Speech Performance of Adaptive Transceivers for PCN

R.Steele, J.E.B. Williams, L.Hanzo [1]

Abstract

Experiments are conducted using a number of adaptive transmission systems that are suitable for third generation personal communications networks (PCN). It becomes apparent that the selection of the type of modem and number of modulation levels of the modem, the speech codec to be used, and whether to deploy a channel codec, depends on the channel SNR and teletraffic demand if the power consumption of the hand held portable is to be minimised. Following the implications of our experiments, we generalise our findings to some of the requirements of an adaptive transceiver that might be deployed in the next generation of mobile radio networks.

1 Introduction

There is much activity world wide in attempting to define the third generation personal communications network (PCN). In designing a third generation system cognizance is given to the second generation systems. Indeed, we may anticipate that some of the sub-systems of GSM and DECT may find their way into PCNs either as a primary sub-system, or as a component to achieve backward compatibility with systems in the field. This approach may result in hand-held transceivers that are intelligent multimode terminals, able to communicate with existing networks, while having more advanced and adaptive features that we would expect to see in the next generation of PCNs.

In this text we present the results of a series of experiments we conducted. We concerned ourselves with the radio link and introduced a menu that includes some second generation sub-systems for backward compatibility. We also included 16-level quadrature amplitude modulation (16-QAM), as we consider that this type of modulation can be made adaptive and reconfigure itself, for example as $\frac{\pi}{4}$-shifted differential quadrature phase shift keying ($\frac{\pi}{4}$-DQPSK) and Gaussian minimum shift keying (GMSK). Essentially our deliberations centre around the concept of adaptive multi-level modulation and the types of appropriate speech and channel coding that may be required. In this short study we restricted ourselves to the microcellular propagation environment.

2 Modem Schemes

The choice of modem is based on the interplay of equipment complexity, power consumption, spectral efficiency, robustness against channel errors, cochannel and adjacent channel interference, as well as the propagation phenomena which depends on the cell size [1]. Equally important are the associated issues of linear or non-linear amplification and filtering, the applicability of non-coherent, differential detection, soft-decision detection, equalisation and so forth.

In our experiments we used three different modems, namely GMSK, $\frac{\pi}{4}$-DQPSK and 16-StQAM, each with a low- and a high-complexity detector. In our GMSK modem a normalised bandwidth bit-timing product of $B_T = 0.3$ was favoured, which was adopted by the Pan-European second generation system known as GSM. The typical bandwidth efficiency of GMSK is about 1.35 bit/Hz. A low-complexity frequency discriminator and a higher-complexity Viterbi detector were deployed [1]. $\frac{\pi}{4}$-DQPSK modulation was favoured by the Americans in their IS-54 D-AMPS network and by the Japanese digital cellular system. We decided to compare the performance of a low-complexity differential detector to a more complex maximum likelihood correlation receiver (MLH-CR). We used square root raised cosine Nyquist filters with a roll-off factor of 0.35 and achieved an approximate modem bandwidth efficiency of 1.64 bit/Hz. The differentially coded 16-level twin-ring star QAM (16-StQAM) included in our studies was shown to work well in microcellular environments, where high SNRs and low dispersion are generally the norm and the high traffic density requirements justify the use of linear amplification. The low-complexity non-coherent differential detector's performance [2] was compared to that of

[1] The authors are with Multiple Access Communications Ltd.

a MLH-CR. When using square root raised cosine Nyquist filters with a roll-off factor of 1 the modem bandwidth efficiency was 2.4 bit/Hz.

3 Speech Codecs

In recent years both speech coding research and IC-technology went through a revolutionary development, culminating in a number of standards with bit rates as low as 4.8 kbps [1]. The selection of the speech codec for mobile applications is based on the appropriate combination of parameters such as speech quality, computational complexity related to power consumption, bit rate, delay and robustness against channel errors.

The two speech codecs included in our experiments are the CCITT G 721 32 kbps adaptive differential pulse code modulation (ADPCM) codec used in the DECT system, and the 13 kbps GSM and DCS 1800 Regular Pulse Excited (RPE), Long Term Predictor (LTP) assisted codec. The ADPCM codec used in our schemes fully complies with the CCITT Recommendation so we will not describe its operation here.

