.6 kbit/s and 4.8 kbit/s) and Racal-Milgo Omni 96. In addition, operation with the CODEX LSI/V.27 (at 4.8 kbit/s) and with the Racal Milgo V.33 modem (9.6 kbit/s) was also tested. Some results are available in two COMSAT contributions to the CCITT [23; 24]. The full results are also available in a 7-volume set from COMSAT that contains complete detailed data covering all measurements, typically 100 error events per data point [25]. Finally, a brief description of the measurement program and some representatives examples are available in a published paper [29].

In the initial phase, four candidate algorithms were tested: a 16-level version of an algorithm proposed by CNET (France), two versions of an algorithm from OKI and KDD (Japan), and an algorithm from ECI (Israel) [23]. In the second phase, the CNET algorithm was modified to operate with a 31-level quantizer, and there was one unified algorithm from OKI/KDD [24] also operating at 15-level. Finally, additional testing with the retained version of the

40 kbit/s algorithm were presented to the CCITT for V.29 modems using the switched carrier operation on the primary and secondary channels, and V.32 modems [3]. These results show that the secondary channel affects performance significantly, although the Block Error Rate (BLER) remains less than $l0^{-2}$ at 30 dB S/N ratio. The results also show that performance at rates higher than 12 kbit/s depends on the characteristics of the modem used.

Analogue performance measurements did not show substantial difference among the various algorithms, except that the restricted 3.2 kHz bandwidth of COM XVIII-102 causes some amplitude and group delay distortions in the frequency response. The algorithm of COM XVIII-101 was marginally better, although more complex, than the algorithms of Recommendation G.726 and COM XVIII-102. Finally, the algorithm of Recommendation G.726 was consistently superior to the algorithm of COM XVIII-102 for the case of V.29 modems operating in the presence of digital link bit errors.

6.2.2 Performance with High Speed Modems

In 1992 and 1993, many activities took place in the area of high speed modems (see Recommendation V.34) and videotelephony on the PSTN. These activities have been based on novel approaches to demodulation. With this situation in mind, some preliminary investigations have been conducted at AT&T Bell Laboratories to evaluate the implication of these new modulation schemes on performance of the various ADPCM algorithms [4].

The following preliminary conclusions can be made when the only noise in the end-to-end system is Gaussian noise:

1) Connections that encounter the 40 kbit/s ADPCM algorithm (Recommendation G.726 or 32-level) can support a data rate up to 19.2 kbit/s with some marging. Connections that contain COM XVIII-102 ADPCM algorithm can support data rates up to 16.8 kbit/s with some margin. Connections that encounter the COM XVIII-101 ADPCM algorithm can support data rates up to 14.4 kbit/s with some margin.

2) Increasing the baud from 2400 to 2800 improves the operating margin for all ADPCM algorithms by about 1-2.5 dB.

3) The algorithm of COM XVIII-102 is not capable of supporting 3200 baud. International connections at 3200 baud that encounter this algorithm will not connect at any data rate.

4) Only 64 kbit/s PCM or 40 kbit/s ADPCM (Recommendation G.726 or the 32-level algorithm), whether packetized or not, can support 19.2 kbit/s operation. However, with 40 kbit/s ADPCM and for a BLER of $l0^{-3}$, the SNR margin is reduced by about 4.8 dB for the conditions of this test.

Because of the assumption that there are no errors from the digital lines, practical connections with a mixture of analog and digital transmission lines may have worse performance than those shown by these results.

6.3 Objective Measurements

While there are not yet standardized objective measurements for ADPCM algorithms, the guidelines of CCITT Recommendation G.712 regarding PCM testing are often used. Before the 1992 revision, G.712 recommended two methods: the band-limited noise method and the sine-wave method. In the band-limited noise method, the input is a random noise shaped with a specified spectrum; in the sine-wave method, the input signal is a single tone of a defined frequency. For PCM, both methods are not exactly equivalent because they respond to different impairments in slightly different ways.

While both methods exercise the algorithmic properties of the ADPCM algorithms, they do not suitably reflect the subjective evaluation of the distorsion that ADPCM introduces to speech waveforms [14].

