BANDWIDTH-EFFICIENT QAM SCHEMES FOR RAYLEIGH FADING CHANNELS

*W. T. Webb, L. Hanzo and *R. Steele

University of Southampton, UK

*also with Multiple Access Communications. UK

1. Introduction

Quadrature Amplitude Modulation (QAM) is a bandwidth efficient transmission method for digital signals. It is expected that the available spectrum for the proposed Personal Communications Network (PCN) will soon be at a premium as the number of subscribers increases, and changing from binary modulation to QAM may significantly ease the problem. The severe amplitude and phase changes introduced by the fading channels, however, make low error transmission of QAM difficult to achieve, unless procedures are introduced at both the transmitter and the receiver to combat the fading.

In this discourse we describe ways to improve the Bit Error Rate (BER) of QAM by using various forms of coding on an increased symbol set. This means that the data throughput, symbol rate and transmission power are unaffected, although the transmitter and receiver are made considerably more complex. In our deliberations we consider Rayleigh, rather than the less severe Ricean fading channel to get worst case performance estimates of mobile radio communications.

2. QAM transmissions over Rayleigh fading channels

Digital communications using 16-level QAM signals and conventional receiver techniques have unacceptably high bit error rates (BER) in a Rayleigh fading environment [1]. The problem is the inability to track absolute phase during fades with the result that on emergence from a fade the phase locked loop (PLL) in the receiver locks onto a different quadrant than the required one [2]. Differential encoding can reduce this false phase locking problem, but the standard square QAM constellation suffers from possible false lock positions at 26 degrees and 53 degrees.

In order to overcome these deficiencies we introduce a constellation having no false lock positions. This constellation we call "Star QAM", and is essentially a twin 8 level Phase Shift Keying (PSK) constellation, as shown in Figure 1. This constellation does not have a minimum least free distance between points in the strict sense, but does allow efficient differential encoding and decoding methods to be used which go some way towards mitigating the effects of Rayleigh fading.

With PLLs some form of differential encoding is essential as the Rayleigh fading channel can introduce phase shifts in excess of 50 degrees between consecutive symbols, making it extremely difficult to establish an absolute phase reference. The encoding method is straightforward. Of the four bits in each symbol, b_1, b_2,b_3 and b_4,the first is differentially encoded onto the QAM phasor amplitude so that a "1" causes a change to the amplitude ring which was not used in the previous symbol, and a "0" causes the current symbol to be transmitted at the same amplitude as the previous symbol. The remaining three bits are differentially Gray encoded onto the phase. Decoding data is now reduced to a comparison test between the previous and current received symbols. Suppose we fix the transmitted rings at amplitude levels A_1 and A_2 as shown in Figure 1. Let the received phasor amplitudes be Z_t and Z_{t+1} at time t and t+1, respectively. The demodulator must identify whether there has been a significant change in amplitude in order to regenerate a logical "1" . The algorithm employed at the demodulator uses two adaptive thresholds to make its decision viz:- if

$$Z_{t+1} \geq \left(\frac{A_1 + A_2}{2}\right) Z_t \qquad (1)$$

or if

$$Z_{t+1} < \left(\frac{2}{A_1 + A_2}\right) Z_t \qquad (2)$$

then a significant change in amplitude is deemed to have occurred and bit b_1 is set to logical "1" at time t+1. Should both of Inequalities (1) and (2) fail to be satisfied, b_1 is assigned logical "0".

If the received symbol phases are θ_t and θ_{t+1} at time t and t+1, respectively, the demodulated angle is

$$\theta_{dem} = (\theta_{t+1} - \theta_t) \; mod \; 2\pi. \qquad (3)$$

This angle is then quantised to the nearest multiple of 45 degrees and a lookup table consulted to derive the remaining three output bits,b_2, b_3 and b_4. This system considerably improves the BERs compared to those for the square constellation because it eliminates long error bursts that occur when a false lock has been made. Of considerable importance is that with differential amplitude encoding there is no longer any need for AGC. This not only simplifies the circuit, but also removes errors caused by an inability of the AGC to follow the fading envelope.