Our RPE-LTP codec is a slightly modified version of the standard GSM scheme [3], transmitting 268 bits per 20 ms frame and hence the overall bit rate is increased from 13 kbps to 13.4 kbps. For robustness against channel errors some speech bits are better protected than others. To identify the effect of bit errors on speech quality in our modified GSM codec we performed a bit sensitivity analysis based on a combination of Segmental SNR (SSNR) and Cepstral Distance (CD) degradation. Our findings are presented in reference [3], which suggested a sensitivity order very similar to those found by GSM, verifying our approach. For the sake of low complexity a simple twin-class error protection scheme is used, where the more important bits are termed as class 1 (C1) bits and the less sensitive ones as class 2 (C2) bits.

In comparison, the 32 kbps ADPCM waveform codec has a segmental SNR (SSNR) of about 28 dB, while the 13 kbps analysis-by-synthesis (ABS) RPE-LTP codec has a lower SSNR of about 16 dB, associated with similar subjective quality rated as a mean opinion score (MOS) of about four. This discrepancy in SSNR values is because the RPE-LTP codec utilises perceptual error weighting, which degrades the objective speech quality in terms of both SSNR and CD, but improves the subjective speech quality. The cost of the RPE-LTP codec's significantly lower bit rate and higher robustness compared to ADPCM is its increased complexity and encoding delay.

4 Channel Coding and Bit-mapping

The cordless telecommunications (CT) schemes considered here are based on the ADPCM codec. No error correction coding is used, because it is not necessary in the microcellular environments at the bit rates employed. In the higher complexity, more robust, lower bit rate speech transmission systems based on the RPE-LTP codec we favour binary Bose-Chaudhuri-Hocquenghem (BCH) block codes. They combine low computational complexity with high error correcting power and have reliable error detection properties, which are used by our speech post-enhancement scheme [4], when the received speech is corrupted due to an FEC decoding failure. The channel codec may also be employed to assist in initiating fast handovers, which are vitally important in small microcells, where they may occur frequently and fast [5]. In our twin-class FEC scheme [3] the 116 C1 bits of the RPE-LTP codec are protected by a shortened binary BCH(62,29,6) code yielding 248 bits, while the 152 C2 bits are coded by a shortened binary BCH(62,38,4) code which also yields 248 bits. Four 62-bit codewords are used in each class, and are rectangularly interleaved to curtail error propagation across 20 ms speech frames. The 496 bits in one speech frame are transmitted during 20 ms, yielding a total transmission bit rate of 24.8 kbits/s.

5 Speech Transmission Systems

In our simulations we used the GMSK, $\frac{\pi}{4}$-DQPSK and 16-StQAM modems combined with both the unprotected low-complexity 32 kbps ADPCM codec (as in DECT and CT2) and the 13 kbps RPE-LTP codec with its FEC. Each modem had the option of either a low or a high complexity demodulator. Synchronous transmissions and perfect channel estimation were assumed in evaluating the relative performances of the systems listed in Table

1. Our results represent performance upper bounds, allowing relative performance comparisons under identical circumstances. The propagation environment was assumed to be Gaussian or Rayleigh. We presumed that the microcells were sufficiently small for the transmitted symbol rate considered that dispersion in the radio channels was a rare event. Returning to Table 1, the first column shows the system classification letter, the next

System	Modulator	Detector	FEC	Speech Codec	Complexity Order	Baud Rate (kBd)	User Bandwidth (kHz)	Min CSNR (dB)	
								AWGN	Rayleigh
A	GMSK	Viterbi	No	ADPCM	2	32	23.7	7	∞
B	GMSK	Freq. Discr.	No	ADPCM	1	32	23.7	21	31
C	$\frac{\pi}{4}$-DQPSK	MLH-CR	No	ADPCM	4	16	19.8	10	28
D	$\frac{\pi}{4}$-DQPSK	Differential	No	ADPCM	3	16	19.8	10	28
E	16-SQAM	MLH-CR	No	ADPCM	6	8	13.3	20	∞
F	16-SQAM	Differential	No	ADPCM	5	8	13.3	21	31
G	GMSK	Viterbi	BCH	RPE-LTP	8	24.8	18.4	1	15
H	GMSK	Freq. Discr.	BCH	RPE-LTP	7	24.8	18.4	8	18
I	$\frac{\pi}{4}$-DQPSK	MLH-CR	BCH	RPE-LTP	10	12.4	15.3	5	20
J	$\frac{\pi}{4}$-DQPSK	Differential	BCH	RPE-LTP	9	12.4	15.3	6	18
K	16-SQAM	MLH-CR	BCH	RPE-LTP	12	6.2	10.3	13	25
L	16-SQAM	Differential	BCH	RPE-LTP	11	6.2	10.3	16	24