6.3.1 Measurements with the Quasi-Random Noise Test Method

The quasi-random noise signal had a frequency spectrum conforming to CCITT Recommendation O.131. This was obtained by passing the noise through a band pass filter with 3 dB points at 350 Hz and 550 Hz. The measurements were obtained from a hardware model of a family of ADPCM algorithms, among which are the fixed rate 32-kbit/s ADPCM algorithm of Recommendation G.726 and the embedded algorithms of Recommendation G.727. Table 8 shows the difference in the signal-to-noise ratio between the 32 kbit/s algorithm of Recommendation G.726 and the (4,2) algorithm of Recommendation G.727 for various input levels.

TABLE 8

ADPCM/Embedded ADPCM Noise Measurements

Input (dBm0)	–5	–10	–15	–20	–25
SNR difference (dB)	0.2	0.3	0.5	0.1	0.3

The results show that the fixed rate algorithm has a slightly higher SNR, which is to be expected. However, results of this test method are not always easy to interpret because of the limitations described below.

In practice, the quasi-random noise can fake a tone and cause a spurious transition detection. Because the prediction coefficients will be reset subsequently, the difference signal may become large, thereby increasing the quantized difference signal. Because the response of the slow speed scale factor $y_l(k)$ can be delayed, the adaptive quantizer may be unable to adapt rapidly. This causes dips in the PCM output of the decoder that the listener may perceive as clicks [42; 43]. The dips are more pronounced for A-law than for μ-law.

Ironically, the better the prediction, the more pronounced is the clicking phenomenon. Clicks are more significant at 40 kbit/s where they occur every few seconds; at 32 kbit/s the average interval between clicks is a few minutes. They are also more pronounced for A-law than for μ-law.

Further investigation was done with the following algorithms for both A-law and μ-law [22]:

1) Fixed rate algorithms of Recommendation G.726 (at the rates of 40 kbit/s, 32 kbit/s, 24 kbit/s and 16 kbit/s).

2) Embedded algorithms for Recommendation G.727, i.e. (5,2), (4,2), (3,2), (5,3), (4,3), (3,3), (5,4), and (4,4).

As previously mentioned, the 16 kbit/s algorithm of Recommendation G.726 is the same as the embedded (2,2) algorithm of Recommendation G.727.

In addition, the (5,5) embedded ADPCM algorithm was also tested. This algorithm is described in Appendix 1 of Annex 1 of Q.24/XV for the 1988-1992 study period, but is not a part of the final Recommendation. The purpose of including the (5,5) case was to compare and contrast a "voice" 40 kbit/s fixed-rate algorithm and the 40 kbit/s algorithm of Recommendation G.726, which was designed specifically to allow the transport of modem data at 9.6 kbit/s.

The measurements were repeated four times for the fixed-rate 40-kbit/s algorithm, and for the embedded (5,4) algorithm. The measurements were repeated twice only for all the remaining algorithms.

The results showed that:

1) The response of the 40 kbit/s fixed rate algorithm shows some randomness for high input signal levels (higher than –20 dBm0). In the same region, the output is not reproducible across runs.

   This happens for both the A-law and μ-law codings.

2) The response for the 32 kbit/s fixed-rate algorithm exhibits the same features as above, although the magnitude of the fluctuations and their randomness are attenuated.

3) For the embedded coding, some fluctuations are apparent for the (5,4) and (5,3) algorithms.

4) Some turbulence occurs for the 24 kbit/s and 16 kbit/s algorithms when the A-law coding is used.

This instability can be traced to the spurious tone detection and tone reset features of the algorithm [42; 43]. The phenomenon is more likely to occur for algorithms which:

1) use more bits for the prediction and therefore track the input signal more accurately;

2) have a higher resolution in their quantizer; and

3) have a propensity to remain in the locked mode of quantization (i.e. are more tuned for voice-band data).

These three conditions are most likely to occur with the 40 kbit/s fixed rate algorithm of Recommendation G.726, because it was optimized for voice-band data, and hence, is more likely to be locked than all other algorithms. This also explains why the fluctuations are more pronounced in the case of the 40-kbit/s algorithm of Recommendation G.726 than for the (5,5) algorithm.

