An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels

James K. Cavers, *Member, IEEE*

Abstract—**Proposals have appeared recently for the use of pilot symbols to mitigate the effects of rapid fading in mobile communications. Unlike the more familiar pilot tone systems, pilot symbol assisted modulation (PSAM) does not affect the transmitted pulse shape or the peak-to-average power ratio, and implementation is straightforward. This paper puts PSAM on a solid analytical basis, a feature missing from previous work. It presents closed form expressions for the BER in BPSK and QPSK, for a tight upper bound on SER in 16QAM, and for the optimized receiver coefficients. The error rates obtained are lower than for differential detection for any combination of *SNR* and Doppler spread, and the performance is within 1 dB of a perfect reference system under slow fading conditions, and within 3 dB when the Doppler spread is 5% of the symbol rate.**

I. Introduction

RAPID fading is a central problem in digital mobile communications. It degrades the bit error rate (BER), and frequently introduces an irreducible BER, or error floor. It also inhibits the use of multilevel modulation formats, with their greater spectral efficiency.

The use of a pilot tone to mitigate the effects of fading has been explored by several authors [1]–[4]. The tone provides the receiver with an explicit amplitude and phase reference for detection, and thereby suppresses the error floor. The question of where in the spectrum to locate the tone is a difficult one. Perhaps the best known solution is transparent tone-in-band, or TTIB [1], [2], [4]. Although it is a general solution, it requires relatively complex signal processing and results in an increased peak-to-average power ratio.

Recently, pilot symbol assisted modulation (PSAM) has been proposed [5]–[7] as an alternative. The transmitter periodically inserts known symbols, from which the receiver derives its amplitude and phase reference. Like pilot tone modulation, PSAM suppresses the error floor and enables multilevel modulation. However, it does so with no change to the transmitted pulse shape or peak to average power ratio. Processing at the transmitter and receiver is also simpler than with TTIB.

Previous studies of PSAM [5]–[7] were based on simulation and experimental implementations. Although they demonstrated feasibility, they did not provide the performance analysis needed before their results can be generalized. The present paper supplies the missing analysis. It includes closed form results for the BER in BPSK and QPSK, and for a tight upper bound on the symbol error rate in 16QAM, as well as expressions for the optimum interpolator at the receiver. No numerical integrations are required, though optimization requires solution of the normal equations. The performance of the optimized system is excellent, within 3 dB of an unachievable perfect reference system, even when the Doppler spread is 5% of the symbol rate. At 1% Doppler, the loss is only 1 dB, and it can be further reduced at lower Doppler values.

Manuscript received December 22, 1989; revised July 19, 1990 and November 6, 1990. This work was supported by the BC Advanced Systems Institute through a Senior Fellowship.

The author is with Communications Science Laboratory, School of Engineering Science, Simon Fraser University, Burnaby, BC Canada V5A 1S6.

IEEE Log Number 9103228.

II. System Model

A. General Description

A PSAM system can be represented by the block diagram of Fig. 1. Known symbols are inserted periodically into the data sequence prior to pulse shaping, and the composite signal is transmitted in the usual way over a channel characterized by flat fading and additive noise. The resulting frame structure is shown in Fig. 2. For simplicity, we will assume Rayleigh fading, though the extension to Rice fading is straightforward.

After matched filter detection, the receiver splits the per-symbol samples into two streams. The reference branch decimates the samples to extract those due to the pilot symbols, and interpolates them to form an estimate of the channel state. It then scales and rotates a reference decision grid (shown in Fig. 1 for 16QAM) with the estimate, and feeds the modified decision boundaries to the data branch. Although this model is equivalent to the usual formulation [1]–[7], in which the receiver normalizes the data branch by dividing it with the estimate from the reference branch, it is more mathematically tractable. Finally, the data branch contains a delay to compensate for the interpolation delay in the reference branch.

There are two immediate consequences of the technique. The first is delay in the receiver; in order to obtain enough pilot samples for a good channel estimate, the receiver must wait and buffer samples for several frames. The second is that the interpolation coefficients used to form the channel estimate depend on the position within the frame of the sample whose identity is to be determined.

B. Transmitted Signal

The transmitted signal has a complex envelope given by

$$s(t) = A \sum_{k=-\infty}^{\infty} b(k)\, p(t - kT) \qquad (1)$$

0018-9545/91$01.00

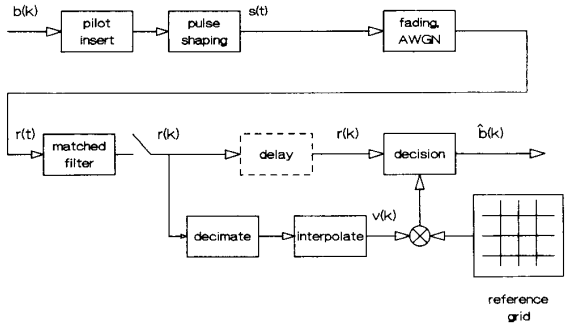


Fig. 1. Organization of a pilot symbol system.

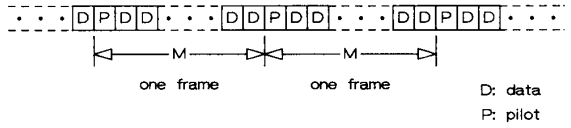


Fig. 2. Transmitted frame structure.

where T is the symbol duration, $b(k)$ is the kth symbol value for BPSK, QPSK, or a general QAM, A is an amplitude factor, and $p(t)$ is a unit energy pulse:

$$\int_{-\infty}^{\infty} |p(t)|^2 \, dt = 1. \tag{2}$$

The symbols are formatted into frames of length M in which the pilot symbols at times $i = kM$ have a known value $\tilde{b}$. As noted in [6], it is better to randomize the pilot values to avoid transmitting tones, but this does not affect the analysis.