Although we have removed the objectionable false locking characteristic of the square QAM constellation, the fading continues to cause problems because there are changes in channel amplitude and phase over consecutive symbol periods, which in general moves the differentially decoded phasors nearer to the decision boundaries. The most likely cause of error in the Star QAM system is when both the noise, and the change in the phasor's amplitude or phase due to fading, combine to drive the incoming signal level over a decision boundary. As the Rayleigh fading envelope is crudely predictable, particularily at low vehicular speeds, a correction factor may be applied to the incoming signal to compensate for the changes in the fading envelope over the last symbol period. However, we should not be too dependant on our previous observations as they will allow errors to propagate. Further, the system must be fast acting so that the sudden change from an amplitude decrease to an amplitude increase experienced at the bottom of a fade can immediately be detected and compensated. To overcome this problem a simple oversampling receiver is used. In this system n observations equally spaced in time are made per symbol period. When E_b/N_0 is low (below 30dB for 16 QAM) the first observation from the current symbol is compared with the last observation from the previous symbol period. This reduces the magnitude of any changes in the channel phase and amplitude and is akin to the advantage gained by transmitting at a higher rate without the problems of frequency selective fading and the loss of a large number of bits during a fade. At higher E_b/N_0 ratios the current phasor can be modified to compensate for the fading. This is done by finding the change in the incoming symbol phase and amplitude over the current symbol period. We do this by subtracting the phase at the end of the symbol from the phase at the beginning of the symbol, and subtracting the amplitude at the end of the symbol from the amplitude at the beginning of the symbol.

We use the same notation as before, but with the addition of subscript n to the symbols, to signify the n^{th} observation of the symbol at time t. Thus

$$Z_{t+1,diff} = Z_{t+1,n} - Z_{t+1,1} \quad (4)$$

where diff means differential, and

$$\theta_{t+1,diff} = \theta_{t+1,n} - \theta_{t+1,1}. \quad (5)$$

These changes can be used to extrapolate back from the first observation in the current symbol to the point in time when the last observation from the previous symbol was made.

To improve the accuracy of Inequalities (1) and (2) we replace Z_t by its last sample value, namely $Z_{t,n}$ and we use an estimate of Z_{t+1} which is formed by extrapolation as

$$Z_{t+1,ext} = Z_{t+1,1} - \frac{Z_{t+1,diff}}{n} \quad (6)$$

rather than Z_t. Notice that $Z_{t+1,1}$ is the first sample of Z_{t+1} and is closest in time to $Z_{t,n}$. By subtracting the average change in Z_{t+1} over a symbol period, i.e. by $\frac{Z_{t+1,diff}}{n}$ from $Z_{t+1,1}$, we obtain an estimate of Z_{t+1} had it been transmitted at time t. This is beneficial as there may have been significant amplitude and phase changes between time t and t+1.

After determining bit b_1 with the aid of Inequalities (1) and (2) with the modified Z_t and Z_{t+1} values, we determine bits b_2, b_3 and b_4 by formulating

$$\theta_{dem} = (\theta_{t+1,ext} - \theta_{t,n}) \ mod \ 2\pi. \quad (7)$$

where $\theta_{t+1,ext}$ and $\theta_{t,n}$ are the phase angles associated with $Z_{t+1,ext}$ and $Z_{t,n}$. Again θ_{dem} is quantised and used to address a lookup table which provides values of b_2, b_3 and b_4.

This oversampling method performs worse at low E_b/N_0 ratios because the noise tends to render the predictions inaccurate. In practice oversampling is only worth implementing for E_b/N_0 ratios in excess of 30dB, where substantial improvements can be made. Previous QAM systems tend to exhibit a residual BER at high E_b/N_0 ratios due to the rapidly changing Rayleigh channel, rather than the additive noise. With this system the residual BER is reduced by approximately two orders of magnitude.

