Pilot Embedding for Channel Estimation and Tracking in OFDM Systems

Shankar Balasubramanian, Behrouz Farhang-Boroujeny and V. John Mathews
ECE Department, University of Utah, Salt Lake City, UT84112
e-mails: balasubr@eng.utah.edu, farhang@ece.utah.edu and mathews@ece.utah.edu

***Abstract*— We consider the problem of channel estimation and tracking in OFDM systems and explore the idea of adding pilot symbols to the data symbols as a means to conserve bandwidth. The term *pilot embedding* (PE) is used to refer to this scheme. Compared to the *pilot insertion* (PI) scheme, i.e., the conventional pilot symbol assisted modulation (PSAM), PE is more bandwidth efficient since no separate subcarriers/timeslots are allocated to pilots. We formalize this by evaluating the capacity of the two schemes and showing that PE indeed has the potential to transmit at a higher rate. The problem of channel tracking using a decision directed approach is reviewed and found to be unreliable, in the sense that the channel estimator fails to track the channel variations after some iterations because of unavoidable decision errors. We propose an ad hoc channel estimation algorithm that uses the embedded pilots along with the past decisions of data for reliable tracking of the channel.**

I. Introduction

Orthogonal frequency division multiplexing (OFDM) [1] has attracted considerable attention in the recent years [2] and [3]. In a mobile communication system, the channel conditions change continuously with time. Thus, in order to equalize for the channel in an OFDM system, one needs to continuously keep track of the channel. The need for channel estimation and tracking in OFDM systems can be avoided by employing differential demodulation. But this results in up to a 3 dB loss in signal-to-noise ratio (SNR) [4]. To perform coherent demodulation, one needs accurate estimates of the channel fading process. This is usually obtained by transmitting known symbols called pilots [4].

In decision directed technique [5], pilot symbols are transmitted in the beginning of each communication session to obtain an initial estimate of the channel. By using this channel estimate to equalize for the channel, one can make decision for the subsequent data symbols. Assuming that the decisions are correct, they may be used to estimate the channel as it evolves. Unfortunately, this method may not work in practice, because accumulation of decision errors may lead to catastrophic failure of the receiver.

An alternate solution is to insert pilot symbols into the data stream. This scheme which in this paper is referred to as pilot insertion (PI) is the conventional pilot symbol assisted modulation (PSAM) [6]. At the receiver, one first estimates the channel at the pilot points. These values are then interpolated, using a two dimensional Wiener filter [7], to estimate the channel at the data points. However, periodic insertion of pilot symbols reduces the effective data transmission rate of the system.

The idea of adding pilot symbols to data symbols as a means to conserve bandwidth was first proposed in [8] in the context of single carrier systems. By doing so, it is possible to transmit data (as well as pilots) at all times, thus avoid any reduction in the data rate. This scheme which we refer to as pilot embedding (PE) was used in conjunction with multicarrier systems in [9], where a semi-blind iterative algorithm was proposed to do the channel estimation. However, due to the iterative nature of the algorithm, the computational complexity of this scheme is much greater than the scheme that is proposed in this paper.

In this paper, we propose a novel decision directed channel estimation algorithm for an OFDM system employing embedded pilots. The embedded pilots are used in such a way that they play the important role of anchoring the channel estimates to their true values. Thus, even while encountering poor channel conditions temporarily, we are assured that decision errors are not going to make the channel estimates deviate away from their true values. We numerically evaluate and compare the maximum achievable rate of the PE and PI schemes and show that the PE scheme has the potential to transmit at a higher rate. Computer simulations that show similar or better performance of the PE scheme, compared to the PI scheme, are presented.

II. System Model

We assume that the channel is constant during each OFDM symbol. Also, we assume that the synchronization is perfect. Moreover, we assume that the duration of the channel impulse response is smaller than or equal to the cyclic prefix length, thus orthogonality of the subcarriers is preserved. Hence, the demodulated signal sample $Y[l,n]$, in the lth OFDM symbol and along the nth subcarrier, can be written as [10]

$$Y[l,n] = H[l,n]X[l,n] + N[l,n] \quad (1)$$

where $H[l,n]$ is the channel gain along the nth subcarrier when the lth OFDM symbol is transmitted, and $N[l,n]$ represents a sample of a zero-mean complex Gaussian white noise process with variance σ_n^2. We also assume that $H[l,n]$ has a variance of $\sigma_h^2 = 1$, for all values of l and n. Accordingly, the SNR of the system is defined as