C. Fading Channel Effects

The fading channel output is given by

$$r_c(t) = c(t)s(t) + n_c(t) \tag{3}$$

in which $n_c(t)$ is AWGN with power spectral density N_o in both real and imaginary components. The channel's complex gain $c(t)$ incorporates both fading and frequency offset:

$$c(t) = \exp(j2\pi f_0 t)\, g(t) \tag{4}$$

where f_o is the residual frequency offset after AFC, and $g(t)$ is the complex Gaussian fading process with variance σ_g^2 and Doppler spread f_D. Its autocorrelation function can be written:

$$R_c(\tau) = \sigma_g^2 \tilde{R}_c(\tau) \tag{5}$$

where $\tilde{R}_c(\tau)$ is the unit power equivalent. Although the analysis is not limited to a specific autocorrelation function, in the numerical results presented later we assume the normalized version to be:

$$\tilde{R}_c(\tau) = \exp(j2\pi f_0 \tau)\, J_0(2\pi f_D \tau) \tag{6}$$

which has the U-shaped power spectrum characteristic of isotropic scattering [8].

D. Receiver Processing

The receiver detects the pulses using a matched filter with impulse response $p^*(-t)/\sqrt{N_o}$, where the noise normalization is for later notational convenience. We assume no intersymbol interference, which implies Nyquist I pulses with no receiver timing error and fading slow enough not to cause appreciable distortion of the pulses. As a point of interest, a 5% Doppler shift results in only 1% maximum ISI in the case of 25% rolloff spectral raised cosine pulses.

It is convenient at this point to introduce conventional received SNR measures. Define the ratio q of pilot power to data power:

$$q = |\tilde{b}|^2/(M-1)E[|b|^2] \tag{7}$$

where the expectation is taken over data symbols only. The expected total energy received over a frame is:

$$E_f = \sigma_g^2 A^2 (|\tilde{b}|^2 + (M-1)E[|b|^2]). \tag{8}$$

With n bits per symbol, and $M-1$ data symbols in a frame, the received energy per data bit is:

$$E_b = \frac{\sigma_g^2 A^2}{n(M-1)} |\tilde{b}|^2 \left(\frac{1+q}{q}\right). \tag{9}$$

The symbol-spaced samples $r(kT)$ of matched filter output are given by

$$r(k) = \frac{Ac(kT)}{\sqrt{N_o}} b(k) + n(k) = u(k)b(k) + n(k) \tag{10}$$

where the Gaussian noise samples $n(k)$ are white with unit variance and the symbol gain $u(k)$, as defined here, has variance:

$$\sigma_u^2 = \frac{1}{2}E[|u|^2] = \gamma_b \frac{n(M-1)}{|\tilde{b}|^2} \frac{q}{1+q} \tag{11}$$

in which $\gamma_b = E_b/N_o$.

Without loss of generality, take $b(0)$ be a pilot symbol, and consider the detection of $b(k)$, $-\lfloor M/2 \rfloor \le k \le \lfloor (M-1)/2 \rfloor$. The channel state estimator prepares an estimate of the symbol gain $u(k)$ in (10) using the K nearest pilot samples:

$$v(k) = \sum_{i=-\lfloor K/2 \rfloor}^{\lfloor K/2 \rfloor} h^*(i,k) r(iM). \tag{12}$$

Note that the coefficients $h(i,k)$ depend on the position k within the frame. The estimation error is denoted by $e(k)$, so that

$$u(k) = v(k) + e(k). \tag{13}$$

E. Optimum Interpolation

Previous descriptions of PSAM used plausible, but arbitrary, interpolation filters: a lowpass in [6] and an approximately Gaussian filter in [7]. Here we use a Wiener filter to

minimize the variance of the estimation error e, in order to discover the limits of the technique.

Define the length K column vector $\boldsymbol{r}$ as the set of pilot samples $r(iM)$, $-\lfloor K/2 \rfloor \le i \le \lfloor K/2 \rfloor$ and $\boldsymbol{h}(k)$ as the corresponding set of coefficients for the kth position in the frame. Then the estimated symbol gain is given by

$$v(k) = \boldsymbol{h}^{\dagger}(k)\boldsymbol{r} \tag{14}$$

where the dagger denotes conjugate transpose, and the interpolation error is given by

$$e(k) = u(k) - \boldsymbol{h}^{\dagger}(k)\boldsymbol{r}. \tag{15}$$

The $K \times K$ autocorrelation matrix is:

$$\boldsymbol{R} = \frac{1}{2}E[\boldsymbol{r}\boldsymbol{r}^{\dagger}] \tag{16}$$

and the $M-1$ length K covariance vectors are:

$$\boldsymbol{w}(k) = \frac{1}{2}E[u^{*}(k)\boldsymbol{r}]. \tag{17}$$

Then the variances of interest are given by

$$\begin{aligned} \sigma_v^2(k) &= \boldsymbol{h}^{\dagger}(k)\boldsymbol{R}\boldsymbol{h}(k) \\ \sigma_{uv}^2(k) &= \boldsymbol{w}^{\dagger}\boldsymbol{h}(k) \\ \sigma_e^2(k) &= \sigma_u^2 - 2\,\mathrm{Re}\left[\sigma_{uv}^2(k)\right] + \sigma_v^2(k). \end{aligned} \tag{18}$$

As is well known (e.g., [9]), an optimum estimate results if the coefficient vector $\boldsymbol{h}$ is selected to satisfy the normal equations:

$$\boldsymbol{R}\boldsymbol{h}(k) = \boldsymbol{w}(k). \tag{19}$$

The estimation error variance then attains its minimum value, and we have:

$$\begin{aligned} \sigma_v^2(k) &= \sigma_{uv}^2(k) = \boldsymbol{w}^{\dagger}(k)\boldsymbol{R}^{-1}\boldsymbol{w}(k) \\ \sigma_e^2(k) &= \sigma_u^2 - \sigma_v^2. \end{aligned} \tag{20}$$

Moreover, the estimation error is uncorrelated with $r(iM)$ and $v(k)$.