Table 1: System Comparison

the modulation used, the third the demodulation scheme employed, the fourth the FEC scheme and the fifth the speech codec deployed. The sixth column gives the estimated relative order of the complexity of the schemes, where the most complex one having a complexity parameter of 12 is the 16-StQAM, MLH-CR, BCH, RPE-LTP arrangement, System K. The speech Baud rate and the TDMA user bandwidth are given next. Observe that the DQPSK and 16-StQAM Baud rate figures have been moderated by the actual bandwidth requirements computed by taking into account their filtering requirements. Explicitly, the 1 bit/symbol, GMSK, 2 bit/symbol $\frac{\pi}{4}$-DQPSK and 4 bit/symbol 16-StQAM modems have moderated spectral occupancy figures of 1.35, 1.62 and 2.4 bit/Hz, respectively. A signalling rate of 400 kBd was chosen for all our experiments, irrespective of the number of modulation levels, to provide a fair comparison for all the systems under identical propagation conditions. The corresponding 400 kBd systems have a total bandwidth of $400/1.35 = 296$ kHz, $2 \cdot 400/1.62 = 494$ kHz and $4 \cdot 400/2.4 = 667$ kHz, respectively. The last two columns of Table 1 list the minimum channel SNR values in dB required for the system to guarantee nearly unimpaired speech quality through Gaussian and Rayleigh channels. These are characteristics closely related to the system's robustness, and they will be derived later after the System Performance Section.

6 Speech Performance

The system performances apply to microcellular conditions. The carrier frequency was 2GHz, the data rate 400 kBd, and the mobile speed 15m/s. At 400 kBd in microcells the fading is flat and usually Rician. The transmission rate in DECT is 1152 kBd and even at this high rate it is not necessary to use FEC codecs or equalisers. The best and worst Rician channels are the Gaussian and Rayleigh fading channels, respectively, and we performed our simulations for these channels to obtain upper and lower bound performances. Our conditions of 2GHz, 400 kBd and 15m/s are arbitrary. They correspond to a fading pattern that can be obtained for a variety of different conditions, for example, at 900 MHz, 271 kBd and 23m/s. Our goal is to compare the objective speech performances of the systems defined in Table 1 when operating according to our standard conditions. In evaluating the speech performance we used both SSNR and CD, hence our conclusions are based on both. However, due to lack of space here we only present SSNR results. The interested reader is referred to reference [6] for full details.

6.1 Cordless Telecommunication Schemes

Cordless telecommunication (CT) systems [7] are typically used in office environments. For the schemes A-F in Table 1 we employed the low-complexity ADPCM codec without the use of FEC coding. Let us consider systems A and B, employing GMSK with $B_T = 0.3$ and the ADPCM speech codec. In error-free conditions this speech codec achieved a segmental SNR of 29 dB with our test speech signal. Figure 1 shows that unimpaired speech