As a consequence, the CCITT does not recommend this method for ADPCM characterization. In fact, the method was removed from the 1992 revision of Recommendation G.712.

6.3.2 Measurements with the Sine-Wave Method

Signal-to-Noise Ratio measurements were taken over the component utilizing a 1.01 kHz signal at 0 Transmission Level Point (TLP). The input signal power levels were then varied from +3 to –45 dBm0. Table 9 gives the measured SNR values for different codings algorithm.

lt is clear from Table 9 that the SNR decreases for lower coding rates, which reflects an increase in quantization noise.

The above measurements were taken with the number of bits in the feed-forward path constant. In real applications, the number of bits in the feed-forward path may vary depending on the traffic.

TABLE 9

Sine-Wave SNR Measurements

Signal level (dBm0)	PCM (64 kbit/s)	ADPCM (32 kbit/s)	Embedded ADPCM (4/2)	Embedded ADPCM (3/2)	Embedded ADPCM (2/2)
+ 3	41.3	36.6	36	30.8	26.4
0	39.4	35.7	35.5	30.8	25.6
– 5	40.2	36.0	35.6	30.7	25.5
–10	40.5	36.5	35.2	30.4	24.8
–15	39.6	36.2	33.6	29.2	23.7
–20	38.7	35.1	33.3	28.5	22.8
–25	38.3	35.5	32.5	27.7	22.3
–30	38.0	34.5	31.0	27.0	21.2
–35	34.7	32.0	29.8	25.5	20.1
–40	32.1	29.5	27.5	23.5	18.6
–45	28.8	27.0	24.7	21.0	15.6

7 Subjective Evaluation of ADPCM

The final criterion for the acceptance of a speech coding algorithm is the user's acceptance. Despite current efforts to derive objective measures for the quality of voice, there are presently no reliable methods that can substitute for subjective tests [46; 30]. Therefore, the main thrust of the evaluation of the algorithm performance was to use subjective testing.

7.1 Subjective Evaluation of 32 kbit/s ADPCM

Results for subjective evaluation of the 32 kbit/s ADPCM algorithm are available [2]. A total of 352 trials were presented to 22 subjects, to test 22 different experimental conditions, each replicated twice with eight different speakers. Six of these conditions were reference conditions recorded through a Modulated Noise Reference Unit as defined in CCITT Recommendation P.70 [14].

7.2 Subjective Evaluation of G.721 Extensions

Subjective performance of both the 24 and 40 kbit/s algorithms are available [3]. The results show that, for a single encoding, there is little difference (less than a tenth of a mean opinion score point) between the ADPCM conditions at the rates of 32 kbit/s and 40 kbit/s. The performance of 24 kbit/s ADPCM is worse, as expected, but in a DCME environment, the use of 24 kbit/s ADPCM is an overload mechanism and the percentage of time spent at this rate should be a small fraction of the total time in properly engineered systems.

7.3 Subjective Evaluation of Embedded ADPCM

Forty-three subjects rated the quality of speech segments recorded through a large number of test conditions [18; 48; 49]. The results indicate that the performance of the five-bit embedded ADPCM with a mid-rise quantizer was essentially identical to that of the 64 kbit/s PCM at both one and four encodings. Furthermore, at four encodings, the rating of the five-bit G.727 algorithm was significantly higher than that of 32 kbit/s of Recommendation G.726. The quality for various bit rates as obtained for embedded ADPCM algorithm with 2 core bits is essentially the same as for the corresponding rate of Recommendation G.726.