2.1. Simulation results

Pseudo-random data generated at 64kb/s was organised into 4 bit words and mapped onto the 16-level Star QAM constellation shown in Figure 1. The QAM carrier was 1.9GHz, and the mobile's speed was 30mph. The channel exhibited Rayleigh fading with additive white Gaussian noise (AWGN). The receiver operated as described above, with the oversampling ratio set at n=8. The variation of BER with E_b/N_0 is curve (c) in Figure 2. Also shown as bench markers are the results for the square QAM constellation (curve (a)) and the Star QAM with differential encoding (curve (b)) both over a Rayleigh fading channel, and the Star QAM with differential encoding over an AWGN channel (curve (e)). The performance of the system having differential encoding and oversampling (curve(c)) can be considerably enhanced by the use of spatial diversity, where two antennas and receiver circuits are used. For these simulations, switched diversity was used, whereby for each phasor received, the receiver with the incoming phasor of largest magnitude was selected. Both receivers must have their own differential decoders. Curve (d) shows the performance of the system.

As can be seen, a very substantial improvement in BER has been obtained over the conventional square QAM by using Star QAM, and at high E_b/N_0 ratios our oversampling technique gave a further significant gain. These changes to the basic QAM were achieved for only a small increase in receiver complexity. By introducing second-order switched diversity the system operated with a channel SNR of only 5dBs above that for an AWGN channel for a BER of 10^{-3}.

3. Block Coding

Block coding was employed by expanding the signal set to cope with the extra bits required for the code. A 2/3 coding rate was considered to be appropriate, causing the number of QAM levels to increase from 16 to 64. An extension of the Star constellation was selected having 4 amplitude rings with 16 points equispaced on each ring.

For block codes to perform well in the Rayleigh fading environment it is necessary to add interleaving to the system in order to randomly distribute the errors in time. The block code and the interleaving process introduce a delay, and the maximium permissible delay depends on the type of information to be transmitted. The integrity constraints of computer data transmissions are at least three orders of magnitude higher than those of digital speech transmissions. Fortunately, data channels can accept longer interleaving delays which allow us to effectively randomise the bursty error statistics of the Rayleigh fading channel. This considerably increases the ability of the forward error correction (FEC) decoder to decrease the BER. For each transmission rate, propagation frequency, vehicular speed and interleaver algorithm there is a minimum interleaving depth or delay to transform the BER statistics of a fading channel into a good approximation of those encountered in a Gaussian channel.For a specific combination of system parameters, an interleaved FEC codeword has to overbridge the channel fades. For the emerging personal communication network(PCN) this is a difficult requirement, since a high proportion of mobile subscribers are moving at walking pace and therefore may spend a relatively long time in a deep fade. One proposed solution which is to be incorporated into the GSM system is the use of frequency hopping. This sets an approximate limit on the length of time that can be spent in a fade. Should a mobile be stationary and in a fade, then the probability of hopping to another fade is very remote. Thus hopping produces a similar effect to setting a lower limit on the mobile speed and reduces the interleaver depth required. Given an appropriate hop rate and a powerful enough interleaver we can randomize the channel errors sufficiently for channel coding to work at any mobile speed. The results we present here would apply to all mobile speeds with the proviso that frequency hopping is carried out often enough to allow the interleaver to work correctly.

For the propagation frequencies (1.7 to 1.9 GHz) used for PCN there is approximately one fade every 10 cm distance. If the data transmission rate is 64 kbit/s, yielding a signalling rate of 16 ksamples/s for the uncoded 16-level QAM,then when the mobile station is travelling at 30mph, there are approximately 200 QAM samples transmitted between two deep fades. Interleaving over three fades randomises the bursty error statistics and has a delay of approximately 600 QAM samples corresponding to less than 40 ms. By increasing the interleaving depth to much higher values as are acceptable for data transmission the BER is further improved. Indeed, for infinite delays the performance of the memoryless AWGN channel is achieved and is represented by curve(e) in Figure 2.