In the case at hand, we have explicit expressions for components of the arrays $\boldsymbol{R}$ and $\boldsymbol{w}(k)$. From (10), (11), and (16), we have:

$$\boldsymbol{R}_{ik} = \gamma_b n(M-1)\frac{q}{1+q}\tilde{R}_c((i-k)MT) + \delta_{ik} \tag{21}$$

where δ_{ik} is the Kronecker delta, and from (7), (10), (11), and (17) we have:

$$\boldsymbol{w}_i(k) = \gamma_b \frac{n(M-1)}{\tilde{b}^{*}}\frac{q}{1+q}\tilde{R}_c((iM-k)T). \tag{22}$$

Calculation of the optimum interpolation coefficients and the resulting variances is straightforward.

III. BPSK and QPSK Signaling

A. Calculation of BER

First, consider BPSK. The pilot symbols have the value $\tilde{b} = 1$, and the data symbols are $b(k) = \pm 1$. The receiver makes decisions by phase correcting the matched filter output with the channel estimate and comparing the result to zero:

$$\mathrm{Re}\left[r(k)v^{*}(k)\right] : 0 \ (\mathrm{BPSK}). \tag{23}$$

If the transmitted bit is $+1$, then an error is made if the real part of the decision variable (23) is negative. Since $r(k)$ and $v(k)$ are correlated, zero mean, Gaussian random variables, we can use a standard result for the probability of this event [9, App. 4B]:

$$P_b(k) = \frac{1}{2}\left(1 - \sqrt{\mathrm{Re}\left[\rho(k)\right]^2/\left(1 - \mathrm{Im}\left[\rho(k)\right]^2\right)}\right) \tag{24}$$

where $\rho(k)$ is the correlation coefficient:

$$\rho(k) = \sigma_{rv}^2(k)/\sigma_r\sigma_v(k). \tag{25}$$

Since $b(k) = 1$, we have $\sigma_{rv}^2(k) = \sigma_{uv}^2(k)$ and $\sigma_r^2 = \sigma_u^2 + 1$. For an arbitrarily selected filter $\boldsymbol{h}(k)$, therefore, the correlation coefficient is given by

$$\rho(k) = \frac{\boldsymbol{w}^{\dagger}(k)\boldsymbol{h}(k)}{\sqrt{(\sigma_u^2+1)\boldsymbol{h}^{\dagger}(k)\boldsymbol{R}\boldsymbol{h}(k)}} \ (\mathrm{BPSK, \ arbitrary} \ \boldsymbol{h}(k)) \tag{26}$$

where σ_u^2 is given by (11). In the case of the optimum filter (19), the correlation coefficient is real:

$$\rho(k) = \frac{\boldsymbol{w}^{\dagger}(k)\boldsymbol{R}^{-1}\boldsymbol{w}(k)}{\sqrt{(\sigma_u^2+1)\boldsymbol{w}^{\dagger}(k)\boldsymbol{R}^{-1}\boldsymbol{w}(k)}} \ (\mathrm{BPSK, \ optimum} \ \boldsymbol{h}(k)) \tag{27}$$

and the error probability becomes:

$$P_b(k) = (1 - \rho(k))/2 \ (\mathrm{BPSK, \ optimum} \ \boldsymbol{h}(k)). \tag{28}$$

Next, consider QPSK, in which the data symbols are $b(k) = \pm 1 \pm j$. Although the pilot symbol would be selected from this set in practice, the resulting $\pi/4$ rotation in $v(k)$ must be corrected before use. To simplify the notation, therefore, we will consider the pilot symbols to be real, with the value $\tilde{b} = \sqrt{2}$. The decision criteria are then:

$$\mathrm{Re}\left[r(k)v^{*}(k)\right] : 0, \ \mathrm{Im}\left[r(k)v^{*}(k)\right] : 0 \ (\mathrm{QPSK}). \tag{29}$$

We will evaluate the BER for the real component only, since it is the same for the imaginary component. However, if σ_{uv}^2 is complex, the resulting phase bias means we have to average the error probabilities for the two cases $b(k) = \pm 1 + j$, since they differ.

The BER calculation is much like that for BPSK, with the change that $\sigma_{rv}^2 = \sigma_{uv}^2 b(k)$, so that from (18):

$$\rho(k) = \frac{\boldsymbol{w}^{\dagger}(k)\boldsymbol{h}(k)b(k)}{\sqrt{(2\sigma_u^2+1)\boldsymbol{h}^{\dagger}(k)\boldsymbol{R}\boldsymbol{h}(k)}} \ (\mathrm{QPSK, \ arbitrary} \ \boldsymbol{h}(k)). \tag{30}$$

with a similar change for optimum $\boldsymbol{h}(k)$ (30). Note that the definition of $\boldsymbol{w}(k)$ (17) makes its value, and that of $\boldsymbol{h}(k)$, different from the BPSK case. Finally, the BER is given by (24), since $\rho(k)$ is complex.

A similar analysis can be applied to M-PSK systems with more than four phases [9, App. 7A]. Again, the performance is determined by a correlation coefficient.

