

On QAM Speech Transmission Schemes for Microcellular Mobile PCNs

Lajos Hanzo, William Webb (*), Raymond Steele (*)

Dept. of Electr. & Comp. Sc., Univ. of Southampton
U.K. SO9 5NH

Redwan A. Salami

Communications Research Center
University of Sherbrooke, Canada, J1K 2R1

Abstract. On the way towards third generation personal communication networks (PCN) the deployment of 16-level quadrature amplitude modulation (QAM) schemes is investigated via flat microcellular Rayleigh-fading mobile radio channels having mild co-channel interference with and without diversity. Our proposed differentially coded non-coherent 16-QAM scheme requires no pilot symbol assistance, no automatic gain control (AGC) and no carrier recovery. High speech quality associated with a mean opinion score of about four ($MOS \approx 4.0$) is ensured at a signalling rate of 6.2 kbd for channel signal to noise ratios (SNR) above 20 - 25 dB and signal to interference ratios (SIR) above 25 - 30 dB by the 13.4 kbit/s medium complexity regular pulse excited (RPE) speech codec, when carefully matched twin-class low complexity binary Bose-Chaudhuri-Hocquenghem (BCH) error correction coding and 16-QAM combined with second order diversity are deployed. The overall speech delay is about 40 ms and the user bandwidth is 9.3 kHz, which allows us to accommodate 21 users in a 200 kHz channel slot.

1. INTRODUCTION

A few years after the definition of the Pan-European digital mobile radio system known as GSM [1] other second generation mobile systems such as the American IS-54 [2] and the Japanese system [3] followed. The key features of these systems are summarised in Table 1, including the half-rate GSM system.

While the GSM system is a complex scheme for hostile, dispersive land mobile channels, the other two systems were contrived for more benign channels and represent a different design philosophy. Indeed, the IS-54 and the Japanese digital mobile systems provide a seamless conceptual transition towards the third generation personal communications network (PCN) concept, where the friendly micro- or pico-cellular environment guarantees high signal-to-noise ratios (SNR) and low signal dispersion. The typically high SNR can be exploit-

ed by deploying 16-level quadrature amplitude modulation (16-QAM), which lowers the signalling rate by a factor of four and hence quadruples the signalling interval's duration. Therefore the relative effects of signal dispersion are mitigated and the need for channel equalization can be removed.

The paper is organised as follows. Section 2 gives a brief overview of QAM research for both Gaussian and Rayleigh fading channels and highlights the key features of two previously proposed QAM-based mobile communication schemes for PCN in contrast to the presently suggested system. Section 3 highlights our proposed non-coherent differentially coded 16-QAM modem. Section 4 concentrates on the RPE speech codec and after a theoretical treatise it provides a rigorous bit sensitivity analysis based on a sensitivity measure derived from the amalgamation of segmental SNR ($SEG-SNR$) and cepstral distance (CD) measures. Section 5 deals with aspects of channel coding, while Section 6 describes the combined transceiver scheme proposed. Lastly, Section 7 summarises our results and provides some discussions.

(*) Also with Multiple Access Communications Ltd, Enterprise Rd., Chilworth Research Centre, Southampton SO1 7NS, U.K.

Table 1 - Comparison of Second Generation Cellular Systems.

	GSM Half-Rate	GSM Full-Rate	IS-54	Japanese
TDMA Channel Bandwidth (kHz)	200	200	30	25
Modulation	GMSK	GMSK	$\pi/4$ -DQPSK	$\pi/4$ -DQPSK
TDMA Burst-Rate (kbit/s)	271	271	48.6	42
Modem Bandwidth Efficiency (bit/s/Hz)	1.35	1.35	1.62	1.68
Number of Users/TDMA Chan.	16	8	3	3
User Bandw.(kHz)	12.5	25	10	8.3
Speech + FEC Rate (kbit/s)	11.4	22.8	13	11.2
Equalizer	Viterbi	Viterbi	Arbitrary	None
Diversity	No	No	No	Yes
Sys. Complexity	Highest	High	Medium	Lowest
Required SNR	Low	Low	Medium	Medium

2. QAM-BASED MOBILE RADIO TRANSCEIVERS

2.1. QAM modems

QAM was first proposed in the early sixties [4, 5] for Gaussian channels, and the ubiquitous "maximum minimum distance" square constellation was first suggested in reference [6]. An excellent review paper was published by Forney et al. [7] in the mid-eighties, covering a variety of constellations suitable for Gaussian channels. This research culminated in the definition of the CCITT telephone circuit modem standards V.29 - V.33 based on various QAM constellations ranging from uncoded 16-QAM to trellis coded (TC) 128-QAM.

Multilevel modem research for narrow- and wideband fading mobile channels has been fuelled by the ever-increasing spectral efficiency demand for PCNs [8 - 20]. Despite all the disadvantages in terms of power-inefficient class A or A B linear amplification and high signal-to-interference (SIR) requirements 16-QAM becomes attractive for friendly high-SNR microcellular mobile channels. In this environment the transmitted power is low and the power consumption of the low-efficiency class-A amplifier [21, 22] is less critical than that of the "power-thirsty" digital speech and channel codecs, equalisers, etc.

In fact, deploying QAM alleviates channel dispersion, thereby removing the need for an equaliser, while the higher SNR demand can be compensated by diversity reception.

If linear amplification must be avoided, Feher's non-linear amplification (NLA) concept can be used [8], where phasors with different magnitudes can be amplified by separate non-linear amplifiers, whose signals are then combined for transmission. Instead of utilising linear-phase Nyquist filters, the concept of non-linear filtering (NLF) joining time-domain signal transitions with a quarter of a period duration raised-cosine segment can also be used. Out-of-band emissions due to class A B amplifier non-linearities and hence adjacent channel interferences can be reduced by some 15 - 20 dB using Stapleton's adaptive predistorter [23 - 25] and a class-A B amplifier with 6 dB back-off, by adjusting the predistorter's adaptive coefficients using the complex convolution of the predistorter's input signal and the amplifier's output signal. Further aspects of linearised power amplifier design are considered in references [26] and [27].

In our previous QAM research for mobile channels reported in references [18 - 20, 28] we used the "maximum minimum distance" square-shaped 16-QAM and 64-QAM constellations proposed by Cahn [5] that were deemed optimum for Gaussian channels. The square-

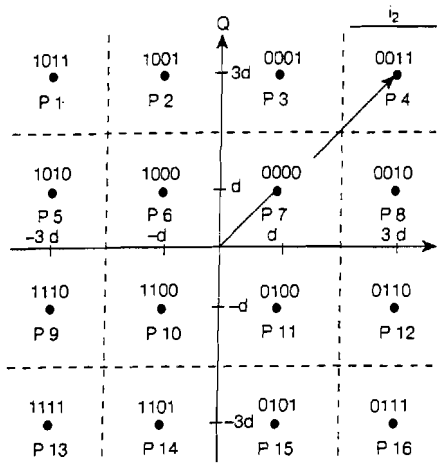


Fig. 1 - 16-level "square" QAM constellation.

shaped 16-QAM constellation displayed in Fig. 1 maps four adjacent input bits onto one of its phasors using Gray-encoding, ensuring that the encoding of nearest neighbour phasor points differs only in a single bit position. Therefore the most likely error events yield a single bit error. The mapping of the in-phase bits i_1 and i_2 , as well as quadrature bits q_1 and q_2 onto the phasor points is interleaved in the order of (i_1, q_1, i_2, q_2) . Observe from the constellation of Fig. 1 that while the bits (i_2, q_2) referred to as class 2 bits are always at a "protection distance" of d from the closest decision boundary, the class 1 bits (i_1, q_1) have a protection distance of d and $3d$ for 50% of the time, respectively. In reference [18] we

derived the class 1 ($C1$) and class 2 ($C2$) bit error probabilities $P_{1R}(\Gamma)$ and $P_{2R}(\Gamma)$ under the assumption of Rayleigh fading in case of an average channel signal-to-noise ratio (SNR) of Γ , defined as the ratio of the average faded signal power to the Gaussian noise power. These bit error probabilities were given by:

$$P_{1R}(\Gamma) = \frac{1}{2\Gamma} \int_0^\infty \left[Q\left(\sqrt{\frac{\gamma}{5}}\right) + Q\left(3\sqrt{\frac{\gamma}{5}}\right) \right] \cdot e^{-\gamma/\Gamma} d\gamma \quad (1)$$

$$P_{2R}(\Gamma) = \frac{1}{2\Gamma} \int_0^\infty Q\left[\left(\frac{2\bar{\alpha}}{\alpha} - 1\right)\sqrt{\frac{\gamma}{5}}\right] e^{-\gamma/\Gamma} d\gamma + \frac{1}{2\Gamma} \int_0^\infty Q\left[\left(\frac{3\bar{\alpha}}{\alpha} - 2\right)\sqrt{\frac{\gamma}{5}}\right] e^{-\gamma/\Gamma} d\gamma \quad (2)$$

where α represents the Rayleigh distributed fading envelope's instantaneous attenuation, $\bar{\alpha}$ is its expected value, γ is the instantaneous SNR value and the Gaussian $Q[\]$ function is defined as:

$$Q(y) = \int_y^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \quad (3)$$

In case of square 64-QAM there are three subchannels referred to as $C1$, $C2$ and $C3$ subchannels and the corresponding bit error probabilities are [19]:

$$P_{C1} = \frac{1}{4\Gamma} \int_0^\infty \left[Q\left(\sqrt{\frac{\gamma}{21}}\right) + Q\left(3\sqrt{\frac{\gamma}{21}}\right) + Q\left(5\sqrt{\frac{\gamma}{21}}\right) + Q\left(7\sqrt{\frac{\gamma}{21}}\right) \right] e^{-\gamma/\Gamma} d\gamma \quad (4)$$