We consider Bose-Chaudhuri-Hocquenghem (BCH) block codes to have favourable properties for PCN transmissions. They can correct both random and bursty errors and also their error detection capability allows the BCH decoder to know when the received codeword contains more errors than the correcting power of the code can cope with. Provided a systematic BCH code is used, the information part of the coded word can be separated from the parity bits so that in code overload the information bits are not corrupted by the decoding process.

A special subclass of BCH codes is the maximum minimum distance Reed-Solomon(RS) codes. These codes operate on non-binary symbols and have identical error-locator and symbol fields. The non-binary RS codes are optimum due to their maximum distance properties, and may also be sufficiently long to overbridge channel fades. However, they are more complex than binary BCH codes. To explore both ends of the complexity/performance trade-off we selected an extremely long RS code, the $RS(252,168,42)$ code over $GF(256)$ using 8 bit symbols, the moderately long $RS(44,30,7)$ code over $GF(64)$, as well as the short binary $BCH(63,45,3)$ code. The nomenclature used here is $RS(m,n,k)$ where m is the number of encoded symbols, n is the number of information symbols and k is the number of corrected symbols in a codeword.

4. Bandwidth efficient coding results

The BER using block coding decreased dramatically compared to our uncoded QAM schemes as shown in Figure 3. The $BCH(63,45,3)$ code and the $RS(44,30,7)$ code have nearly identical performances, inspite of the considerably higher block length and complexity of the RS code. The $RS(252,168,42)$ code offers an extra 2 dB coding gain for a large increase in complexity. We therefore recommend the BCH code which provides virtually error-free communications for E_b/N_0 in excess of 30 dB, a value that may be realisable in the small microcells to be ultimately found in a fully developed PCN. By using error correction coding and compensating for the increased bit rate by using higher level modulation, we are able to provide transmissions of higher integrity, but with identical bandwidth compared to an uncoded system.

5. Overall coding strategy

We now consider 16 and 64-level QAM schemes which are error protected, where the error coding decreases the useful information transmission rate whilst increasing the integrity of the transmitted data. The 64-level QAM having a 2/3 rate code had its code rate reduced to 1/2, decreasing the information content rate by 1/4. We selected this overall 1/2 coding rate as it is widely used in mobile radio [3]. A further coding rate reduction does not bring corresponding coding gains and squanders channel capacity which cannot be compensated for by using more modulation levels. We also employed a 3/4 rate code with the 16-level QAM in order to provide the same overall bit rate as the coded 64-level QAM.

Our results are depicted in Figure 4. The uncoded 16-level QAM and the 64-level QAM RS (44,30,7) curves are repeated for comparison along with the 64-level QAM RS (60,30,15) and 16-level QAM RS (60,44,12) arrangements. There is a consistent and remarkable improvement in both the 16-level QAM and 64-level QAM performance. For E_b/N_0 in excess of 25 dB there is an almost 5 dB extra coding gain improvement due to the stronger RS code in case of the 64-level QAM scheme. By employing error correction coding, the performance of the 16-level QAM arrangement is dramatically improved and becomes superior to that of the 1/2-rate coded 64-level QAM system. Depending on the integrity required, an E_b/N_0 value of 25 dB is sufficient for reliable signalling via the 16-level QAM RS (60,44,12) system.

6. Summary and Conclusions

We have proposed various signal processing techniques to facilitate reliable QAM transmissions in Rayleigh fading environments. Our findings indicate that these bandwidth efficient modulation schemes can be used for the emerging high traffic density PCNs. The well-known 16-level square constellation was found to be unsuitable for the PCN environment. The sub-optimal Star scheme with differential encoding and oversampling signal estimation dramatically improved the BER performance, rendering the channel appropriate for speech transmissions. The lower BERs essential for data transmissions were achieved by expanding the 16-level QAM signal set to 64 levels and using the extra channel capacity acquired for error correction coding. For E_b/N_0 values in excess of 25dB the performance of the coded 64-level QAM was superior to that of the 16-level QAM and at values of E_b/N_0 in excess of 30dB it was virtually error free. When the overall coding rate was lowered to allow the 16-level QAM scheme to incorporate a 3/4 rate RS code and the 64-level QAM system had a 1/2 rate RS code, the 16-QAM arrangement out-performed the more complex 64-level scheme. A virtually error free performance was achieved above E_b/N_0=25dB - a value that should be easy to achieve in the small microcells expected in PCN.