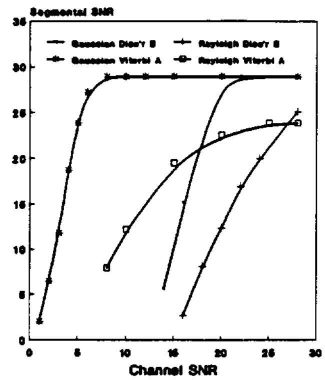

Figure 1: Speech Performance of Systems A & B: SSNR versus Channel SNR

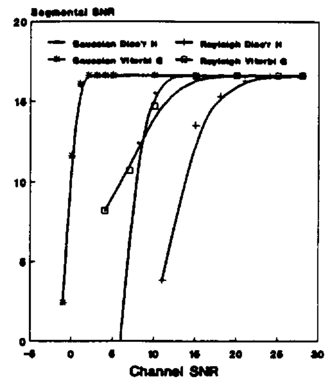

Figure 4: Speech Performance of Systems G & H: SSNR versus Channel SNR

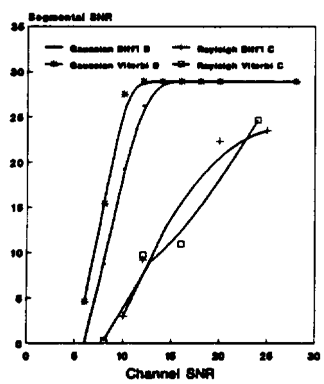

Figure 2: Speech Performance of Systems C & D: SSNR versus Channel SNR

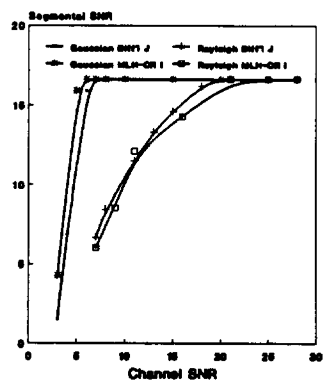

Figure 5: Speech Performance of Systems I & J: SSNR versus Channel SNR

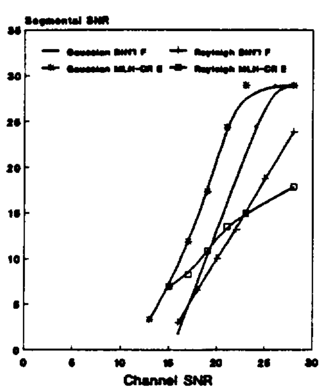

Figure 3: Speech Performance of Systems E & F: SSNR versus Channel SNR

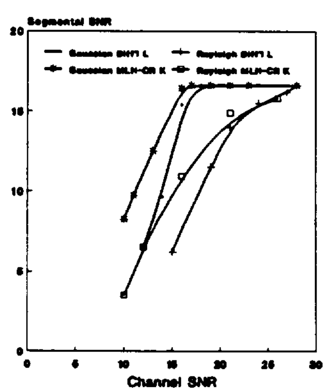

Figure 6: Speech Performance of Systems K & L: SSNR versus Channel SNR

quality was achieved for the Gaussian channel with the Viterbi detector (System A) for SNRs in excess of about 6 dB, whereas the discriminator (System B) required about 20 dB. For transmissions over the Rayleigh fading channel, the Viterbi detector never reached the error-free SSNR=29 dB, since the modem had an irreducible error rate at high SNRs. Despite its poorer performance at low SNRs, System B did not suffer from the problem of irreducible error rate. Consequently, above 25 dB channel SNR its speech SSNR performance was better than that achieved by System A for Rayleigh fading channels.

The objective speech quality of schemes C and D is shown in Figure 3 in terms of SSNR. The more complex MLH-CR of system C required some 2 dB lower channel SNR to achieve the same SSNR performance as the differential detector of system D for transmissions over the Gaussian channel, a trade-off not necessarily worthwhile. For Rayleigh fading channels the two detectors performed similarly, requiring about 25 dB channel SNR for unimpaired speech quality.

The performance of Systems E and F evaluated in terms of SSNR vs. channel SNR is given in Figure 4. Over the Gaussian channel system E requires about 22 dB channel SNR for unperturbed speech quality, while system F is some 3 dB less robust in terms of channel SNR. For transmissions over the Rayleigh channel the lower complexity system F recovers faster from low channel SNR conditions, as the SNR improves, reaching almost error-free conditions for SNRs of about 30 dB.

From our assessment of systems using the ADPCM speech codec, we conclude that for the Gaussian channel the GMSK system offers the best performance in terms of SNR, but has the lowest throughput. Doubling the number of bits per symbol, by moving to the DQPSK scheme, sacrifices 5 dB in channel SNR. Changing to Star QAM, with four bits per symbol, results in a further 12 dB penalty. For high channel SNR values the lower complexity frequency discriminator and differential decoder schemes are preferable to the more complex MLH-CR. This is because they have similar or, for high channel SNR values, superior BER and SSNR performances. However, if the channel was fading, the less complex discriminator provided better speech quality. For the multi-level modems the low-complexity differential approach was preferred. A minimum of about 25 dB channel SNR was required for unprotected, unimpaired speech quality. Our findings are summarised in the last two columns of Table 1.

6.2 Robust Systems

We now consider the more sophisticated schemes with FEC coding and the RPE-LTP speech coder. Again, the GMSK, $\frac{\pi}{4}$-DQPSK and 16-StQAM modems were used, each with either a simple or a complex detector.

Systems G and H consisted of a BCH coded RPE-LTP speech codec in conjunction with a GMSK modem using either Viterbi detection or a frequency discriminator. Because of the FEC, the SSNR reached the error-free unimpaired condition for all the scenarios, as shown in Figure 5. The poor performance of the frequency discriminator was apparent. The Viterbi detector required some 10 dB less channel SNR than the frequency discriminator for both types of channels for an unimpaired SSNR of about 16 dB. The discriminator operating over the Gaussian channel appeared to have an extremely inconvenient systematic distribution of errors, causing the FEC to be frequently overloaded. Surprisingly, this meant that the bursty errors over the Rayleigh channel were more easily accommodated. Thus the Viterbi detector was the most appropriate demodulator for this scheme.