Appendix I

(to Appendix III to Recommendation G.726
and to Appendix II to Recommendation G.727)

The discrete function $W[I(k)]$ defines a quantizer multiplier whose value depends on the code word I for Recommendation G.726 and on the core bits I_c for Recommendation G.727 as shown:

For 2-core-bit operation (1 sign bit), the discrete function $W[I_c(k)]$ is defined as follows:

$\lvert I_c(k)\rvert$	1	0
$W[I_c(k)]$	27.44	–1.38

For 3-core-bit operation (1 sign bit), the discrete function $W[I_c(k)]$ is defined as follows:

$\lvert I_c(k)\rvert$	3	2	1	0
$W[I_c(k)]$	36.38	8.56	1.88	–0.25

For 4-core-bit (1 sign bit) operation, the discrete function $W[I_c(k)]$ is defined as follows:

$\lvert I_c(k)\rvert$	7	6	5	4	3	2	1	0
$W[I_c(k)]$	69.25	21.25	11.50	6.13	3.13	1.69	0.25	–0.75

For 40 kbit/s ADPCM, the discrete function $W(I)$ is defined to optimize performance with a variety of voiceband data signals, in particular for V.29 modems:

$\|I(k)\|$	15	14	13	12	11	10	9	8
$W[I(k)]$	43.50	33.06	27.50	22.38	17.50	13.69	11.19	8.81

$\|I(k)\|$	7	6	5	4	3	2	1	0
$W[I(k)]$	6.25	3.63	2.56	2.50	2.44	1.50	0.88	0.88

For the 32 kbit/s optimized ADPCM algorithm of COM XVIII-102, the discrete function $W(I)$ is defined:

$\|I(k)\|$	15	14	13	12	11	10	9	8
$W[I(k)]$	245.94	136.88	76.98	55.54	41.73	28.61	15.14	8.81

$\|I(k)\|$	7	6	5	4	3	2	1	0
$W[I(k)]$	11.55	8.65	5.71	2.58	1.28	0.04	–1.59	–2.40

For 32 kbit/s ADPCM, the discrete function $W(I)$ is defined as follows (infinite precision values):

$\|I(k)\|$	7	6	5	4	3	2	1	0
$W[I(k)]$	70.13	22.19	12.38	7.00	4.00	2.56	1.13	–0.75

For 24 kbit/s ADPCM, the discrete function $W(I)$ is defined as follows (infinite precision values):

$\|I(k)\|$	3	2	1	0
$W[I(k)]$	36.38	8.56	1.88	–0.25

For 16 kbit/s ADPCM, the discrete function $W(I)$ is defined as follows (infinite precision values):

$\|I(k)\|$	1	0
$W[I(k)]$	27.44	–1.38

$F[I(k)]$ is defined using $I(k)$ for the fixed ADPCM algorithm G.726, COM XVIII-101, COM XVIII-102 and using $I_c(k)$ for embedded ADPCM of Recommendation G.727 by:

$\|I(k)\|$	1	0
$F[I(k)]$	7	0

for 2-core-bit (1 sign bit) or 16 kbit/s operation; or

$\|I(k)\|$	3	2	1	0
$F[I(k)]$	7	2	1	0

for 3-core-bit (1 sign bit) or 24 kbit/s operation; or

$\|I(k)\|$	7	6	5	4	3	2	1	0
$F[I(k)]$	7	3	1	1	1	0	0	0

for 4-core-bit (1 sign bit) and 32 kbit/s operation.

For 40 kbit/s ADPCM, $F[I(k)]$ is defined by:

$\|I(k)\|$	15	14	13	12	11	10	9	8
$F[I(k)]$	6	6	5	4	3	2	1	1

$\|I(k)\|$	7	6	5	4	3	2	1	0
$F[I(k)]$	1	1	1	0	0	0	0	0

The function $F[I(k)]$ for COM XVIII-102 is not known.

The functions $W_s[I(k)]$ and $W_d[I(k)]$ of the algorithm of COM XVIII-101 are:

$\|I(k)\|$	15	14	13	12	11	10	9	8
$W_s[I(k)]$	1.2695	0.4883	0.1294	–0.0708	–0.0977	–0.1123	–0.1709	–0.2495
$W_d[I(k)]$	1.0825	0.0	0.0	0.0	0.0	0.0	0.0503	0.1235

$\|I(k)\|$	7	6	5	4	3	2	1	0
$W_s[I(k)]$	1.2695	0.4883	0.1294	–0.0708	–0.0977	–0.1123	–0.1709	–0.2495
$W_d[I(k)]$	1.0825	0.0	0.0	0.0	0.0	0.0	0.0503	0.1235