$$\gamma = \frac{\sigma_x^2}{\sigma_n^2}. \quad (2)$$

III. Channel Estimation

A. PE Scheme

In the PE scheme, at the transmitter, pilot symbols $P[l,n]$ are added to data symbols $D[l,n]$ prior to modulation. Accordingly, the transmitted symbol $X[l,n]$ can be written as

$$X[l,n] = D[l,n] + P[l,n]. \quad (3)$$

Assuming that pilot and data symbols are independent of one another, the transmit power is the sum of data power and pilot power, viz., $\sigma_x^2 = \sigma_d^2 + \sigma_p^2$. We define the pilot-to-total power ratio $\rho = \sigma_p^2/\sigma_x^2$. Substituting (3) in (1), we obtain

$$Y[l,n] = H[l,n]D[l,n] + H[l,n]P[l,n] + N[l,n]. \quad (4)$$

Since, here, pilot and data symbols are transmitted concurrently, we may note that the data-to-noise power ratio, which determines the BER when the channel is known, is slightly lower than the SNR defined in (2). We define the effective SNR

$$\gamma_{\text{eff}} = \sigma_d^2/\sigma_n^2 = (1-\rho)\gamma. \quad (5)$$

In (4), the challenge is to solve for $H[l,n]$ and thereby determine $D[l,n]$, inspite of having three unknowns in $H[l,n]$, $D[l,n]$ and $N[l,n]$. We solve this problem with the help of pilots and by making use of the known statistics of the channel and noise and also the fact that the data symbols are from a finite set. We propose a three step algorithm as follows:

Step 1: Prediction - A priori estimates
We start by predicting the channel gains for each of the subcarriers on the kth (present) frame. For predicting the channel at the mth subcarrier of the kth frame, we select a few demodulated signal samples from its adjacent subcarriers on a few of the previous frames. From each of these, we remove the pilot component and then normalize by the corresponding data estimates, viz.,

$$\breve{H}_1[l,n] = \frac{Y[l,n] - \hat{H}[l,n]P[l,n]}{\hat{D}[l,n]}. \quad (6)$$

We arrange all the selected $\breve{H}_1[l,n]$ in a column and denote this vector by $\breve{\mathbf{h}}_1$. We also define the autocorrelation matrix $\mathbf{R}_1 = E\{\breve{\mathbf{h}}_1\breve{\mathbf{h}}_1^{\text{H}}\}$ and the crosscorrelation vector $\mathbf{r}_1 = E\{\breve{\mathbf{h}}_1 H^*[k,m]\}$, where the superscripts $*$ and H denote complex conjugate and Hermitian, respectively. The Wiener filter coefficients are then given by $\mathbf{w}_1 = \mathbf{R}_1^{-1}\mathbf{r}_1$ and we obtain the channel estimate $\hat{H}_1[k,m] = \mathbf{w}_1^{\text{H}}\breve{\mathbf{h}}_1$. Using this estimate, we remove the pilots, equalize the channel and make a hard decision $\hat{D}_1[k,m]$ for the transmitted data symbol $D[k,m]$.

Step 2: Tracking - Pilot based estimates
To get reliable estimates of the channel based on pilot symbols, we remove the data portion and then normalize by the pilot symbols to obtain the channel estimates

$$\breve{H}_2[l,n] = \begin{cases} \frac{Y[l,n]-\hat{H}[l,n]\hat{D}[l,n]}{P[l,n]}, & l < k \\ \frac{Y[l,n]-\hat{H}_1[l,n]\hat{D}_1[l,n]}{P[l,n]}, & l = k, \end{cases} \quad (7)$$

where $\hat{D}[l,n]$, for $l < k$, are the final estimates of data in the previous frame (obtained in Step 3), and $\hat{D}_1[k,n]$ are the data estimates obtained in the previous step. Arranging all $\breve{H}_2[l,n]$ in a column vector $\breve{\mathbf{h}}_2$, similar to $\breve{\mathbf{h}}_1$ above, and defining $\mathbf{R}_2 = E\{\breve{\mathbf{h}}_2\breve{\mathbf{h}}_2^{\text{H}}\}$ and $\mathbf{r}_2 = E\{\breve{\mathbf{h}}_2 H^*[k,m]\}$, we obtain $\mathbf{w}_2 = \mathbf{R}_2^{-1}\mathbf{r}_2$ and the second channel estimate $\hat{H}_2[k,m] = \mathbf{w}_2^{\text{H}}\breve{\mathbf{h}}_2$. This is then used to obtain the corresponding decision $\hat{D}_2[k,m]$, after removing the pilots.