Systems I and J incorporated RPE-LTP speech coding, BCH FEC and DQPSK with either MLH-CR or differential detection. Figure 6 shows that the SSNR objective speech performance of the two schemes was very similar, hence the less complex differential detector was favoured. For Gaussian channels an SNR of about 5 dB is sufficient for error-free speech reception, while the Rayleigh channel requires about 20 dB channel SNR.

The SSNR versus channel SNR performances of systems K and L are shown in Figure 7. The systems with differential or MLH-CR detection had similar overall SSNR performances for both Gaussian and Rayleigh channels, requiring some 15 and 25 dB channel SNRs to achieve unperturbed speech quality, respectively. However, the complexity advantage of the differential scheme made it more attractive.

As anticipated, the BCH-protected RPE-LTP schemes were more robust. For the GMSK modem the Viterbi detector was very attractive for both the channels we investigated, requiring a mere 10 dB channel SNR for unimpaired speech quality in the presence of Rayleigh fading. For the multi-level modems the more simple differential detector was preferred. These BCH-protected systems removed the modem's residual BER, yielding essentially unimpaired, error-free speech quality, which was rarely achieved without FEC coding via fading channels for unprotected CTs. The more robust, higher quality speech transmissions are attainable at a channel coded rate of 24.8 kbps, instead of the unprotected and less complex 32 kbps ADPCM transmissions, a trade-off available to system designers. Over non-fading Gaussian channels, typically 5 dB lower channel SNR was sufficient for similar speech quality, when using the FEC-protected robust schemes, assuming identical modems. Our results are summarised in the last two columns of Table 1.

7 Discussion

In our discourse we have examined a number of radio links formed with different sub-systems some of which are to be found in the second generation cordless telecommunication (CT) systems of CT2 and DECT, and in the cellular systems of GSM, DCS 1800 and IS54. We consider that third generation systems will need backward compatibility with these second generation systems, and therefore third generation transceivers will contain in firmware formats sub-systems that can be reconfigured under software control to emulate the second generation transceivers as well as those of the more advanced third generation PCNs [8].

We envisage an adaptive transceiver storing the software of all the sub-systems listed in Table 1 in order to provide services matching the predominant optimisation criteria. The number of modulation levels and detection algorithm, the activation of the BCH channel codec and loading of the appropriate speech codec algorithm and other transceiver features are programmable. As channel conditions worsen, or as the offered teletraffic increases, the transceiver reconfigures itself to match the optimisation criterion, which might be minimum power consumption, maximum robustness against channel errors or minimum transmission bandwidth requirement. The definition of these optimisation algorithms requires further research.

8 Acknowledgement

The authors are grateful to the Radiocommunications Agency for their support of this work.

References

[1] R. Steele, "Mobile Radio Communications", Pentech Press, London, 1992.

[2] W. T. Webb, L. Hanzo and R. Steele, "Bandwidth efficient QAM schemes for Rayleigh fading channels", IEE Proc Part I, Volume 138, No. 3, June 1991, pp. 169-175.

[3] W. Webb, L. Hanzo, R.A. Salami and R. Steele, "Does 16-QAM Provide an Alternative to a Half-Rate GSM Speech Codec?", Proc. IEEE Vehicular Technology Conference, St. Louis, U.S.A., 19-22 May 1991, pp. 511-516.

[4] L. Hanzo, R. Steele, P.M. Fortune: A Subband Coding, BCH Coding and 16-QAM System for Mobile Radio Communication, IEEE Tr. on VT, Nov 1990, Vol 39., No 4, pp 327-340.

[5] R. Steele, D. Twelves, L Hanzo: Effect of Cochannel Interference on Handover in Microcellular Mobile Radio, Electr. Let., 1989, Vol 25, No 20, pp 1329-1330

[6] J.E.B. Williams, L. Hanzo, R. Steele, J.C.S. Cheung: Adaptive terminals for mobile radio, submitted to IEE Proc. Part-I, 1992

[7] W. H. W. Tuttlebee (Ed.), "Cordless Telecommunications in Europe", Springer-Verlag, London, 1990.

[8] R.Steele, W. T. Webb: Variable rate QAM for data transmissions over mobile radio channels, Keynote paper, Wireless 91, Calgary Alberta, July 1991