The function $F[I(k)]$ of the algorithm of COM XVIII-101 is:

$\|I(k)\|$	15	14	13	12	11	10	9	8
$F[I(k)]$	7	3	1	1	1	0	0	0

$\|I(k)\|$	7	6	5	4	3	2	1	0
$F[I(k)]$	7	3	1	1	1	0	0	0

References

[1] DAUMER (W. R.) and SPARRELL (D.): FSK enhancements for proposed T1Y1 ADPCM algorithm, *Doc. No. T1Y1.2/86-006*, January 1986.

[2] AT&T *G.721 ADPCM algorithm modification and test results*, CCITT Contribution No. D.725/XVIII, July 1986.

[3] AT&T *G.721 algorithmic modification and test data for 24 and 40 kbit/s operation*, CCITT Contribution to SG XVIII, July 1987.

[4] AT&T *Performance of various ADPCM algorithms with very high speed modems*, ITU-T Contribution No. COM 15-D.9, Q.4/15, 9/15, 13/15, September 1993.

[5] BANERJEE (S.) and BERTOCCI (G.): Testing implementations of the 32 kbit/s CCITT and ANSI ADPCM algorithm, *ICC'87*, Vol. 3: 1487-1490, Seattle, Washington, June 7-10, 1987.

[6] BENVENUTO (N.), BERTOCCI (G.), DAUMER (W. R.) and SPARRELL (D. K.): The 32-kbit/s ADPCM coding standard, *AT&T Tech. J.*, Vol. 65 (5): 12-22, Sept./Oct., 1986.

[7] BENVENUTO (N.) and BERTOCCI (G.): Prevention of quantizer overloading in the 32 kbit/s ADPCM algorithm, *ICC Proceedings*, Vol. 3: 1483-1486, June 7-10, 1987.

[8] BENVENUTO (N.) and BERTOCCI (G.): Enhancement in the 32 kbit/s ADPCM algorithm for FSK voice-band data signals, *IEEE Trans. Commun.*, Vol. COM-36: 626-628, May 1988.

[9] BENVENUTO (N.) and DAUMER (W. R.): Two approaches for waveform coding of 9.6 kbit/s voiceband data signals through 32 kbit/s ADPCM, *IEEE Trans. Commun.*, Vol. COM-36: 382-385, March 1988.

[10] BIALLY (T.), GOLD (B.) and SENEFF (S.): A technique for adaptive voice flow control in integrated packet networks, *IEEE Trans. Commun.*, Vol. COM-28: 325-333, March 1980.

[11] BOWKER (D. O.) and DVORAK (C. A.): Speech Transmission Quality of Wideband Packet Technology, *GLOBECOM'87*, Vol.3: 1887-1889, Tokyo, Japan, November 15-18, 1977.

[12] BOX (G. E. P.) and JENKINS (G. M.): Time Series Analysis Forecasting and Control, Holden-Day, San Francisco, 1976.

[13] CCITT SG XVIII, Annex 1 to the Report of the Meeting of Working Party XVIII/2 (Speech Processing), Geneva 21, 22 and 25 November 1983, COM XVIII-R 28, December 1983.

[14] CCITT Recommendation P.70 *Modulated Noise Reference Unit*, VIIIth Plenary Assembly, Malaga-Torremolinos, 8-19 October 1984.

[15] CCITT SG XVIII, Question 25,31, Report to Working Party XVIII/8, July 1987.

[16] CCITT Recommendation G.721 *32 kbit/s adaptive differential pulse code modulation (ADPCM)*, IXth Plenary Assembly, Melbourne, 14-25 November 1988.

[17] CCITT Recommendation G.723 *Extensions of Recommendation G.721 to 24 and 40 kbit/s for DCME Application*, IXth Plenary Assembly, Melbourne, 14-25 November 1988.

[18] CCITT Plenary Assembly, 1988 *Questions allocated to Study Group XV (Transmission systems and equipment)*, Question 24/XV, Speech Packetization, Appendix 3 to Annex 1.