$$P_{C2}(\Gamma) = \frac{1}{4\Gamma} \left\{ \int_0^\infty Q\left[\left(\frac{4\bar{\alpha}}{\alpha} - 1\right)\sqrt{\frac{\gamma}{21}}\right] e^{-\gamma/\Gamma} d\gamma + \int_0^\infty Q\left[\left(\frac{4\bar{\alpha}}{\alpha} - 3\right)\sqrt{\frac{\gamma}{21}}\right] e^{-\gamma/\Gamma} d\gamma + \int_0^\infty Q\left[\left(\frac{4\bar{\alpha}}{\alpha} - 5\right)\sqrt{\frac{\gamma}{21}}\right] e^{-\gamma/\Gamma} d\gamma + \int_0^\infty Q\left[\left(\frac{4\bar{\alpha}}{\alpha} - 7\right)\sqrt{\frac{\gamma}{21}}\right] e^{-\gamma/\Gamma} d\gamma \right\} \quad (5)$$

$$P_{C3}(\Gamma) = \frac{1}{4\Gamma} \left\{ \int_0^\infty \left[Q\left[\left(\frac{6\bar{\alpha}}{\alpha} - 3\right)\sqrt{\frac{\gamma}{21}}\right] + Q\left[\left(3 - \frac{2\bar{\alpha}}{\alpha}\right)\sqrt{\frac{\gamma}{21}}\right] \right] e^{-\gamma/\Gamma} d\gamma + \int_0^\infty \left[Q\left[\left(\frac{2\bar{\alpha}}{\alpha} - 1\right)\sqrt{\frac{\gamma}{21}}\right] - Q\left[\left(\frac{6\bar{\alpha}}{\alpha} - 1\right)\sqrt{\frac{\gamma}{21}}\right] \right] e^{-\gamma/\Gamma} d\gamma + \int_0^\infty \left[Q\left[\left(\frac{6\bar{\alpha}}{\alpha} - 5\right)\sqrt{\frac{\gamma}{21}}\right] + Q\left[\left(5 - \frac{2\bar{\alpha}}{\alpha} - 1\right)\sqrt{\frac{\gamma}{21}}\right] \right] e^{-\gamma/\Gamma} d\gamma + \int_0^\infty \left[Q\left[\left(-7 - \frac{6\bar{\alpha}}{\alpha}\right)\sqrt{\frac{\gamma}{21}}\right] - Q\left[\left(7 - \frac{2\bar{\alpha}}{\alpha}\right)\sqrt{\frac{\gamma}{21}}\right] \right] e^{-\gamma/\Gamma} d\gamma \right\} \quad (6)$$

The corresponding BER versus channel SNR curves for coherent detection of 16-QAM and 64-QAM via Rayleigh channels were verified by simulations and can be found in references [18] and [19], respectively. While the $C1$ bit error rates (BER) evaluated from Eqs. 1 and 4 for 16-QAM and 64-QAM were sufficiently low to be further reduced by forward error correction (FEC) coding, the excessively high $C2$ and $C3$ BERs could not be efficiently reduced by practical FEC codecs, since the FEC code was usually overloaded by bursty errors. In reference [18] we opted for automatic gain control (AGC) techniques in order to further reduce the $C2$ residual BER, while in references [19, 28] we deployed pilot symbol assisted modulation (PSAM) [29].

2.2. QAM-based speech systems

The BER differences of the $C1$ and $C2$ subchannels were exploited in references [18, 20] and [28] in order to convey the more vulnerable speech or video bits via the higher integrity $C1$ route, while the less sensitive bits were transmitted using the FEC-protected higher BER $C2$ subchannel.

Under the assumption of perfectly coherent demodulation, in reference [18] we used a variety of medium complexity schemes based on a Bose-Chaudhuri-Hoc-

squenghem (BCH) coded 16 kbit/s subband codec (SBC) and our square-constellation 16-QAM modem. The overall signalling rate for the proposed schemes was between 4 and 7 kbd and we favoured the 5 kbd diversity assisted system, yielding a speech quality comparable to a mean opinion score (MOS) of about 3.5, if the channel SNR was in excess of about 24 dB. The features of this scheme designated by System 1 are summarised in the second column of Table 2.

In reference [19] we proposed an extremely bandwidth efficient system using a 4.8 kbit/s code excited linear predictive (CELP) codec and a perfectly coherent 64-QAM modem, which we refer to as System 2. Features of this scheme are summarised in column 3 of Table 2. Observe that this highly complex system offers lower speech quality at higher minimum channel SNR and about five times lower bandwidth requirement than that of System 1. The speech quality of both systems was lower than that of the 13 kbit/s regular pulse excited (RPE) GSM speech codec. This is because the SBC is essentially a low-complexity open-loop codec, while the 4.8 kbit/s CELP codec suffers from its extremely low bitrate, in spite of being an intelligent perceptually motivated closed-loop scheme.

In this contribution we contrive a compromise scheme based on an RPE speech codec [30], which offers high speech quality, associated with an MOS of four, while maintains low complexity due to a direct excitation computation procedure to be described in Section 4. The features of System 3 based on the RPE codec

Table 2 - QAM-based speech systems.

	System 1	System 2	System 3
Speech Codec	16 kbit/s SBC	4.8 kbit/s CELP	13.4 kbit/s RPE
FEC	Binary BCH	Binary BCH	Binary BCH
FEC coded rate	20 kbit/s	6.6 kbit/s	24.8 kbit/s
Constellation	Square 16-QAM	Square 64-QAM	Star 16-QAM
Demodulation	Coherent	Coherent	Non-coherent
Diversity	Yes	Yes	Yes
Equalizer	No	No	No
Complexity	Medium	High	Medium
Min. Channel SNR	24 dB	30 dB	20 dB
User Bandwidth	10 kHz	2.2 kHz	9.3 kHz
Estimated MOS	3.5	2.5	4.0
Reference	[18]	[19]	

and a novel differentially coded non-coherent 16-QAM modem scheme to be described in Section 3 are also given in Table 2 for the sake of comparison, although its parameters will be derived later. The most important

and a novel differentially coded non-coherent 16-QAM modem scheme to be described in Section 3 are also given in Table 2 for the sake of comparison, although its parameters will be derived later. The most important

characteristic of our proposed modem in System 3 is that its operation does not require pilot symbols. AGC or coherent demodulation, which significantly reduces the receiver's complexity and increases its robustness against false locking problems. Let us initially focus our attention on our proposed 16-QAM modem schemes.

3. QAM MODEMS FOR PCN

3.1. Coherent square 16-QAM

As we have seen in Table 2, our previously proposed 16-QAM modems used the square-shaped 16-QAM constellation seen in Fig. 1, where it is essential to be able to separate the information modulated onto the *I* and *Q* quadrature carriers. This can only be achieved, if a quadrature carrier source perfectly coherent with that of the transmitter is available at the receiver. If there is a frequency difference between the transmitter and receiver, the demodulator's reference coordinate axis will be rotating with respect to the transmitter's according to the difference frequency, yielding catastrophic communication break-down. Albeit a phase error yields a constant coordinate axis rotation offset, the demodulated BER will be equally high. Hence the performance of the carrier recovery is crucial as regards to the overall modem BER.

Excellent treatises are available in the literature for carrier recovery [31 - 33], but all known methods require an accurate AGC in order to prevent the phasor constellation from collapsing in the centre of the coordinate axis. Furthermore, most carrier recovery schemes have false locking problems in case of the maximum minimum distance square 16-QAM. Specifically, rotating the coordinate axis seen in Fig. 1 it can be shown that false locking can appear at rotations $(26 + n 90^\circ)$ and $(52 + n 90^\circ)$, where $n = 0, 1, 2, 3$ [34]. These false locking problems can be mitigated using only those phasors to update the voltage controlled oscillator (VCO), for which the ratio of the *I* and *Q* components fulfils one of the following conditions:

$$\begin{aligned} (1 - \Delta) < I/Q < (1 + \Delta) \\ (3 - \Delta) < I/Q < (3 + \Delta) \\ (3 - \Delta) < Q/I < (3 + \Delta) \end{aligned} \tag{7}$$

where Δ is a small quantity. Clearly, the *I/Q* ratio is not allowed to significantly deviate from the ideal values of 1 or 3, if it is to be used for VCO updates. Hence for this carrier recovery method to work reliably unacceptably high *SNR* values ought to be maintained, which cannot be guaranteed in the mobile environment.

A better approach is to find a 16-QAM constellation, which eventually has a lower minimum distance amongst constellation points, but does not require coherent demodulation.

3.2. Non-coherent star 16-QAM

The key to non-coherent demodulation is to find a rotationally symmetric 16-QAM constellation, where all constellation points are rotated by the same amount. Such a constellation is given in Fig. 2. Although this phasor diagram suggests that there are eight false locking positions at $(n 45^\circ) n = 0...7$ positions, using differential coding removes this ambiguity. In Gaussian channels differential coding has a 3 dB *SNR* disadvantage, but via Rayleigh fading channels this becomes more negligible. Furthermore, if coherent demodulation is required, the previously mentioned VCO-adjustment algorithm becomes fairly simple. This is because the *I/Q*-ratio is unity for all constellation points and there is no need to introduce the deadzone Δ , as we have done for square 16-QAM. Hence this carrier recovery can operate at lower *SNR* values than that of the square constellation and does not require an AGC arrangement.

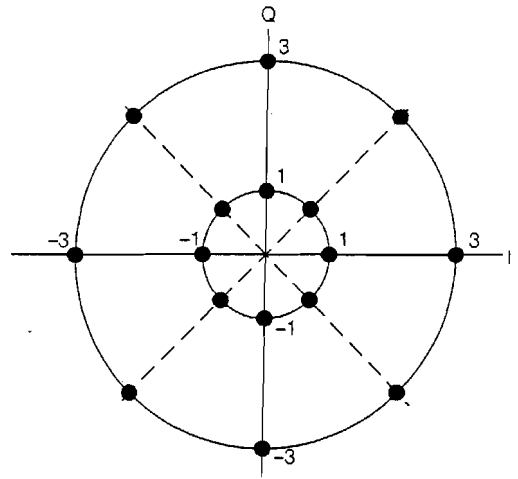


Fig. 2 - 16-level "star" QAM constellation.