B. Effect of Interpolation Position

The position of a symbol within the frame determines the estimation error, and hence the BER. Numerical experimentation, though, showed a small position dependence, observable only at very high SNR, where AWGN did not mask the effect. For example, for $K = 7$, in an $M = 10$ symbol frame, with $f_DT = 0.01$ and $\gamma_b = 50$ dB, there was only a 0.8% variation in BER across the frame. For a very short filter ($K = 3$), fast fading ($f_DT = 0.05$), and 50 dB SNR we found a factor of 3 variation across a 5-symbol frame. However, for normal operating parameters and K greater than 7, we determined that position dependence is negligible. Subsequent sections evaluate the BER only for $k = 1$, that is, for the symbol immediately following the pilot.

C. Effect of Pilot Symbol Spacing

The frame size, or pilot spacing, M, has an optimum value, which represents a trade-off between wasting energy in unnecessary pilot symbols and not sampling the fading process often enough for good estimation. Fig. 3 illustrates the effect for BPSK at $\gamma_b = 30$ dB, with a variety of fade rates, and a 63 coefficient filter optimized at every point on the graph. Not surprisingly, the BER rises steeply when the frame size causes sampling to fall below the Nyquist rate, i.e., when $M < 1/2f_DT$.

We selected a frame of $M = 7$ symbols as the benchmark for the rest of the paper; although it represents a 14% loss of capacity, it does accommodate fade rates up to 5% of the symbol rate. If slower fade rates are expected—for example, 100 Hz Doppler in a 16 ksymbol/s system gives $f_DT = 6.25 \times 10^{-3}$—then much larger frames can be adopted, to reduce the loss of capacity.

The loss of capacity is somewhat less than in TTIB [1], a well-known pilot tone technique. TTIB requires a bandwidth increase of $2f_D$ for the tone, plus an overlap region of comparable size for synchronization, for a total of about $4f_D$. With symbol spacing T, this is a fractional increase in bandwidth of about $4f_DT$. In comparison, as we saw above, PSAM requires a fractional increase of $1/M$, or about $2f_DT$.

As noted earlier, there is some startup delay in PSAM. The receiver must discard several data symbols at the start of each transmission, until it has accumulated enough pilot symbols for useful interpolation. There are about $KM/2$ such wasted data symbols—about 39 for the values $M = 7$ and $K = 11$ we have adopted as a benchmark. This nevertheless compares favorably with TTIB, which has been shown previously [10] to require on the order of 150 symbols to acquire phase lock at the start of a transmission.

The delay of $KM/2$ symbols also has a potentially adverse effect on speech communication. However, for QPSK and an 8 kb/s codec, the 39-symbol delay in our benchmark system represents only 10 ms of speech. For many applications, the delay can be ignored.

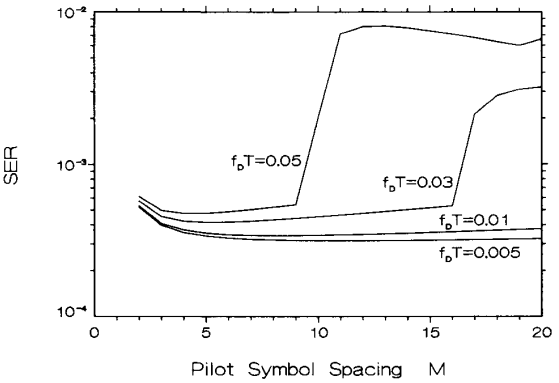


Fig. 3. Effect of frame size on BPSK ($\gamma_b = 30$ dB, $K = 63$).

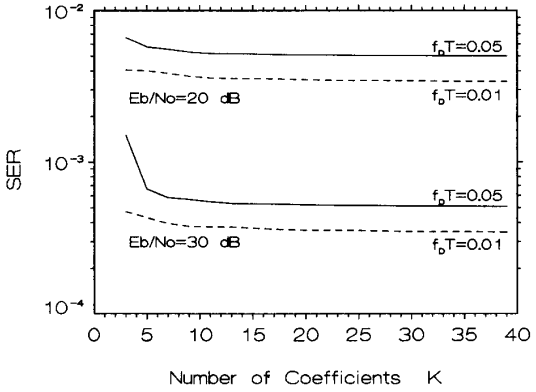


Fig. 4. Effect of interpolator size on BPSK ($M = 7$).

D. Effect of Interpolator Size

The number of interpolator coefficients is a significant issue, as it affects delay and computational load, as well as *BER*. Fig. 4 shows the dependence on K for BPSK and a seven-symbol frame. Again, the coefficients are optimized at every point. The improvement beyond five or ten coefficients is very slight. In the remaining sections, we use $K = 11$, though a reduction even to five coefficients would not cause serious harm.

E. Performance with Optimized Coefficients

Fig. 5 shows BER curves for PSAM BPSK with a seven-symbol frame and 11 coefficient interpolation. The coefficients are optimized at every point. We see that for 0 Hz Doppler, PSAM is only ¼ dB poorer than the unachievable coherent BPSK. Of the loss, 0.7 dB is attributable to the pilot symbols themselves, which are unnecessarily frequent for $f_DT = 0$, and the remainder is due to the limit of 11 coefficients in the interpolation. Even at 5% Doppler, the loss compared with coherent BPSK is only 3.5 dB and, as hoped, there is no error floor, at least in the useful range.

Comparison with DPSK, as calculated from expressions given in [9], shows that PSAM is better than DPSK at all values of SNR and Doppler spread. DPSK, of course, develops a significant error floor at large values of Doppler.