References

[1] Transmission of SBC Speech Via 16-level QAM Over Mobile Radio Channels - P.M.Fortune, L.Hanzo, R.Steele, Proc. of GLOBECOM 88, pp.832-836

[2] Carrier Recovery for 16-level QAM in mobile radio - E.Issman,W.T.Webb. To be published in IEE colloquium on Multi-level modulation techniques for point-to-point and mobile radio, March 1990.

[3] Channel coding for satellite mobile channels - K.H.H Wong,L.Hanzo, R.Steele. Int Journal of Satellite Communications, Vol 7,pp143-163,1989.

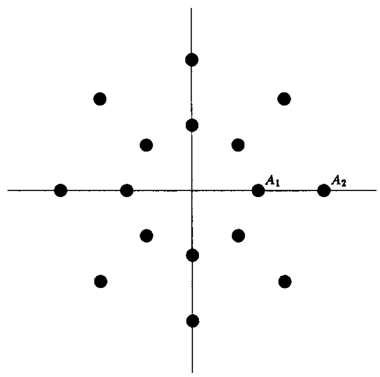

Figure 1

Star 16 QAM Constellation at Transmitter

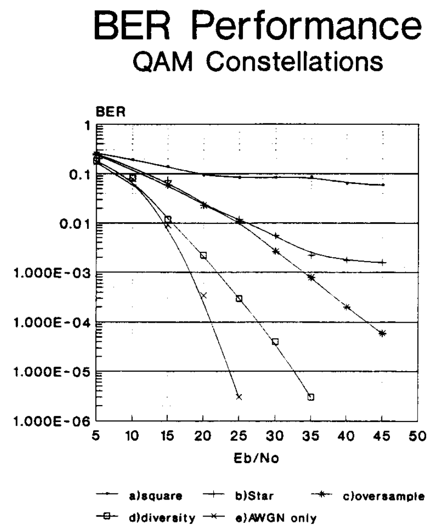

Figure 2

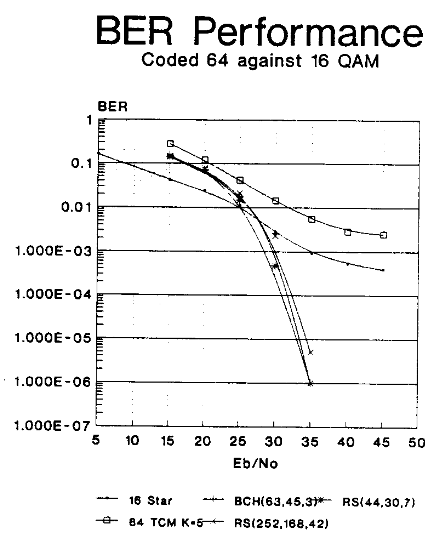

Figure 3

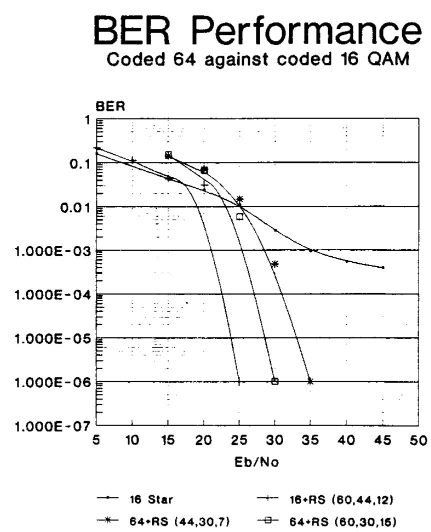

Figure 4