A disadvantage of the proposed star 16-QAM (16-St-QAM) constellation is its lower average energy. While square 16-QAM had an average phasor energy of $10 d^2$, the 16-StQAM halves this value to $5 d^2$. This implies a 3 dB disadvantage over Gaussian channels, a price worth sacrificing for high robustness and lower complexity over Rayleigh channels.

Our differential encoder obeys the following rules. The first bit b_1 of a four-bit symbol is differentially encoded onto the phasor magnitude, yielding a ring-swap for an input logical one and maintaining the current magnitude, i.e., ring for $b_1 = 0$. Bits (b_2, b_3, b_4) are then differentially Gray-coded onto the phasors of the particular ring pinpointed by b_1 . Accordingly, $(b_2, b_3, b_4) = (0, 0, 0)$ implies no phase change, $(0, 0, 1)$ a change of 45° , $(0, 1, 1)$ a change of 90° , etc.

The non-coherent differential 16-StQAM demodulation is equally straightforward, having decision boundaries at a concentric ring of radius 2 and at phase rotations of $(22.5^\circ + n 45^\circ) n = 0...7$. Assuming received phasors of P_i and P_{i+1} at consecutive sampling instants

of t and $t + 1$, respectively, bit b_1 is inferred by evaluating the condition:

$$|P_{t+1}|/|P_t| \geq 2 \quad (8)$$

If this condition is met, $b_1 = 1$ is assigned, otherwise $b_1 = 0$ is demodulated. Bits (b_2, b_3, b_4) are then recovered by computing the phase difference

$$\Delta\Theta = (\Theta_{t+1} - \Theta_t) \bmod 2\pi \quad (9)$$

and comparing it against the decision boundaries ($22.5^\circ + n \cdot 45^\circ$) $n = 0 \dots 7$. Having decided which rotation interval the received phase difference $\Delta\Theta$ belongs to, Gray-decoding delivers the bits (b_2, b_3, b_4) .

From our previous discourse it is plausible that the less dramatic the fading envelope and phase trajectory fluctuation between adjacent signalling instants, the better this differential scheme works. This implies that lower vehicular speeds are preferred by this arrangement, if the signalling rate is fixed. Alternatively, for a fixed vehicular speed higher signalling rates are favourable, since the relative amplitude and phase changes introduced by the fading channel between adjacent information symbols are less drastic.

These observations led us to a further improvement of our 16-StQAM modem. Namely, taking several samplings during any signalling interval allows us to detect the envelope and phase changes between its beginning and end. These samples can then be used to predict and remove the effects of the envelope and phase changes due to the channel, retaining only those that are due to modulation. This can be done by detecting the phasor magnitude and phase changes between the last sample $P_{t,N}$ from the previous signalling interval and the first of the present one, $P_{t+1,1}$. The phasor change is now mainly due to modulation and only fractionally caused by the channel between observations $P_{t,N}$ and $P_{t+1,1}$.

The above procedure can be further improved using linear magnitude and phase trajectory interpolation in order to eliminate the effect of channel fading between observations $P_{t,N}$ and $P_{t+1,1}$. This can be achieved with the aid of deriving the extrapolated estimate

$$P_{t+1,\text{ext}} = P_{t+1,1} - \frac{P_{t+1,N} - P_{t+1,1}}{N} \quad (10)$$

by linearly extrapolating from the observation $P_{t+1,1}$ using the gradient $(P_{t+1,N} - P_{t+1,1})$ inferred during the sampling interval $(t + 1)$, in order to identify and eliminate the phasor change due to fading during the whole sampling interval. The oversampled observations for $n = 1 \dots 5$ and the extrapolated estimate $Z_{t+1,\text{ext}}$ are shown in Fig. 3. Having eliminated the effect of fading, we are in the position to infer the magnitude and phase change encountered purely due to modulation. Note that in our discourse on oversampling so far we have assumed rectangular time-domain modulation waveforms with no pulse-shaping.

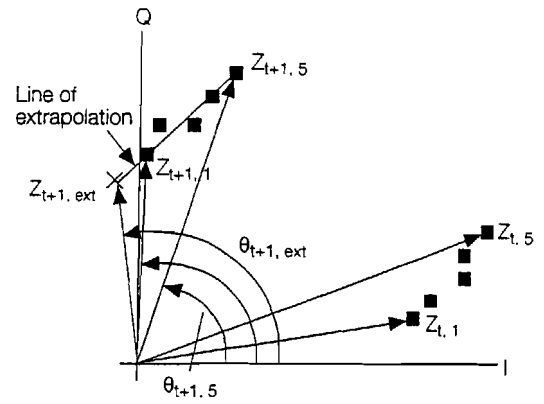


Fig. 3 - Fade-tracking oversampling in 16-StQAM.

Incorporating pulse-shaping into the oversampling scheme requires us to include its inverse function in the extrapolating linear prediction, since the effects of fading have to be estimated without any other time-domain shaping operation, which complicates its implementation. In our candidate schemes we have experimented with both conventional square-root raised-cosine Nyquist filters using roll-off factors of 0.5 and 1, as well as non-linear filtering (NLF) using the previously mentioned time-domain quarter raised cosine segments. Due to its implementational simplicity and adequate performance we decided to favour the NLF principle, where the originally instantaneously changing rectangular signalling pulses are smoothed using a 90°-segment of a fitted raised-cosine wave. Spectral spillage into adjacent channels was then further mitigated by a low-pass filter (LPF) with a 3 dB down cut-off frequency of $1.5/T$, where $1/T$ is the signalling rate. This filtering scheme has an identical bandwidth to that of a Nyquist filter with a roll-off factor of 0.5. However, it is implementationally more convenient and does not affect the BER performance of our modem, when compared with the scheme using unshaped pulses, which would require an infinite bandwidth.

4 THE RPE SPEECH CODEC

In the widespread analysis-by-synthesis (ABS) speech codecs [35] the typically $N = 40$ samples (5 ms) innovation sequences $v(n)$ are filtered through the long-term predictor (LTP) synthesis filter $1/P(z)$, short-term predictor (STP) synthesis filter $1/A(z)$ and perceptual weighting filter $H(z) = A(z)W(z) = A(z)/A(z/\gamma)$, $\gamma \approx 0.8$ to produce the weighted synthetic speech $\hat{s}_w(n)$:

$$\hat{s}_w(n) = v(n) * h_p(n) * h_w(n) = v(n) * h_c(n) \quad (11)$$

where $h_p(n)$ and $h_w(n)$ are the impulse responses of the filters $1/P(z)$ and $W(z)$, respectively, and $h_c(n)$ is that of the cascaded filter complex $1/P(z)$, $W(z)$ while $*$ means convolution. Similarly, a 5 ms input speech segment about to be encoded is weighted by the identical

perceptual weighting filter, and their difference is computed for each legitimate innovation sequence to find the particular one yielding the minimum weighted error and hence subjectively best synthetic speech.

Depending on the construction of the innovation vectors rather different complexities, bit rates and speech qualities arise. Historically most important is the multi-pulse excited (MPE) codec [36] yielding good speech quality between 9.6 kbit/s and 16 kbit/s at moderate complexity. Variants of the typically low-bitrate (4.8 – 8 kbit/s) CELP codec yield medium quality at high complexity, while the moderate complexity, medium bitrate (13 kbit/s) RPE codec used also in the GSM system provides high speech quality (MOS ≈ 4.0). Spectrally it is almost three times more efficient while maintaining higher robustness than the 32 kbit/s ADPCM codec of the DECT system, and therefore represents an attractive compromise for PCN.

In RPE codecs the innovation sequence $v(n)$ holds M number of equidistant excitation samples with amplitudes β_k and positions m_k :

$$v(n) = \sum_{k=1}^M \beta_k \delta(n - m_k) \tag{12}$$

The minimisation of the total weighted mean squared error (*mse*) over a block of N speech samples is achieved by setting its partial derivatives with respect to the pulse amplitudes β_k , $k = 1 \dots M$ to zero, leading to the set of M simultaneous equations:

$$\Psi(m_i) = \sum_{k=1}^M \beta_k \Phi(m_i, m_k) \quad \text{for } i = 1 \dots M \tag{13}$$

In Eq. (13) $\Psi(m_i)$ represents the cross-correlation of $x(n)$ and $h_c(n)$, given by:

$$\Psi(m_i) = x(n) * h_c(-n) \tag{14}$$

with $x(n)$ being the weighted input speech after subtracting the zero-input response of the filter $W(z) = 1/A(z/\gamma)$ due to its initial state from the previous excitation, and

$$\Phi(m_i, m_k) = \sum_{n=0}^{N-1} h_c(n - m_k) \cdot h_c(n - m_i) \tag{15}$$

is the covariance of the impulse response $h_c(n)$. Eq. (13) is a set of M equations that should be solved for M pulse positions plus M pulse amplitudes, which is unviable. A

computationally attractive, high quality sub-optimum solution was proposed by Kroon et al. [30].

Accordingly the innovation sequence is derived as a sub-sampled version of the STP residual. The excitation pulses are d samples apart and there are d decimated candidate excitation sequences according to the d possible initial grid-positions. If a frame of N prediction residual samples is processed, the number of excitation pulses is given by $M = (N) \text{ Div } (d)$. The legitimate excitation pulse positions are $m[k, i] = k + (i - 1)d$, $i = 1, 2, \dots, M$, where $k = 0, 1, \dots, (d - 1)$ are the initial grid-positions. With the pulse-positions fixed the key-equation Eq. (13) is solved d times for each candidate excitation pattern, yielding d sets of M pulse amplitudes.