[19] CCITT Recommendation G.726 *40, 32, 24, 16 kbit/s Adaptive Embedded ADPCM*, July 1990.

[20] CCITT Recommendation G.727 *5-, 4-, 3-, 2-Bit/Sample Embedded ADPCM*, July 1990.

[21] CCITT Recommendation G.764 *Packet Voice Protocol*, July 1990.

[22] CCITT SG XV Special Rapporteur for Q.24/XV (D. Sparrell) *Objective measurements of the fixed rate and embedded ADPCM algorithms using the pseudorandom noise method of G.712*, CCITT SG XV Temporary Documents 21 (XV/2) and 49 (XV/3), 18 February-1 March 1991.

[23] COMSAT, *Testing/Phase I: Results of an objective performance evaluation of ADPCM coders for DCME voiceband data transparency*, CCITT Contribution to 25/XVIII DCME Ad Hoc Group No. 74, January 1987.

[24] COMSAT, *Testing/Phase II: Results of an objective performance evaluation of ADPCM coders for DCME voiceband data transparency*, CCITT Contribution to 25/XVIII DCME Ad Hoc Group No. 98, April 1987.

[25] COMSAT, *Evaluation of ADPCM Coders for Digital Circuit Multiplication Equipment*, Final Report on Task TTA-22, CTD 87/125, August 1987, Volumes 1-7. Vol. l: Performance Curves (for all algorithms), Vol. 2: PCM Reference, Vol. 3: OKI/KDD 32 kbit/s ADPCM, Vol. 4: ECI 32 kbit/s ADPCM, Vol. 5: OKI/KDD 40 kbit/s ADPCM, Vol. 6: CNET 40 kbit/s ADPCM, Vol. 7: 4800 bit/s modem (for all algorithms). Available from COMSAT Laboratories, Records Dept., 22300 Comsat Dr., Clarksburg, MD 20871, USA.

[26] DAUMER (W. R.), MERMELSTEIN (P.), MAITRE (X.) and TOKIZAWA (I.): Overview of the ADPCM coding algorithm, *GLOBECOM'84*, Vol. 2: 774-777, Atlanta, GA, November 26-29, 1984.

[27] DÈCINA (M.) and MODENA (G.): CCITT standards on digital speech processing, *IEEE J. Sel. Areas. Commun*., Vol. 6 (2): 227-234, Feb. 1988.

[28] DIMOLITSAS (S.) and BHASKAR (U.): Hurwitz stability analysis of an ADPCM system, *COMSAT Technical Review*, Vol. 17 (1): 107-125, Spring 1987.

[29] DIMOLITSAS (S.), CORCORAN (F. L.), ONUFRY (M.) and SUYDERHOUD (H. G.): Evaluation of ADPCM coders for digital circuit multiplication equipment, *COMSAT Technical Review*, Vol. 17 (2): 325-345, Fall 1987.

[30] DIMOLITSAS (S.): Objective speech distortion measures and their relevance to speech quality assessments, *IEE Proceedings*, Vol. 136, Pt. 1. No. 5: 317-324, Oct. 1989.

[31] HOSODA (H.), NOGUCHI (O.) and YATSUZUKA (Y.): A 32 kbit/s ADPCM algorithm having a high performance for both voice and 9.6 kbit/s modems, *J. Select. Areas Commun.*, Vol. 6, No. 2: 262-272, Feb. 1988.

[32] IEEE Recommended Practice, *Speech Quality Measurements*, No. 297, 1969.

[33] Israel, Voiceband data optimized algorithm- Detailed description, Document No. COM XVIII/92 CCITT Study Group XVIII- Ad Hoc Group DCME-Base Solution for Voiceband Data Transparency, Stockholm, April 1987.

[34] JAIN (V. K.), SKRZYPKOWIAK (S. S.) and HEATHCOCK (R. B.): Full duplex speech and data coder: algorithm enhancement test bed, *GLOBECOM'88*, Vol. 3: 1409-1413, Hollywood, Florida, Nov. 28-Dec. 1, 1988.