Step 3: Smoothing - Decision directed estimates
The pilot-based channel estimates obtained in Step 2 may be very noisy, since the pilot power is usually a very small fraction of the total signal power. Nevertheless, this channel estimate and the corresponding data estimates are reliable in the sense that they are not subject to catastrophic failure. In this final step, we employ one more filtering operation to finish up with a good, noise suppressed estimate of the channel. The normalization of $Y[l,n]$ at this step is described by

$$\breve{H}_3[l,n] = \begin{cases} \frac{Y[l,n]-\hat{H}[l,n]P[l,n]}{\hat{D}[l,n]}, & l < k \\ \frac{Y[l,n]-\hat{H}_2[l,n]P[l,n]}{\hat{D}_2[l,n]}, & l = k. \end{cases} \quad (8)$$

Again, if we arrange the channel estimates $\breve{H}_3[l,n]$ in a column vector $\breve{\mathbf{h}}_3$, and define $\mathbf{R}_3 = E\{\breve{\mathbf{h}}_3\breve{\mathbf{h}}_3^{\text{H}}\}$ and $\mathbf{r}_3 = E\{\breve{\mathbf{h}}_3 H^*[k,m]\}$, the final estimate of the channel is obtained as $\hat{H}[k,m] = \hat{H}_3[k,m] = \mathbf{w}_3^{\text{H}}\breve{\mathbf{h}}_3$, where $\mathbf{w}_3 = \mathbf{R}_3^{-1}\mathbf{r}_3$. After equalizing with this channel estimate, we will get the final data estimate $\hat{D}[k,m]$.

B. PI Scheme

In the PI scheme, the first estimate of the channel is obtained through the traditional PSAM scheme [6]. This is equivalent to Step 2 in the PE case – there is no Step 1 here. This will be followed by a better estimate of the channel which is obtained by using the decisions obtained after equalization using the first estimates of the channel. Wiener filtering similar to Step 3 in the PE case is then applied to improve on channel estimates before the final equalization and data estimates are performed.

IV. Capacity Comparison of the PE and PI Schemes

In order to bring out the advantages of the PE scheme, we need to show that its performance compares favorably with the conventional PSAM scheme, i.e., the PI scheme. In this section, we take an information theoretic approach and compare PE and PI schemes in terms of their achievable capacity. Since in practical applications symbols of finite alphabet are used, we evaluate the channel capacities when data symbols are from a finite alphabet.

A. Capacity of a Rayleigh fading channel

Consider a system where the demodulated signal can be written as

$$Y = HD + N. \tag{9}$$

This is the same as the channel model (1), except that we have removed the time and frequency indices for convenience of notations. In what follows, we first assume that perfect knowledge of the channel H is available. The results are then extended to the cases with imperfect channel estimates by subtracting the corresponding loss in SNR due to estimation errors. We also restrict ourselves to the case where D is selected from an m-ary PSK alphabet with equal probability $1/m$. The equal probability follows from the symmetry of PSK symbols.

The capacity C can be related to the mutual information I as [11], [12]

$$\begin{aligned} C &= I(D;Y,H) = I(D;Y|H) + I(D;H) \\ &= I(D;Y|H) \end{aligned} \tag{10}$$

since independence of D and H implies $I(D;H) = 0$. Equation (10) can be expanded as [11]

$$\begin{aligned} C &= \int_h I(D;Y|H=h) f_H(h)\, dh \qquad (11) \\ &= \int_h \sum_d \int_y \frac{1}{m} f_Y(y|d,h) \log\left[\frac{f_Y(y|d,h)}{f_Y(y|h)}\right] f_H(h)\, dy\, dh \end{aligned}$$

where $f_Y(y)$ and $f_H(h)$ are the probability density functions of Y and H, respectively, and d denotes an element in the input alphabet set. Since

$$f_Y(y|h) = \frac{1}{m}\sum_d f_Y(y|d,h) \tag{12}$$

we see that the capacity depends only on the single function $f_Y(y|d,h)$ which happens to be a two dimensional Gaussian distribution function; when D and H are known, then $E\{Y\} = HD$ and the only ambiguity in Y is the noise N. We can thus evaluate (11) using numerical integration in two dimensions [13].