Now the total *mse* E_w can be computed over the interval of N samples for all of the d sequences, and the one minimising the weighted *mse* E_w is elected as innovation sequence:

$$E_w = \sum_{n=0}^{N-1} e_w^2(n) = \sum_{n=0}^{N-1} [s_w(n) - \hat{s}_w(n)]^2 = \sum_{n=0}^{N-1} \left[x(n) - \sum_{k=1}^M \beta_k h_c(n - m_k) \right]^2 \tag{16}$$

which can be simplified to

$$E_w = \sum_{n=0}^{N-1} x^2(n) - \sum_{k=1}^M \beta_k \Psi(m_k) \tag{17}$$

where E_w is minimised if the second term of Eq. (17) is maximised.

The key-equation Eq. (13) can be solved either by matrix inversion using Gaussian elimination or deploying Cholesky-decomposition, which has a complexity proportional to M^3 . For the typical values of $N = 40$ and $d = 4$ $M = 10$ equations have to be solved four times during the 5 ms subsegment period. Besides, Eq. (14) and Eq. (15) have to be evaluated as well.

The computational complexity incurred in solving Eq. (13) can be considerably reduced, while maintaining high speech quality. Significant simplification is achieved at almost imperceptible quality degradation assuming that the speech is stationary, rendering the covariance $\Phi(i, j)$ to become $\Phi(i - j) = \Phi(k)$. With this assumption the key-equation Eq. (13) in matrix form is simplified to:

$$\begin{pmatrix} \Phi(0) & \Phi(d) & \Phi(2d) & \dots & \Phi((M-1)d) \\ \Phi(d) & \Phi(0) & \Phi(d) & \dots & \Phi((M-2)d) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \Phi((M-1)d) & \cdot & \cdot & \cdot & \Phi(0) \end{pmatrix} \begin{pmatrix} \beta(k,1) \\ \beta(k,2) \\ \cdot \\ \cdot \\ \beta(k,M) \end{pmatrix} = \begin{pmatrix} \Psi(m(k,1)) \\ \Psi(m(k,2)) \\ \cdot \\ \cdot \\ \Psi(m(k,M)) \end{pmatrix} \tag{18}$$

where the correlation matrix $\underline{\Phi}$ becomes Töplitz. For Töplitz-type systems a plethora of computationally efficient iterative algorithms is available, amongst which the Levinson-Durbin, Robinson, Berlekamp-Massey, Schur and Burg algorithms are best known, all exhibiting a complexity proportional to M^2 rather than M^3 .

As we have seen, $h_c(n)$ is a sharply decaying function, therefore

$$\Phi(i) = \sum_{n=1}^{N-1} h_c(n) h_c(n-i) \quad (19)$$

decays even faster. This allows us to set all off-diagonal elements in the covariance matrix $\underline{\Phi}$ to zero, resulting in tremendously reduced complexity when solving the key-equation Eq. (13) as evidenced by

$$\Phi(0) \begin{pmatrix} \beta & (k,1) \\ \beta & (k,2) \\ \cdot \\ \cdot \\ \beta & (k,M) \end{pmatrix} = \begin{pmatrix} \Psi & (m(k,1)) \\ \Psi & (m(k,2)) \\ \cdot \\ \cdot \\ \Psi & (m(k,M)) \end{pmatrix} \quad (20)$$

$$\beta(k,i) = \frac{\Psi[m(k,i)]}{\Phi(0)} \quad (21)$$

Assuming equal memory contributions in the original and synthetic speech paths enables us to compute $\Psi(m_i)$ in Eq. (14) as

$$\begin{aligned} \Psi(n) &= x(n) * h_c(-n) = s_w(n) * h_c(-n) = \\ &= r(n) * h_w(n) * h_c(-n) \end{aligned} \quad (22)$$

If the minimum LTP delay is restricted to $D > N = 40$, then the LTP does not influence the impulse response $h_c(n)$ for values of $n < N = 40$ and therefore $h_w(n) = h_c(n)$, yielding

$$\Psi(n) = r(n) \cdot h_c(n) \cdot h_c(-n) = r(n) \cdot \Phi(n) \quad (23)$$

Substituting $\Psi(m_i)$ from Eq. (23) into Eq. (21) gives:

$$\begin{aligned} \beta(k,i) &= r[m(k,i)] \cdot \frac{\Phi[m(k,i)]}{\Phi(0)} = \\ &= r[m(k,i)] \cdot \phi[m(k,i)] \end{aligned} \quad (24)$$

where the normalised autocorrelation $\phi[m(k,i)]$ of the impulse $h_c(n)$ response represents a symmetric smoother or low-pass filter (LPF). Further simplification accrues, if we derive a "compromise smoother" from the long-term averaged weighting filter correlations or deploy simply an ideal LPF. For a pulse-spacing of $d = 3$, as in the GSM-standard RPE codec, a cut-

toff frequency of $f_c = 1.3$ kHz has to be used. For an ideal "rectangular" low-pass FIR-filter of order 11 the symmetric impulse response coefficients are simply derived from the Hamming-windowed sinc function:

$$\begin{aligned} \phi(0) &= \phi(10) = -0.016256 \\ \phi(1) &= \phi(9) = -0.045649 \\ \phi(2) &= \phi(8) = 0 \\ \phi(3) &= \phi(7) = 0.250793 \\ \phi(4) &= \phi(6) = 0.70079 \\ \phi(5) &= 1 \end{aligned} \quad (25)$$

Having determined the pulse amplitudes in Eq. (21) and with the aid of Eq. (17) the minimum weighted mse $E_w^{(j)}$ using the j -th of each candidate excitation vector is computed as:

$$\begin{aligned} E_w^{(i)} &= \sum_{n=0}^{N-1} s_w^2(n) - \Phi(0) \sum_{k=1}^M \beta^2(k,i) = \\ &= \sum_{n=0}^{N-1} s_w^2(n) - \Phi(0) E(j) \end{aligned} \quad (26)$$

where $E(j)$ is the energy of the j -th candidate excitation vector. It becomes plausible now that the specific excitation vector having the highest energy, i.e., maximising the second term of Eq. (26) minimises E_w . This is in harmony with our expectations, since this very vector gives the best representation of the prediction residual $r(n)$ when generating the synthetic speech.

Details of the RPE-LTP GSM codec's implementation are readily available in the literature [1, 35], therefore only those necessary for unequal source-matched FEC are discussed here. The LPC analysis framelength is 20 ms covering 160 input speech samples. We found using line spectrum frequencies (LSFs) rather than logarithmic area ratios (LARs) slightly more advantageous in terms of robustness due to their ordering properties, as well as in terms of subjective speech quality, which is a slight deviation from the standardised RPE codec. Instead of utilising 8 LARs quantised with a total of 36 bits we use 10 LSFs encoded similarly with a total of 36 bits. The other subjectively and objectively important variation is the deployment of 4 bits to quantise the long term predictor gains (LTPG) instead of the standardised 2 bits. These have ramifications as regards to bit-sensitivity and increase the number of bits to 268/20 ms, resulting in a bitrate of 13.4 kbit/s, but the reward in terms of speech quality makes it worthwhile. The final bit allocation scheme is summarised in Table 3.

The speech delay introduced by the RPE encoder is 20 ms since 160 speech samples must be buffered before speech encoding can take place. In our proposed system the transceiver is designed so that no additional delay is introduced by interleaving, FEC, modulation, etc. Nonetheless, there will be an implementation-dependent

processing delay, which is typically of the order of a few ms. Also inevitable is the introduction of some wave propagation delay, which can be mitigated by adaptive time frame alignment in large rural cells, as in the GSM system [1]. However, in our micro- or pico-cellular system the propagation delay is typically negligible. The

Table 3 - 13.4 kbit/s RPE codec bit allocation.

Parameter	Bit No.	Bit Position in Frame
8 LFSs	36	1 - 36
RPE gridpos.	2	37,38
Block max.	6	39 - 44
RPE exc. pulses	$13 \times 3 = 39$	45 - 83
LTP delay (LTPD)	7	84 - 90
LTP gain (LTPG)	4	91 - 94
Per Subsegment	58	
Total bitrate: $36 + 4 \times 58 = 268 / 20 \text{ ms} = 13.4 \text{ kbit/s}$		

maximum RPE decoding delay is also 20 ms, since the interleaving and FEC coding memory is limited to a single speech frame. Hence the total speech delay is about 40 ms, which is lower than that of the GSM system [1].

4.1. RPE bit-sensitivity analysis

Due to the variations in the RPE-LTP implementation proposed and the different FEC scheme to be deployed we cannot use the subjectively motivated bit-sensitivity order specified by GSM [1]. Instead, with the aim of providing appropriate embedded bit-protection we systematically corrupted each bit of the RPE-LTP codec in the 268-bit frame, and evaluated the speech quality degradation in terms of both segmental signal-to-noise ratio (*SEG-SNR*) and cepstral distance (CD) [19], as shown in Fig. 4 and Fig. 5 respectively, for the first 94 bits representing the 36 LSF bits as well as the 58 bits of the first subsegment. Impairments caused by the second, third and fourth subsegments for the bits 95-268 in the frame are identical to those of the first one, when they are not depicted here.