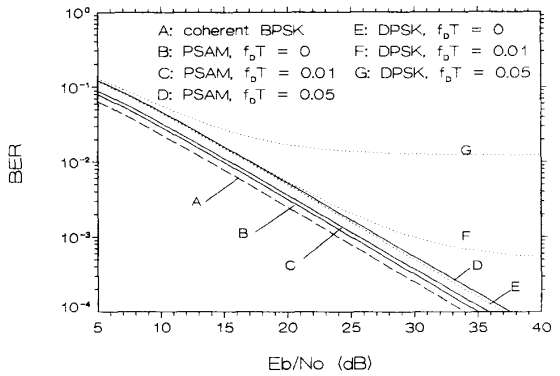


Fig. 5. BER performance of optimized PSAM for BPSK.

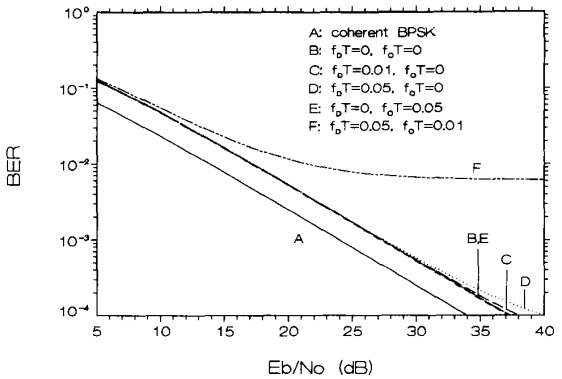


Fig. 6. PSAM BPSK: Mismatched with coefficients optimized for γ_b = 20 dB, $f_oT = 0$, $f_DT = 0.05$.

Comparison with the pilot tone based TTIB is also interesting. It has been shown previously [10] that the phase lock required by TTIB when used for data forces differential encoding, and hence a performance penalty of 3 dB. In contrast, PSAM produces an absolute phase reference, and therefore does not need differential encoding.

The frequency offset f_o has no effect on PSAM BER. By assumption, f_oT and f_DT are both small enough not to introduce intersymbol interference by distorting the pulses, and the optimized coefficients simply introduce compensating phase shifts.

The performance of QPSK has not been illustrated. Its BER is within 2% of that of BPSK for error rates of 10^{-2} or less, as both numerical evaluation and series expansions demonstrate.

F. Effect of Mismatch

The results of Section III-E above are somewhat unrealistic, in that the interpolator coefficients are optimum at every point on the graph. In practice, we would either design an adaptive interpolator or, more likely, optimize for one operating point, and use those coefficients even if conditions change.

Fig. 6 illustrates one such trial, in which the coefficients are optimized for γ_b = 20 dB, 5% Doppler, and no frequency offset. The result is that lower values of Doppler give no improvement in performance over most of the useful range. Even frequency offset, acting alone, had a negligible effect. However, the curve for combined 5% Doppler and 1% frequency offset shows a very substantial error floor, caused by the fact that the total bandwidth exceeds that of the interpolation filter. We have a clear warning that the filters must be designed for worst case conditions. This observation also applies to pilot tone systems, of course.

IV. 16QAM Signaling

A. Calculation of SER

In 16QAM we use symbols in which both real and imaginary parts of $b(k)$ take on the values -3, -1, 1, 3, and the expected value $E[|b|^2] = 10$. In the absence of fading, we would use the following six decision boundaries: $\mathrm{Re}[r(k)] = 0$, $\mathrm{Im}[r(k)] = 0$, $\mathrm{Re}[r(k) \pm 2] = 0$, and $\mathrm{Im}[r(k) \pm 2j] = 0$. As shown in Fig. 7, however, fading scales and rotates the signal constellation. The receiver compensates by scaling and rotating the decision boundaries, using the channel estimate $v(k)$. The six decision boundaries become:

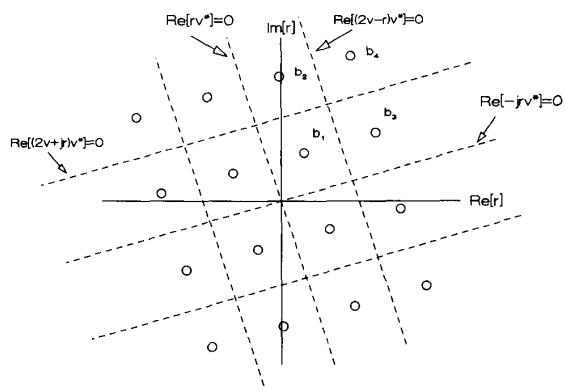


Fig. 7. 16QAM constellation and decision boundaries in fading.

$$\mathrm{Re}\left[rv^*\right] = 0 \qquad \mathrm{Re}\left[-jrv^*\right] = 0$$
$$\mathrm{Re}\left[(r \pm 2v)v^*\right] = 0 \qquad \mathrm{Re}\left[(-jr \pm 2v)v^*\right] = 0 \quad (31)$$

where we have replaced $\mathrm{Im}[\cdot]$ with $\mathrm{Re}[-j \cdot]$ for simplicity, and have dropped the dependence on k, since we determined it to be weak in Section III-B.

A symbol error occurs if r falls outside the decision cell determined by the appropriate boundaries. The probability of this event depends on the transmitted symbol; clearly, it is greater for an interior cell than it is for a corner cell, which has only two boundaries. There are four distinct cell types, for which representative symbols are: $b_1 = 1 + j$ (interior cell); $b_2 = 1 + j3$ (edge cell); $b_3 = 3 + j$ (another edge cell); and $b_4 = 3 + j3$ (corner cell). We shall carry the calculations through only for the interior cell with symbol b_1. Calculations for the rest are similar, and are summarized in the Appendix for completeness.