[35] JAYANT (N. S.) and NOLL (P.): Digital Coding of Waveforms: Principles and Applications to Speech and Video, Prentice-Hall, Englewood Cliffs, New Jersey, 1984.

[36] Kokusai Denshin Denwa Co. Ltd. (KDD), *High-level description of advanced 32 kbit/s ADPCM*, Document Ad-Hoc No. 27 CCITT Study Group XVIII/WP8- Ad Hoc Group "DCME-Base Solution for Voiceband Data Transparency," September 22-24, 1986.

[37] Kokusai Denshin Denwa Co. Ltd. (KDD) and OKI Electric Industry Co., Ltd., *Computational details of advanced 32 kbit/s ADPCM*, Document Ad-Hoc No. 86 CCITT Study Group XVIII/WP8- Ad Hoc Group "DCME-Base Solution for Voiceband Data Transparency," January 5-9, 1987.

[38] MEDAUGH (R. S.): Digital code converter, US Patent 4,475,213, October 2, 1984.

[39] MILLAR (D. J.) and MERMELSTEIN (P.): Prevention of predictor mistracking in ADPCM coders, *ICC'84*, Vol. 3: 1508-1512, Amsterdam, Netherlands, June 1984.

[40] MILLS (J. D.) and MELSA (J. L.): Performance improvements using synchronous code correction in ADPCM/PCM tandem connections, *IEEE Trans. Commun*.: Vol. 38, No. 5, May 1990, pp. 561-564.

[41] MILLS (J. D.), MELSA (J. L.) and MAGILL (R. B.): T1.EMB single encoding distortion using synchronous code correction, *ICC'90*, Vol. 3: 1019-1024, Atlanta, Georgia, April 1990.

[42] Mitsubishi Electric Corporation, *The problems on objective test results of Recommendations G.721/G.723 ADPCM transcoder*, CCITT contribution to SG XV, July 1990.

[43] Mitsubishi Electric Corporation, *The characteristics of the turbulences appeared in objective test and voice signal transmission of Rec. G.726*, CCITT contribution D.301/XV, July 1991.

[44] NISHITANI (T.): Tandem transcoding without distortion accumulation, *IEEE Trans. Commun*.: Vol. COM-34, No. 3, March 1986, pp. 278-284.

[45] NISHITANI (T.), AIKOH (S.), ARASCKI (T.), OZAWA (K.) and MARITA (R.): A 32 kbit/s toll quality ADPCM codec using a single chip processor, *IEEE Int. Conf. Acoust. Speech. Sig. Process. (ICASSP'82)*: Vol. 2: 960-963, Paris, France, 1982.

[46] PASCAL (D.) and COMBESCURE (P.): Evaluation de la qualité de la transmission vocale, *L'Echo des Recherches*, No. 132: 31-40, 1988.

[47] PETR (D. W.): 32 kbit/s ADPCM-DLQ for network applications, *GLOBECOM'82*, Vol. 1: 239-243, Miami, Florida, Nov. 29-Dec. 2, 1982.

[48] SHERIF (M. H.), BOWKER (D. O.), BERTOCCI (G.), ORFORD (B. A.) and MARIANO (G. A.): Overview of CCITT/ANSI embedded ADPCM algorithms, *ICC'90*, Vol. 3, pp. 1014-1018, Atlanta, GA, April 15-19, 1990.

[49] SHERIF (M. H.), BOWKER (D. O.), BERTOCCI (G.), ORFORD (B. A.) and MARIANO (G. A.): Overview and performance of CCITT/ANSI embedded ADPCM algorithms, accepted for publication in *IEEE Trans. Commun*.

[51] SUZUKI (J.) and TAKA (M.): Missing packet recovery techniques for low-bit rate coded speech, *J. Select. Areas Commun*., Vol. 7, No. 5: 707-717, June 1989.

[52] WINER (B. J.): Statistical Principles in Experimental Design (Second Edition), *McGraw-Hill*, New York, 1971.

[53] SHAPIRO Zeev: Adaptive differential pulse code modulation (ADPCM) systems, *US Patent 4,860,313*, Aug. 22, 1989.