Our basic observation is that corruption of the LSFs and LTP parameters as seen at beginning and end of both of the *SEG-SNR* and CD figures results in more serious degradations than that of the excitation bits. The LSFs determine the spectral envelope, while the LTP parameters the spectral fine-structure. However, the *SEG-SNR* sensitivities seem to emphasize the low-frequency LSFs and the LTP parameters, while the CD suggests that almost all LSFs are equally important, but

underestimates the role of the LTP parameters. This is plausible, since the CD excels in measuring the envelope distortion, but not in estimating waveform fidelity. This is why the excitation bits hardly inflict CD degradation, although their role is clearly recognised down to least (LSB), medium and most significant bit (MSB) level in terms of *SEG-SNRs* in Fig. 4. Nevertheless, CD objective measures are regarded to have the highest cor-

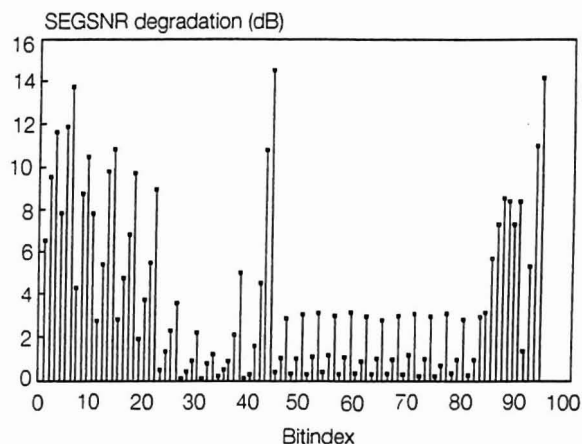


Fig. 4 - Bit sensitivities in *SEG-SNR* degradation for the 13.4 kbit/s RPE codec.

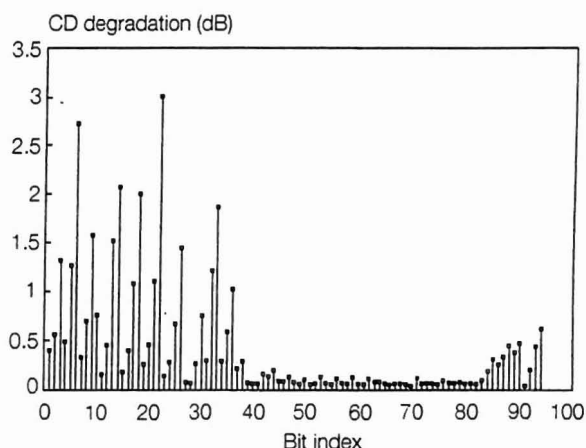


Fig. 5 - Bit sensitivities in CD degradation for the 13.4 kbit/s RPE codec.

relation with subjective listening tests.

To form a balanced view of objective bit-sensitivities and take proper account of subjective quality degradations we decided to amalgamate the *SEG-SNR* and CD measures into a combined sensitivity figure *S*, which is a dimensionless number derived as follows. The bits were assigned a sensitivity rank in terms of both CD and *SEG-SNR*, where the most sensitive speech bit was allocated the *SEG-SNR* and CD sensitivity Fig. 1 and the least sensitive the figure of 94. Then for each bit between 1 and 94 the *SEG-SNR* and CD sensitivity indices were added to form the combined sensitivity figure *S*. Following this the bits were sorted in ascending order of *S*, as tabulated in Table 4, where the constituent

Table 4 - Combined bit-sensitivity figures for the 13.4 kbit/s RPE-LTP codec.

No. in Fr.	Index in Fr.	SEGSNR Sensit.	CD Sensit.	S	No. in Fr.	Index in Fr.	SEGSNR Sensit.	CD Sensit.	S
1	6	3	2	5	51	47	45	52	97
2	14	8	3	11	52	11	49	48	97
3	22	13	1	14	53	62	44	55	99
4	3	5	9	14	54	77	36	67	103
5	5	4	10	14	55	29	71	39	110
6	9	9	6	15	56	80	46	68	114
7	18	11	4	15	57	68	43	74	117
8	13	10	7	17	58	46	62	60	122
9	94	2	19	21	59	74	42	80	122
10	8	14	17	31	60	65	47	77	124
11	2	12	21	33	61	23	75	49	124
12	93	6	27	33	62	34	90	36	126
13	10	18	16	34	63	64	65	61	126
14	17	22	13	35	64	31	93	34	127
15	21	25	12	37	65	76	73	58	131
16	90	17	22	39	66	58	61	70	131
17	26	33	8	41	67	73	66	65	131
18	88	16	25	41	68	79	68	64	132
19	4	19	23	42	69	67	64	71	135
20	44	1	43	44	70	52	60	76	136
21	87	15	31	46	71	82	67	69	136
22	1	23	28	51	72	41	54	83	137
23	89	21	30	51	73	45	78	59	137
24	12	26	26	52	74	55	59	84	143
25	43	7	47	54	75	48	81	63	144
26	20	32	24	56	76	63	84	62	146
27	85	24	33	57	77	91	55	93	148
28	16	29	29	58	78	78	76	73	149
29	86	20	38	58	79	70	57	94	151
30	38	28	35	63	80	54	77	75	152
31	7	31	32	63	81	49	63	89	152
32	33	58	5	63	82	61	70	86	156
33	30	51	15	66	83	27	94	66	160
34	25	50	18	68	84	57	82	79	161
35	92	27	42	69	85	39	92	72	164
36	42	30	46	76	86	69	87	82	169
37	84	34	44	78	87	72	91	78	169
38	32	72	11	83	88	81	88	81	169
39	36	69	14	83	89	28	80	90	170
40	59	35	50	85	90	40	85	85	170
41	71	38	51	89	91	60	79	92	171
42	53	37	53	90	92	51	83	88	171
43	15	48	45	93	93	75	89	87	176
44	19	53	40	93	94	66	86	91	177
45	37	52	41	93					
46	24	56	37	93					
47	35	74	20	94					
48	50	39	56	95					
49	56	41	54	95					
50	83	40	57	97					

SEG-SNR and CD sensitivities are shown as well.

It is extremely instructive to compare our combined objective bit-sensitivities (S) with the tendencies of subjective GSM results [1]. As expected, the most sensitive bits in both objective (SEG-SNR and CD results

in Figs. 4 and 5) and subjective terms appear to be the most significant bits (MSB) of the spectral envelope parameters, whether it is LARs or LSFs. Also the LTPD, LTPG and block maxima MSBs are rated similarly highly by both the objective S measure and the subjective tests, and therefore come in the more vulnerable class one category. In contrast to the GSM system, where 182 class one bits are half-rate coded and 78 bits are not protected, here 116 bits will be approximately 1/2-rate coded, while 152 bits 2/3-rate coded. Clearly, all bits have some protection to prevent unprotected class two bits from leaking through corrupted via the channel. Conceptually this is quite important to yield high, error-free speech performance beyond the target SNR and SIR values to be maintained by adaptive power control. Observe that in the GSM system the unprotected 78 class two bits will always have a residual BER and hence degrade speech quality. In general, the LSBs of most parameters are in the second protection class and on the whole, although there are slight differences between our objective S measure and the whole-scale GSM subjective tests, the basic tendencies are identical and the advantages of the objective analysis are unquestionable.

5. CHANNEL CODING

The type of FEC coding still remains to be resolved. It is widely recognised that the highest coding gain over AWGN channels is achieved using trellis-coded modulation [43]. Recent advances in coded modulation and bandwidth efficient transmission published in references [44, 45] provide a plethora of attractive transceiver schemes for fading mobile channels. However, for speech transmissions source sensitivity-matched unequal error protection has to be used. Cox, Hagenauer et al. used in their classic paper [37] rate compatible punctured convolutional (RCPC) codecs [38] to provide bit sensitivity matched FEC protection for a subband speech codec. Bit puncturing was used to gradually increase the coding rate from an $R = 1/2$ rate mother code to achieve a finely graded set of coding rates, approaching unity, as the source bit sensitivity decreased. A further attractive feature of their solution was that a single decoder was capable of decoding all the RCPC codes derived from the $R = 1/2$ rate mother code.

In a similar but less ambitious study we compared RCPC codecs with rate compatible shortened Reed-Solomon (RS) codecs [39], which possessed all the attractive RCPC features. The $R = 1/2$ rate mother code for the RCPC codecs was generated using the GSM system's generator polynomials and Hagenauer's puncturing matrices were used to generate coding rates of $R = 8/15$, $8/14$, $8/13$, $8/12$, $8/11$ and $8/10$. In order to arrive at a set of appropriate RS codes for this comparative test the primitive double error correcting RS (15, 11, 2) code over the finite Galois field GF (16) was systematically shortened to rates $R = 10/14$, $9/13$, $8/12$, $7/11$, $6/10$, $5/9$ and

$4/8$. In terms of decoded BER both codes had similar performances, with the RS codec slightly outperforming the RCPC codec for $BER < 1\%$ over Gaussian channels, while above this point the RCPC codec performed marginally better. When these FEC codecs were used to provide unequal protection for the 148 bits/20 ms sequences of a 7.4 kbit/s code excited linear predictive (CELP) codec with an FEC-coded rate of 11.4 kbit/s, the $SEGSNR$ versus average Rayleigh fading channel SNR performances were again quite similar.

When we used RS or binary Bose-Chaudhuri-Hocquenghem (BCH) codes having higher block lengths, we found that the subjective speech quality improved due to longer unimpaired speech segments, even if the objective $SEGSNR$ and BER performances of the RCPC and block codes were similar. This was because the speech quality is typically more strongly dependent on the frame error rate (FER) than on decoded BER. Furthermore, powerful block codes also have reliable error detection capability, which can be advantageously exploited to control speech post-enhancement, as suggested in [18] or to initiate handovers, etc. The GSM system for example employs an external error detecting block code in addition to a half-rate constraint-length five convolutional code in order to provide a means of channel BER monitoring.

For all these advantages [46], in harmony with the Japanese second generation system, in this contribution we favoured block codes [35]. The channel interleaver has to disperse errors concentrated around deep fades so that the FEC codec can have approximately equal number of errors in each block. Naturally, long block codes inherently possess this randomising property, but are complex to implement. A very good compromise in terms of implementational complexity and error correcting power is constituted by the family of binary BCH codes of 63 bits length. The 13 kbit/s RPE-GSM codec has a frame-length of 20 ms, whence we limit the interleaving depth, i.e., the transmitter's delay to one linear predictive coding (LPC) frame of 20 ms. As an example, let us relate this to a vehicular speed of 30 miles/h or 13.3 m/s giving a travelling distance of 26.6 cm/20 ms. For propagation frequencies of 1.9 GHz, as in the planned PCN, the wavelength is around 15 cm, and therefore interleaving over an interval of 26.6 cm/20 ms or approximately two wavelengths ensures efficient randomisation. However, for slowly walking pedestrians there is a danger of idling in deep fades, which is detrimental as regards to reception quality. In this situation a switch-diversity scheme is essential, unless slow frequency hopping is used, as in the GSM system [1], hopping to a new carrier every new TDMA burst and thereby reducing the concentration of error bursts.