For the given interior cell, an error occurs if any of the

following are true:

$$\begin{aligned} \mathrm{Re}\left[y_{11}v^*\right] &= \mathrm{Re}\left[rv^*\right] < 0 \\ \mathrm{Re}\left[y_{12}v^*\right] &= \mathrm{Re}\left[-jrv^*\right] < 0 \\ \mathrm{Re}\left[y_{13}v^*\right] &= \mathrm{Re}\left[(2v - r)v^*\right] < 0 \\ \mathrm{Re}\left[y_{14}v^*\right] &= \mathrm{Re}\left[(2v + jr)v^*\right] < 0 \end{aligned} \tag{32}$$

where y_{11} to y_{14}, as defined here, are zero mean Gaussian variates, correlated with v. The probability of each event can therefore be calculated as in Section III-A:

$$P_e(\rho_{1m}) = \frac{1}{2}\left(1 - \sqrt{\mathrm{Re}\left[\rho_{1m}\right]^2/(1 - \mathrm{Im}\left[\rho_{1m}\right]^2)}\right) \tag{33}$$

where ρ_{1m} is the correlation coefficient relating y_{1m} and v. From the definitions of y_{1m} (32) and r (10), the four correlation coefficients are easily determined to be:

$$\begin{aligned} \rho_{11} &= b_1\sigma_{uv}^2\left(\left(|b_1|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2} \\ \rho_{12} &= -j\rho_{11} \\ \rho_{13} &= \left(2\sigma_v^2 - b_1\sigma_{uv}^2\right)\left(\left(4\sigma_v^2 - 4\,\mathrm{Re}\left[b_1\sigma_{uv}^2\right]\right.\right. \\ &\qquad \left.\left. + |b_1|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2} \\ \rho_{14} &= \left(2\sigma_v^2 + jb_1\sigma_{uv}^2\right)\left(\left(4\sigma_v^2 - 4\,\mathrm{Im}\left[b_1\sigma_{uv}^2\right]\right.\right. \\ &\qquad \left.\left. + |b_1|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2} \end{aligned} \tag{34}$$

where the variances σ_u^2, σ_{uv}^2, and σ_v^2 are given by (11) and (18). The union bound gives a close approximation to the error rate for symbol b_1:

$$P_{s1} \le \sum_{m=1}^{4} P_e(\rho_{1m}). \tag{35}$$

In practice, the bound is almost indistinguishable from the true value if the SER is below about 0.01, and is useful—within about 10% of the true value—for SER as high as 0.4.

As noted above, a similar calculation must be performed for the other three cell types. Since the four cell types are equally represented, the final calculation for SER is just:

$$P_s = \frac{1}{4}\sum_{i=1}^{4} P_{si}. \tag{36}$$

B. Higher Order Constellations

For larger constellations like 64QAM, calculations like those above become very tedious, and upper bounding the overall SER with that of a single interior cell becomes attractive. However, not all interior cells have the same error rate, as can be seen readily from (13). We can write:

$$r = bv + be + n$$

so that the additive disturbance has variance $|b|^2\sigma_e^2 + 1$. Therefore the cells farthest from the origin—the interior corners—have the highest error rate, and provide the upper bound.

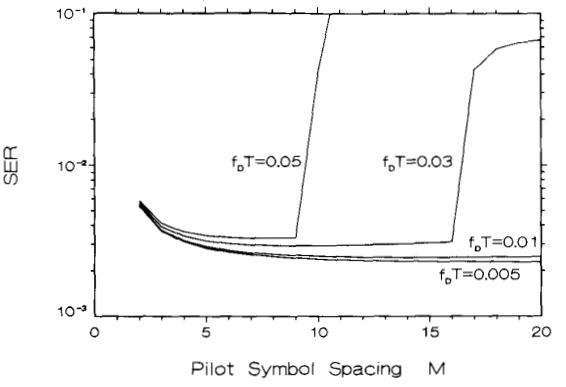


Fig. 8. Effect of frame size on 16QAM (γ_b = 30 dB, K = 63).

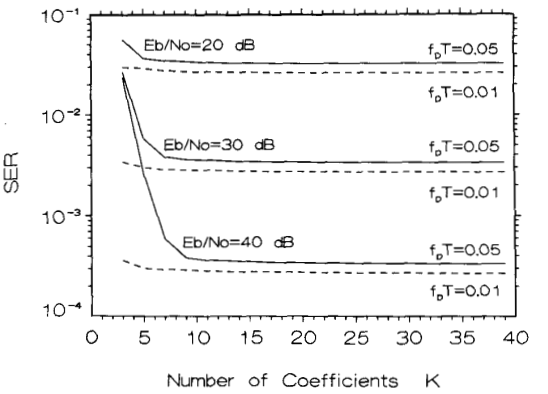


Fig. 9. Effect of interpolator size on 16QAM (M = 7).

C. Frame Size and Interpolator Size

Figs. 8 and 9 illustrate the effect of frame size M and number K of interpolator coefficients for 16QAM. The pilot symbols were selected as the corner symbols $b_4 = 3 + j3$; more precisely, they were real symbols of the same magnitude $\sqrt{18}$. It can be seen that $M = 7$ and $K = 11$ is a reasonable choice for the baseline system, just as it was for BPSK/QPSK in Section III.

D. 16QAM With Optimized Coefficients

Fig. 10 shows the SER of PSAM with coefficients optimized at every point. For comparison, the lowest curve represents the performance of an ideal system, with perfect phase and amplitude reference, and was obtained by letting M approach infinity in σ_u^2, then setting σ_v^2 and σ_{uv}^2 equal to σ_u^2 before evaluating (34–36). PSAM is about 0.9 dB worse than the ideal system at zero Doppler, and only about 2.4 dB worse than ideal at 5% Doppler. Since 0.7 dB of the loss is due to the power spent in the pilot symbols, performance can be improved further at low Doppler values by spacing the pilot symbols more widely, though this increases the delay.

In comparing 16QAM with BPSK (Fig. 5), note that the 16QAM SER should be divided by four to obtain the equivalent BER of a Gray coded constellation. Thus the performance penalty of 16QAM is 3 dB in comparing ideal reference

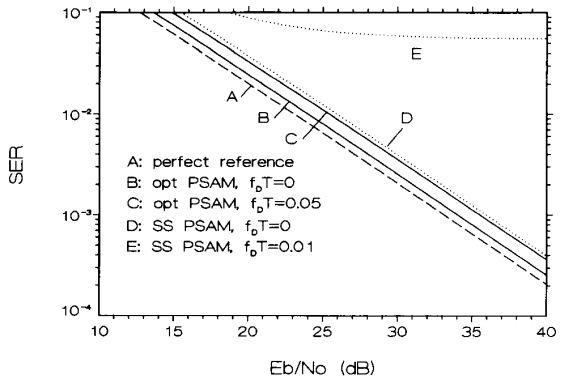


Fig. 10. SER performance of optimized PSAM for 16QAM.

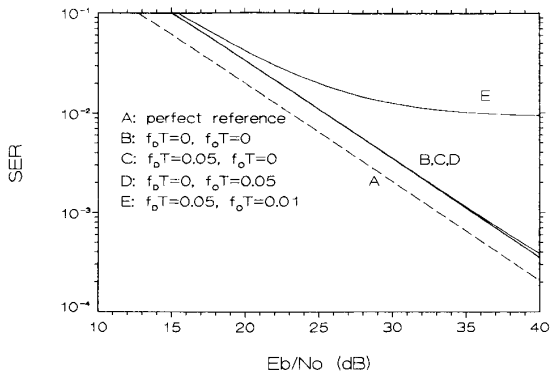


Fig. 11. PSAM 16QAM: Mismatched with coefficients optimized for $\gamma_b = 30$ dB, $f_oT = 0$, $f_DT = 0.05$.

systems, and about 2.1 dB in comparing 16QAM to BPSK at 5% Doppler.

Also shown on Fig. 10 as SS PSAM is the performance of a system using the three-point interpolator coefficients specified in [7], calculated with the method of Section IV-A after substituting those coefficients in (18). The frame size is 7. For zero Doppler, the performance is slightly worse than that of the optimized coefficients at 5% Doppler. At 1% Doppler, the performance is unacceptable. It is worth noting that these coefficients produce a strong dependence of SER on position within the frame: about 33% at zero Doppler, and over an order of magnitude at 1% Doppler. The comparison demonstrates the value of both the optimization and the use of a longer interpolator.

E. 16QAM With Mismatched Coefficients

As noted in Section III-F, the operating point of a system could well differ from that for which the coefficients are optimized. Fig. 11 shows the performance of a system in which the optimization point is $\gamma_b = 30$ dB, $f_oT = 0$, $f_DT = 0.05$. The general behavior is like that in Fig. 6, for BPSK: the SER is unaffected by frequency offset and Doppler spread up to 5%; however, if the combination of offset and Doppler exceeds the interpolation filter bandwidth, a significant error floor appears.

V. Conclusion

Pilot symbol assisted modulation is relatively simple to implement. The transmitter just inserts known symbols periodically, so there is no change in pulse shape or peak to average power ratio. The receiver interpolates the channel measurements provided by the pilot symbols to obtain an amplitude and phase reference for detection. The pilot symbols lower the effective bit rate by about 14% if the Doppler spread is 5% of the symbol rate. For smaller Dopper values, though, the loss of capacity is much less: about 5% for a 1% Doppler spread. Minor drawbacks include the delay and buffer space required at the receiver for interpolation. However, the good error performance, the removal of the error floor, and the enabling of multilevel signal formats outweigh these deficiencies.

PSAM is straightforward to analyze. We obtain analytical expressions for BER in BPSK and QPSK, for a tight upper bound on SER in 16QAM, and for the optimum interpolation coefficients. For BPSK and QPSK, optimized PSAM's loss compared with the unachievable coherent detection is only 1 dB for very slow fading, and it can be reduced further by transmitting pilot symbols less frequently. Even at 5% Doppler, the loss is only 3 dB. There is no error floor, as there is in differential detection. Moreover, optimized PSAM outperforms differential detection at any values of Doppler and SNR.

Comparison with TTIB is also favorable. PSAM is less complex at both the transmitter and receiver, and does not introduce envelope fluctuations. The loss of capacity, or bandwidth, is somewhat lower in PSAM, and the header preceding each transmission can be much shorter than in TTIB. In addition PSAM does not suffer the 3 dB performance penalty incurred by TTIB due to the latter's requirement for differential encoding.

For 16QAM, optimized PSAM is 0.9 dB poorer in very slow fading than the unachievable system with perfect phase and amplitude reference, and is 2.4 dB poorer at 5% Doppler. A comparison PSAM system using previously published interpolation coefficients [7] has poorer performance in very slow fading than optimized PSAM at 5% Doppler, and has a large irreducible error rate at 1% Doppler.

In an operational system, both Doppler and frequency offset are variable, so that PSAM is unlikely to be optimized at all times. When the coefficients are optimized for a worst case Doppler value, the performance for smaller Doppler and offset values is the same as for that worst case. However, if the combined Doppler and frequency offset exceeds the filter bandwidth, a significant error floor appears. This indicates the importance of correctly identifying the worst case. A possible remedy is to make the coefficients adaptive.