Armed with the sensitivity figures S tabulated in Table 4, an arbitrarily complex embedded FEC scheme can be implemented. For practical reasons, similarly to the GSM system, we opted for two protection classes. Observe that assigning any subsegment bit in the RPE frame with an index $95 > X > 36$ to a specific category implies the in-

clusion of the corresponding bits ($X + 58$), ($X + 116$) and ($X + 174$) of the other subsegments in the same class.

After careful consideration of the sensitivity figures given in Table 4, we accommodated 116 bits in the high protection class and assigned a shortened binary BCH (62, 29, 6) code, correcting up to six bit errors in a 62 bit codeword, embracing 29 speech bits. The remaining 152 less significant speech bits are protected by a less powerful shortened BCH (62, 38, 4) code. The codes are appropriately shortened to embrace the 116 and 152 class one and class two speech bits by four rectangularly interleaved BCH codewords in each category, which conveniently curtails error propagation between adjacent LPC frames. Then in case of serious speech impairments the reliable error detection capability of the high protection BCH code [35] is invoked to activate the correlative speech post-processing algorithm portrayed in [18].

The signalling rate of our RPE/BCH/16-QAM system is computed as follows. The 116 highly protected bits are coded into four consecutive BCH (62, 29, 6) codewords giving a total of 248 bits. Similarly, the 152 low-protection bits are encoded by four BCH (62, 38, 4) codewords, yielding equally 248 bits. Then a total of 124 16-QAM symbols per each 20 ms LPC frame are transmitted, and so the signalling rate is 6.2 kBd.

As regards to overall system delay, the transceiver design ensures that no additional delay is introduced by the FEC, interleaving and modem arrangement. Therefore the total speech delay is attributable to the 20 ms speech buffering at both the encoder and decoder, plus the comparatively low processing and propagation delay. Hence the overall speech delay is approximately 40 ms, lower than that of the GSM system [1].

6 THE 6.2 kBd PCN TRANSCIVER

The schematic diagram of our PCN transceiver is depicted in Fig. 6. The RPE speech encoder's bits are sorted by the "Mapper" into two sensitivity classes according to their vulnerability, as described in Section 4 and summarised in Table 4. Twin-class unequal error protection is deployed following the principles outlined in [37] and the FEC coded bits are rectangularly interleaved over the current 20 ms frame of speech. The interleaved bits are mapped back in one stream by another bit "Mapper", multiplexed (MPX) by the Time Division Multiple Access (TDMA) arrangement, and 16-StQAM modulated onto a 1.9 GHz radio carrier and transmitted via the Rayleigh fading channel. The receiver carries out the inverse functions of the transmitter. Observe that the error detection capability of the stronger C1 FEC decoder is exploited to invoke speech post-processing, if the decoder happens to be overloaded. Furthermore, this error detection capability can also be utilised to monitor the channel's BER statistics and control handovers to other base stations (BS) providing a higher signal-to-noise ratio.

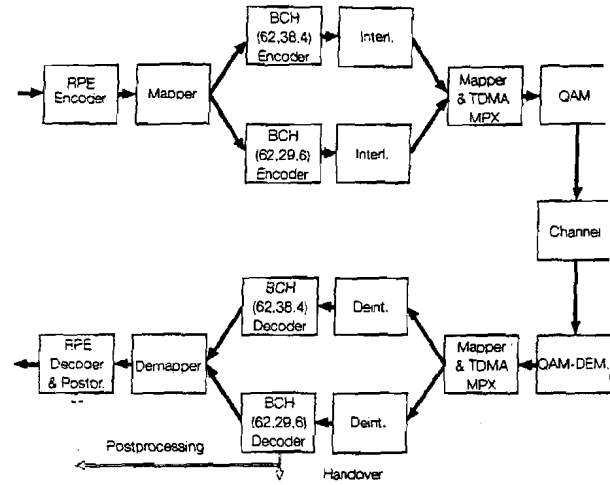


Fig. 6 - System's Schematic.

7 RESULTS AND DISCUSSION

7.1. Spectral occupancy

Let us now compare our proposed 4 bits/symbol 6.2 kBd RPE/BCH/16-StQAM system with a 2 bits/symbol RPE/BCH/ $\pi/4$ -DQPSK and a binary RPE/BCH/GM-SK system in terms of spectral occupancy. The attributes of these transceiver schemes that will be referred to as Systems A, B and C are summarised in Table 5. The speech source rates are 13.4 kbit/s and the FEC-coded rates are 24.8 kbit/s. However, due to variable 1-4 bits/symbol transmissions the source Baud rates are different. Assuming an overall time division multiple access (TDMA) channel bandwidth of 200 kHz, as in the GSM system, the signalling rates become 271 kBd, 162 kBd and 133 kBd, respectively.

Table 5 - Comparison of PCN Transceivers.

	System A	System B	System C
Modem	GMSK	$\pi/4$ -DQPSK	16-StQAM
Speech rate (kbit/s)	13.4	13.4	13.4
FEC-coded rate (kbit/s)	24.8	24.8	24.8
Source Baud rate (kBd)	24.8	12.4	6.2
Channel Baud rate (kBd)	271	162	133.33
No. of TDMA users carrier	10	13	21
TDMA user bandwidth (kHz)	20	15.4	9.3

Explicitly, for GMSK the GSM rate was assumed, which ensures a modem bandwidth efficiency of 1.35 bits/s/Hz. The $\pi/4$ -DQPSK modem requires a Nyquist roll-off factor of 0.35, as in the IS-54 system. In order to have an infinite spectral attenuation at the transmission band edge the maximum signalling rate is $200/1.35 = 148$ k Bd, which implies a modem bandwidth efficiency of $148 \text{ k Bd} \times 2 \text{ bit/symbol}/200 \text{ kHz} = 1.48 \text{ bits/s/Hz}$. However, the IS-54 system for example transmits at a Baud rate of 48.6 k Bd in a bandwidth of 30 kHz, implying a modem efficiency of 1.62 bits/s/Hz. From our experience we found that this is equivalent to the assumption of a 24 dB spectral decay at the transmission band edge. Assuming therefore a modem efficiency of 1.62 bits/s/Hz, System B can transmit at $200 \text{ kHz} \times 1.62 \text{ bits/s/Hz} = 324 \text{ k bit/s} = 162 \text{ k Bd}$. For the 16-StQAM modem of System C we used a LPF limiting the spectrum at 1.5 times the Baud rate. Hence the bandwidth of 200 kHz supports a 16-StQAM signalling rate of 133.33 k Bd.

Using our previously derived source and channel Baud rates the number of TDMA users accommodated within the 200 kHz bandwidth is given in the last but one row of Table 5. The corresponding TDMA user bandwidths are listed in the last row of Table 5. Explicitly, in an indoors cordless telecommunication (CT) system, where the partitioning walls form natural pico-cell boundaries with high attenuation of co-channel interferers, and the higher SNR required by the 16-StQAM scheme can be provided by appropriate power budget, the 16-StQAM-based System C can accommodate more users within the same bandwidth than Systems A and B.

7.2. 16-StQAM performance

In this subsection we characterise the performance of our proposed 16-StQAM schemes in terms of BER versus channel SNR and SIR. Linear amplification was assumed, no AGC and carrier recovery were invoked, and our previously described NLF pulse-shaping with the concatenated LPF was utilised to contain the spectrum within a band of 1.5 times the Nyquist frequency.

Micro- and pico-cellular indoors wideband propagation measurements in the 1.9 GHz band, where PCN is likely to be allocated, are not yet widely available, nonetheless the standardised Digital European Cordless Telephone (DECT) scheme operates without a channel equalizer at a signalling rate of 1152 k bit/s. This implies that the micro- and pico-cellular channel's coherence bandwidth is well above 1 MHz and therefore its impulse response dispersion is below 1 μ s. In this contribution we used a TDMA signalling rate of 133.33 k Bd, about an order of magnitude lower than that of the DECT scheme, and hence no equalizer was required.

The BER versus channel SNR performance of our 16-StQAM modem is depicted with and without diversity for vehicular speeds of 48 and 6 km/h in Fig. 7. Observe that at 48 km/h, which is a typical urban microcellular speed, the modem's differential codec was not able to

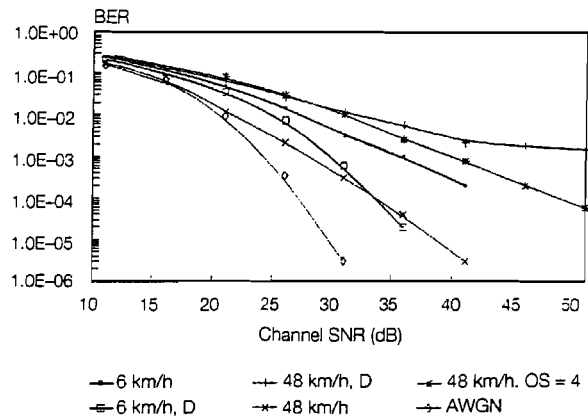


Fig. 7 - BER versus channel SNR performance of 16-StQAM with and without diversity at speeds of 48 and 6 km/h.

adequately trace the fading envelope and phase trajectory in the steepest fading sections, which resulted in a residual BER of about 10^{-3} . However, when an oversampling ratio (OS) of 4 was used, the residual BER disappeared. The performance of this oversampled scenario with $OS = 4$ was similar to that of a four times slower vehicular speed. The typical pedestrian scenario of 6 km/h speed is also portrayed in the figure and is expected to be similar to an $OS = 8$ scenario, since the PS speed was reduced by a factor of eight from 48 km/h to 6 km/h. Hence it is seen to be slightly more favourable in terms of BER than the $OS = 4$ oversampled modem performance. Again, the surprisingly low BER is attributed to the differential coding deployed, which effectively traces the fading envelope. When second order switched diversity (D) was invoked, the BER performance further improved both at 48 and 6 km/h. However, in the low SNR region at 6 km/h the diversity scheme was not able to improve the performance as dramatically, as at 48 km/h. This is because at low speeds, i.e., low fading rates both diversity channels received more correlated signals, and hence both received signals were simultaneously of high or low received signal power. Nonetheless, at both speeds near-Gaussian BER performance was achieved, when diversity was deployed. Since our 16-StQAM-based system is proposed for cordless telecommunications (CT) in predominantly indoors cells, its favourable performance at 6 km/h is convenient for our application. Furthermore, the additional signal processing complexity of the oversampling is not required at walking speed, but the effect of diversity is extremely positive. Hence in our final proposed scheme, System C, the non-oversampled, diversity assisted modem was used, but we emphasize that for higher vehicular speeds oversampling might be required.

Although in indoors cells the partitioning walls and floors form natural microcell boundaries and hence conveniently curtail co-channel interferences, it is important to quantify the expected performance degradations in cellular systems, where high co-channel interference prevails. An explicit formula for the overall spectral efficiency of cellular systems with frequency re-use was de-

rived by Gejji [40], which is based on Lee's former work [41] as well as on Shannon's channel capacity equation, where inverse fourth power wave propagation law and six closest co-channel interfering cells were assumed.

In this contribution we characterise our 16-StQAM modem in terms of its channel bit error rate (BER) versus channel SNR performance, parameterised by the SIR values of 10 dB, 20 dB and 30 dB as well as ∞ corresponding to no co-channel interference, as seen in Fig. 8. A Rayleigh-faded single interferer was used, which corresponds to the worst case scenario, and the portable station (PS) speed was 6 km/h, as in pedestrian microcells. Observe that with no co-channel interference there was no residual BER at the assumed low PS speed, since our differentially coded modem was able to effectively trace the Rayleigh fading envelope and phase. However, with SIR values of 10 dB and 20 dB the residual BER became excessively high for the proposed BCH codes and the speech quality was impaired. An SIR of 25-30 dB was required for channel SNRs in excess of about 25 dB for error-free speech quality.

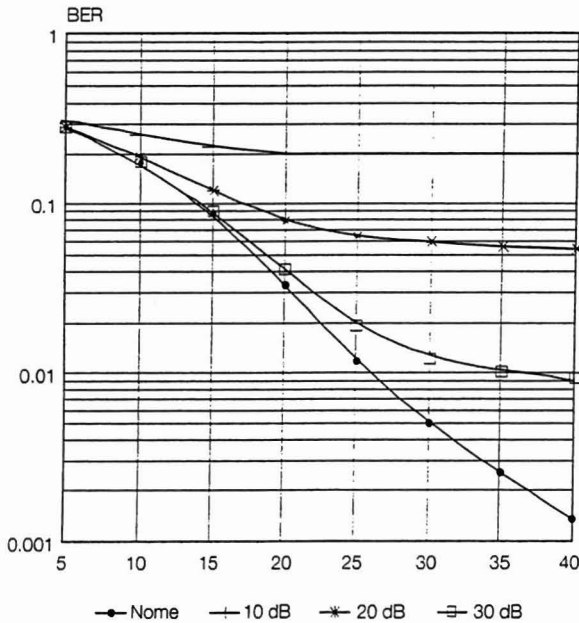


Fig. 8 - 16-StQAM BER vs channel SNR and SIR.

For these reasons it should be noted that the number of channels supported by higher level modulation schemes, such as 16-StQAM would be reduced in a cellular system with frequency re-use. A study concerned with the efficiency of a range of modulation schemes when both the number of modulation levels and the required cluster sizes were varied was given in reference [42] for cellular systems. This showed that the efficiency is dependent on the required BER, but in cellular systems having severe co-channel interference multi-level schemes can become less efficient than binary or quaternary constellations. Therefore the relative gains of the 16-StQAM scheme proposed in our CT system must be appropriately moderated, when deployed in a cellular

system with frequency re-use.

In Fig. 9 we portray the BER performance of our BCH-coded modem both without and with diversity (D) at 6 km/h, using the chosen BCH (62, 29, 6) and BCH (62, 38, 4) codecs. In speech communications the frame error rate (FER) is also of high importance, since the perceived speech quality depends on the percentage of

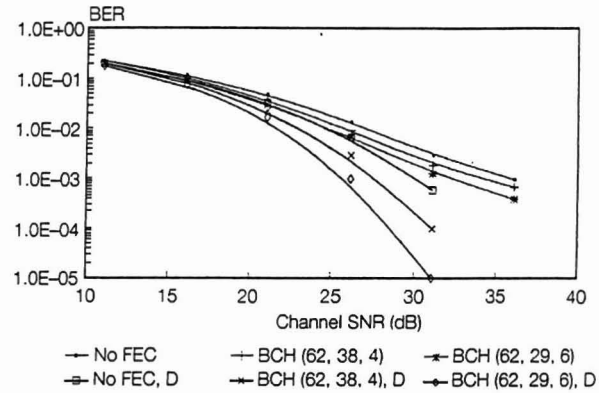


Fig. 9 - FEC coded 16-StQAM BER versus channel SNR at 6 km/h with and without diversity.

time, during which unimpaired speech quality is experienced. Therefore in Fig. 10 we plotted the FER of the BCH (62, 29, 6) and BCH (62, 38, 4) codecs versus channel SNR for our diversity assisted modem at a vehicular speed of 6 km/h for both protection classes. Note that the RPE-coded speech FER would be higher than the BCH-coded FER, since a 20 ms speech frame was represented by four BCH-coded frames of each protection class, but the perceived speech quality is proportional to the FER of the high protection class, since the errors inflicted in the low-protection class are subjectively much less objectionable. Observe that while with diversity a BCH (62, 29, 6) FER of 1% required by the speech codec was achieved at a channels SNR of about 25 dB for the more sensitive bit protection class, without diversity an SNR of about 30 dB was needed.

Lastly, in Fig. 11 we plotted the objective speech SEGSNR measure against channel SNR, in order to cha-

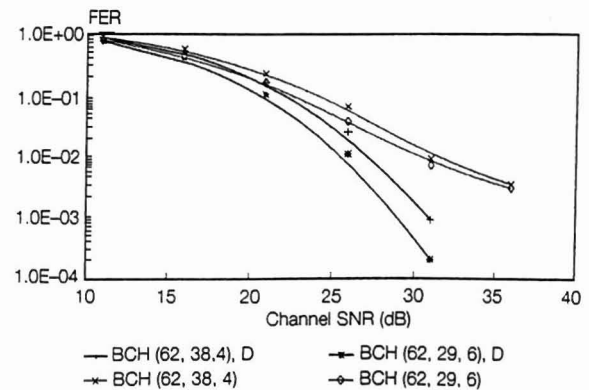


Fig. 10 - BCH-coded C1 and C2 FER versus channel SNR with and without diversity.

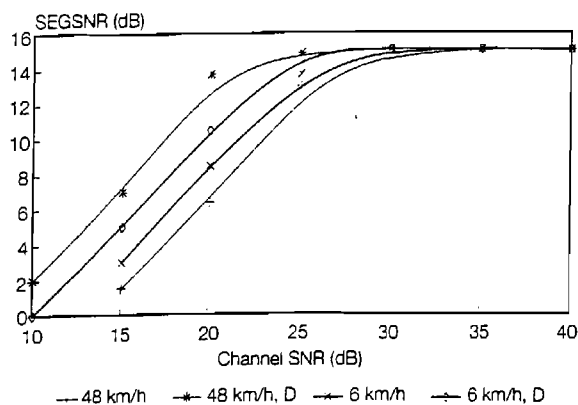


Fig. 11 - SEGSNR versus channel SNR of the proposed RPE/BCH/16-StQAM systems at 48 and 6 km/h with and without diversity.

racterise the achievable speech quality both with and without diversity at 48 and 6 km/h. As expected, the SEGSNR preference order followed the BER performances evaluated in Fig. 7 and Fig. 9. Best performance was achieved by the diversity assisted 48 km/h scenario, which degraded by about 5 dB without diversity. As anticipated from the BER curves, the performance curves with and without diversity at 6 km/h were fairly close to each other, and an SNR of about 25 dB was high enough in order to achieve virtually unimpaired speech quality.

8 SUMMARY AND CONCLUSIONS

In this contribution a variety of robust 16-StQAM modem schemes for CT have been proposed. We have shown that the suggested differentially coded non-coherent modem can effectively trace the fading magnitude and phase trajectories and hence does not require the assistance of a pilot symbol aided AGC. Its performance greatly benefited from second order diversity at higher vehicular speeds. At walking speeds the fade-tracking differentially coded modem's performance improved, while the efficiency of the diversity scheme degraded and hence the performance gap between the diversity aided and non-diversity scenarios was reduced.

The 16-StQAM modem is sensitive to co-channel interference and an SIR of about 25-30 dB was required in order to reduce the BER floor below 1%, which implies that in cellular systems the modem's bandwidth efficiency gain would be eroded by inevitably higher frequency re-use distances. However, for indoors CT, where mild interference prevails this SIR value can be achieved.

The RPE speech codec proposed was subjected to bit-sensitivity analysis using a combination of SEGSNR and CD objective speech measure degradations and the bits were assigned into two sensitivity classes. The cost- and time-efficient, objectively motivated bit-sensitivity analysis provided results in line with the expensive formal subjective GSM sensitivity tests [1]. In our twin-class BCH FEC scheme no bits were transmitted unprotected via the channel, which allowed unimpaired speech

communications for channel SNRs in excess of about 25 and 30 dB with and without diversity, respectively. The speech SEGSNR was about 15 dB associated with a Mean Opinion Score of about four, while the overall speech delay was confined to about 40 ms. The number of users supported within a 200 kHz channel slot was increased from 10 to 21, when deploying 16-StQAM facilitated by the benign CT propagation environment, while maintaining low complexity. The user bandwidth was 9.3 kHz. Further details on QAM for mobile radio communications can be found in reference [47].