Appendix
Summary of Correlation Coefficients for 16QAM

For the reader's convenience, we have listed below the correlation coefficients of the four distinct cell types, for which representative symbols are: $b_1 = 1 + j$ (interior cell);

$b_2 = 1 + j3$ (edge cell); $b_3 = 3 + j$ (another edge cell); and $b_4 = 3 + j3$ (corner cell).

$$\rho_{11} = b_1\sigma_{uv}^2\left(\left(|b_1|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{12} = -jb_1\sigma_{uv}^2\left(\left(|b_1|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{13} = \left(2\sigma_v^2 - b_1\sigma_{uv}^2\right)\left(\left(4\sigma_v^2 - 4\,\mathrm{Re}\left[b_1\sigma_{uv}^2\right] + |b_1|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{14} = \left(2\sigma_v^2 + jb_1\sigma_{uv}^2\right)\left(\left(4\sigma_v^2 - 4\,\mathrm{Im}\left[b_1\sigma_{uv}^2\right] + |b_1|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{21} = b_1\sigma_{uv}^2\left(\left(|b_2|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{23} = \left(2\sigma_v^2 - b_2\sigma_{uv}^2\right)\left(\left(4\sigma_v^2 - 4\,\mathrm{Re}\left[b_2\sigma_{uv}^2\right] + |b_2|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{22} = \left(-2\sigma_v^2 - jb_2\sigma_{uv}^2\right)\left(\left(4\sigma_v^2 - 4\,\mathrm{Im}\left[b_2\sigma_{uv}^2\right] + |b_2|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{31} = -jb_3\sigma_{uv}^2\left(\left(|b_3|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{32} = \left(-2\sigma_v^2 + b_3\sigma_{uv}^2\right)\left(\left(4\sigma_v^2 - 4\,\mathrm{Re}\left[b_3\sigma_{uv}^2\right] + |b_3|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{33} = \left(2\sigma_v^2 + jb_3\sigma_{uv}^2\right)\left(\left(4\sigma_v^2 - 4\,\mathrm{Im}\left[b_3\sigma_{uv}^2\right] + |b_3|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{41} = \left(-2\sigma_v^2 + b_4\sigma_{uv}^2\right)\left(\left(4\sigma_v^2 - 4\,\mathrm{Re}\left[b_4\sigma_{uv}^2\right] + |b_4|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

$$\rho_{42} = \left(-2\sigma_v^2 - jb_4\sigma_{uv}^2\right)\left(\left(4\sigma_v^2 - 4\,\mathrm{Im}\left[b_4\sigma_{uv}^2\right] + |b_4|^2\sigma_u^2 + 1\right)\sigma_v^2\right)^{-1/2}$$

where the variances σ_u^2, σ_v^2, and σ_{uv}^2 are given in (11) and (18).

References

[1] J. P. McGeehan and A. J. Bateman, "Phase locked transparent tone-in-band (TTIB): A new spectrum configuration particularly suited to the transmission of data over SSB mobile radio networks," *IEEE Trans. Commun.*, vol. COM-32, pp. 81-87, Jan. 1984.

[2] A. J. Bateman *et al.*, "Speech and data communications over 942 MHz TAB and TTIB single sideband mobile radio systems incorporating feed-forward signal regeneration," *IEEE Trans. Veh. Technol.*, vol. VT-34, pp. 13-21, Feb. 1985.

[3] F. Davarian, "Mobile digital communications via tone calibration," *IEEE Trans. Veh. Technol.*, vol. VT-36, pp. 55-62, May 1987.

[4] P. M. Martin *et al.*, "The implementation of a 16-QAM mobile data system using TTIB-based fading correction techniques," in *Proc. IEEE Veh. Technol. Conf.*, Philadelphia, PA, 1988, pp. 71-76.

[5] J. H. Lodge and M. L. Moher, "Time diversity for mobile satellite channels using trellis coded modulations," *IEEE Global Telecommun. Conf.*, Tokyo, 1987.

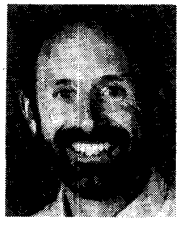
[6] M. L. Moher and J. H. Lodge, "TCMP—a modulation and coding strategy for Rician fading channels," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 1347-1355, Dec. 1989.

[7] S. Sampei and T. Sunaga, "Rayleigh fading compensation method for 16QAM in digital land mobile radio channels," in *Proc. IEEE Veh. Technol. Conf.*, San Francisco, CA, May 1989, pp. 640-646.

[8] W. C. Y. Lee, *Mobile Communications Engineering*. New York: McGraw-Hill, 1982.

[9] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1983.

[10] J. K. Cavers, "The performance of phase locked transparent tone-in-band with symmetric phase detection," *IEEE Trans. Commun.*, vol. 39, pp. 1389-1399, Sept. 1991.



James K. Cavers (M'90) was born in Port Alice, BC, Canada on September 19, 1944. He received the B.A.Sc. degree in engineering physics and the Ph.D. degree in electrical engineering, both from the University of British Columbia, in 1966 and 1970, respectively.

From 1970 to 1979 he was an assistant, then an associate, professor in the Department of Systems Engineering at Carleton University in Ottawa. He spent 1979 to 1982 as a program manager at MacDonald Dettwiler and Associates, in Vancouver, followed by a year as senior engineer at Glenayre Electronics, also in Vancouver. In 1983 he joined the School of Engineering Science at Simon Fraser University, Burnaby, where he holds the rank of Professor. His research interests include modulation and protocols for mobile communications, and integrated RF/DSP design.

Dr. Cavers is a Fellow of the Advanced Systems Institute of British Columbia.