Manuscript received on April 20, 1993.

REFERENCES

- [1] Group Mobile Special (GSM) ETSI 05.03. Recommendation. Apr. 1988.
- [2] G. Larson, B. Gudmundson, K. Raith: *Receiver performance for the North American digital cellular system*. "Proc. of IEEE-VTC", St. Louis, U.S.A. 19-22 May, 1991, p 1-6.
- [3] M. Ikura, K. Ohno, Y. Yamao, F. Adachi: *Field experiments on TDMA mobile radio signal transmissions*. "Proc. of IEEE-VTC", St. Louis, U.S.A. 19-22 May, 1991, p 669-674.
- [4] J.C. Hancock, R.W. Lucky: *Performance of combined amplitude and phase modulated communications systems*. "IRE Trans. Commun. Systems", Vol CS-8, 1960, p 232-237.
- [5] C.R. Cahn: *Combined digital phase and amplitude modulation communication systems*. "IRE Trans. Commun. Systems", Vol CS-8, 1960, p 150-154.
- [6] C.N. Campopiano, B.G. Glazer: *A coherent digital amplitude and phase modulation scheme*. "IRE Trans. Commun. Systems", Vol CS-10, 1962, p 90-95.
- [7] G.D. Forney et al.: *Efficient modulation for band-limited channels*. "IEEE Journal on Selected Areas in Communications", Vol SAC-2, No 5, Sept. 1984, p. 632-647.
- [8] K. Feher: *Modems for emerging digital cellular mobile systems*. "IEEE Tr. on VT", Vol 40, No 2, May 1991 p 355-365.
- [9] M. Iida, K. Sakniwa: *Frequency selective compensation technology of digital 16-QAM for microcellular mobile radio communication systems*. Proc. of VTC'92, Denver, Colorado, p 662-665.
- [10] R.J. Castle, J.P. McGeehan: *A multilevel differential modem for narrowband fading channels*. Proc. of VTC'92, Denver, Colorado, p 104-109.
- [11] D.J. Purle, A.R. Nix, M.A. Beach, J.P. McGeehan: *A preliminary performance evaluation of a linear frequency hopped modem*. Proc. of VTC'92, Denver, Colorado, p 120-124.
- [12] Y. Kamia, S. Sampei: *Performance of reduced complexity DFE using bidirectional equalizing in land mobile communications*. Proc. of VTC'92, Denver, Colorado, p 372-376.
- [13] T. Nagayasu, S. Sampei, Y. Kamio: *Performance of 16-QAM with decision feedback equalizer using interpolation for land mobile communications*. Proc. of VTC'92, Denver, Colorado, p 384-387.
- [14] E. Malkamaki: *Binary and multilevel offset QAM, spectrum efficient modulation schemes for personal communications*. Proc. of VTC'92, Denver, Colorado, p 325-378.
- [15] Z. Wan, K. Feher: *Improved efficiency CDMA by constant envelope SQAM*. Proc. of VTC'92, Denver, Colorado, p 51-55.
- [16] H. Sasaoka: *Block coded 16-QAM/TDMA cellular radio system using cyclical slow frequency hopping*. Proc. of VTC'92, Denver, Colorado, p 405-408.
- [17] W.T. Webb, L. Hanzo, R. Steele: *Bandwidth-efficient QAM schemes for rayleigh-fading channels*. "IEE Proceedings", Vol 138, No 3, June 1991, p 169-175.

- [18] L. Hanzo, R. Steele, P.M. Fortune: *A subband coding, BCH coding and 16-QAM system for mobile radio speech communications*. "IEEE Vehicular Technology", Nov., 1990, Vol 39, No 4, p 327-339.
- [19] L. Hanzo, R. Salami, R. Steele, P.M. Fortune: *Transmission of digitally encoded speech at 1.2 kBd for PCN*. "Proc. of the IEE-Part I", Vol 139, No 4, Aug. 1992, p 437-447.
- [20] P.M. Fortune, L. Hanzo, R. Steele: *On the computation of 16-QAM and 64-QAM performance in rayleigh-fading channels*. "IEICE Tr. on Comms.", Vol E75-B, No 6, June 1992, p 466-475.
- [21] P.B. Kennington et al.: *Broadband linear amplifier design for a PCN base-station*. "Proc. of 41st IEEE VTC", May 1991, p 155-160.
- [22] R.J. Wilkinson et al.: *Linear transmitter design for MSAT terminals*. 2nd Int. Mobile Satellite Conference, June 1990.
- [23] S.P. Stapleton, F.C. Costescu: *An adaptive predistorter for a power amplifier based on adjacent channel emissions*. "IEEE Tr. on VT", Vol 41, No 1, Febr. 1992, p. 49-57.
- [24] S.P. Stapleton, F.C. Costescu: *An adaptive predistorter for a power amplifier based on adjacent channel emissions*. "IEEE Tr. on VT", Vol 41, No 1, Febr. 1992, pp. 49-57.
- [25] S.P. Stapleton, G.S. Kandola, J.K. Cavers: *Simulation and analysis of an adaptive predistorter utilizing a complex spectral convolution*. "IEEE Tr. on VT", Vol 41, No 4, Nov. 1992, p. 387-394.
- [26] A.S. Wright, W.G. Durtler: *Experimental performance of an adaptive digital linearized power amplifier*. "IEEE Tr. on VT", Vol 41, No 4, Nov. 1992, p. 395-400.
- [27] M. Faulkner, T. Mattson: *Spectral sensitivity of power amplifiers to quadrature modulator misalignment*. "IEEE Tr. on VT", Vol 41, No 4, Nov. 1992, p. 516-525.
- [28] R. Stedman, H. Gharavi, L. Hanzo, R. Steele: *Transmission of subband coded images via mobile channels*. "IEEE Tr. on Circuits and Systems for Video Technology", Vol 3, No 1, Febr. 1993, p. 15-27.
- [29] J.K. Cavers: *An analysis of pilot symbol assisted modulation for rayleigh fading channels*. "IEEE Tr. on VT", Vol 40, No 4, Nov. 1991, p 686-693.
- [30] P. Kroon, F. Deprettere, R.J. Sluyter: *Regular-pulse excitation - A novel approach to effective multipulse coding of speech*. "IEEE Tr. ASSP", Vol-ASSP-34, No 5, Oct. 1986, p 1054-1062.
- [31] L.E. Franks: *Carrier and bit synchronisation in data communications - A tutorial review*. "IEEE Tr. on Comms.", Vol-COM 28, No 8, Aug. 1980, p 1107-1121.
- [32] R.L. Cupo, R.D. Gitlin: *Adaptive carrier recovery systems for digital data communications receivers*. "IEEE-JSAC", Vol 7, No 9, Dec. 1989, p. 1328-1339.
- [33] W.C. Lindsey, M.K. Simon: *Carrier synchronisation and detection of polyphase signals*. "IEEE Tr. on Comms.", June 1972, p. 441-454.
- [34] E. Issman, W.T. Webb: *Carrier recovery for 16-level QAM in mobile radio*. "IEE Colloquium on Multi-Level Modulation Techniques", March 1990, p.1990/045.
- [35] R. Steele, Ed.: *Mobile radio communication*. Pentech Publishers, Sept. 1992.
- [36] B.S. Atal, J.R. Remde: *A new model of LPC excitation for producing natural-sounding speech at low bitrates*. "Proc. ICASSP'82", p. 614-617.
- [37] R.V. Cox, J. Hagenauer, N. Seshadri, C.E.W. Sundberg: *Subband speech coding and matched convolutional channel coding for mobile radio channels*. "IEEE Tr. on Sig. Proc.", Vol 39, No 8, Aug. 1991, p. 1717-1731.
- [38] J. Hagenauer: *Rate-compatible punctured convolutional codes (RCPC) and their applications*. "IEEE Tr. on Comms.", Vol 36, No 4, Apr. 1988, p. 389-400.
- [39] R. A. Salami, K. H. H. Wong, R. Steele, D. Appleby: *Performance of error-protected binary pulse excitation coders at 11.4 kbit/s over mobile radio channels*. "Proc. of ICASSP'90", p. 473-476.
- [40] R.R. Gejji: *Channel efficiency in digital cellular communications systems*. Proc. of the 42nd IEEE VTC, Denver, Col. U.S.A., May 1992, p. 1005-1005.
- [41] W.Y.C. Lee: *Spectrum efficiency in cellular*. "IEEE Tr. on VT", Vol 38, No 2, May 1989, p. 69-75.
- [42] W.T. Webb: *Modulation methods for PCN*. "IEEE Comm's Mag.", Vol 30, No 12, Dec. 1992, p. 90-95.
- [43] G. Ungerböck: *Trellis-coded modulation with redundant signal sets. Part I&II*. "IEEE Comm's Mag.", Vol 25, No 2, Febr. 1987, p. 5-21.
- [44] E. Biglieri, M. Luise: *Coded modulation and bandwidth-efficient transmission*. Proc. of the Fifth Tirrenia Intern. Workshop, 8-12 Sept. 1991, Elsevier, Netherlands, 1992.
- [45] IEEE Communications Magazine, Special Issue on Coded Modulation, Vol 29, No 12, Dec. 1991.
- [46] K.H.H. Wong, L. Hanzo, R. Steele: *Channel coding for satellite mobile channels*. "Int. Journal on Satellite Communications", Vol 7, No 2, 1989, p. 143-163.
- [47] W.T. Webb, L. Hanzo: *Quadrature amplitude modulation*. Research Monograph, to appear Dec. 1993, Pentech Publishers, London, U.K.