B. Effects of channel estimation errors

Here, we derive the SNR loss due to channel estimation errors. The channel estimation mean square error (MSE) of the PE scheme can be calculated by considering the smoothing stage and assuming that all decisions made are correct. Once again dropping the time and subcarrier indices for the sake of convenience, we define the channel estimation error as $\epsilon = H - \hat{H}$ and the channel estimation MSE as $\sigma_\epsilon^2 = E\{|\epsilon|^2\}$. From the Wiener filter theory [14], we have

$$\sigma_\epsilon^2 = \sigma_h^2 - \mathbf{w}_3^{\mathrm{H}} \mathbf{r}_3. \tag{13}$$

This is for the PE scheme. It can be readily extended the PI scheme simply by letting $\rho = 0$ while evaluating (13).

Recall that in the PE scheme, we first subtract the quantity $\hat{H}P$ from the demodulated signal before equalizing. The equalizer input is thus given by $Y' = \hat{H}D + N_{\text{eff}}$, where $N_{\text{eff}} = \epsilon(D + P) + N$. Here, we define the effective SNR $\gamma_{\text{eff}} = \frac{E|\hat{H}D|^2}{E|N_{\text{eff}}|^2}$, which can be evaluated as (recalling the assumption $\sigma_h^2 = 1$)

$$\gamma_{\text{eff}} = (1 - \rho)\eta\gamma \tag{14}$$

where $\eta = \frac{1-\sigma_\epsilon^2}{1+\gamma\sigma_\epsilon^2}$. Finally, one can show that N_{eff} and $\hat{H}$ are weakly dependent, thus can be assumed to be independent [15]. Noting this, one can calculate the capacity of a system with imperfect channel estimates by substituting γ_{eff} for SNR.

C. Capacity curves

For the PE scheme, we allot 10% of the total power for the pilots, i.e., $\rho = 0.1$. Computer simulation and also theoretical analysis show that $\rho = 0.1$ is a near optimum choice for minimizing the BER [16]. Moreover, we note that $\rho = 0.1$ translates to a loss of about 0.5 dB in SNR. In addition to this loss, we have the SNR loss due to the channel estimation as discussed above. Given γ_{eff}, we can evaluate the achievable rate of a QPSK signal constellation using (11).

For the PI scheme, we only have the SNR loss due to channel estimation. However, the achievable data rates have to be multiplied by a factor that takes care of the rate loss due to pilot insertion. The percentage of pilot symbols in most of the practical cases [17], [18] and also in broadcasting standards [2], [3] is usually around 10%. Accordingly, we multiply the resulting capacity values by 0.9.

The capacity results of the two schemes evaluated as above are shown in Fig. 1 for the fading rate of $f_d = 0.05$ normalized to the OFDM frame rate, T_s. Similar curves are obtained if f_d is varied over a wide range. At low SNR, the two schemes have almost identical capacity. But as SNR increases, the difference becomes more prominent. This is because at high SNR, the discrete (here, QPSK) capacity is relatively flat and hence the loss in SNR due to pilot power in the PE scheme is insignificant. However, for the PI scheme, the rate loss due to pilot insertion can never be compensated.

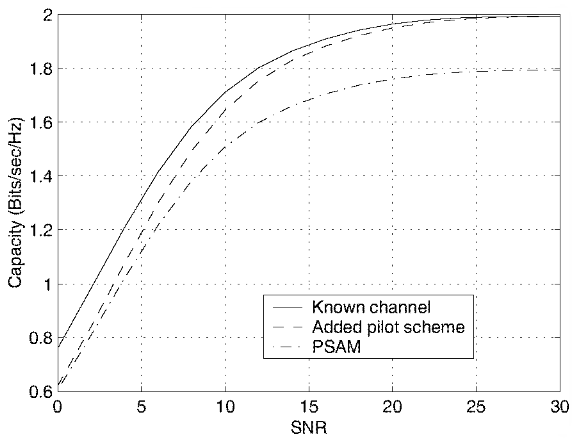

Fig. 1. Capacity of a QPSK system as a function of SNR for $f_d T_s = 0.05$.

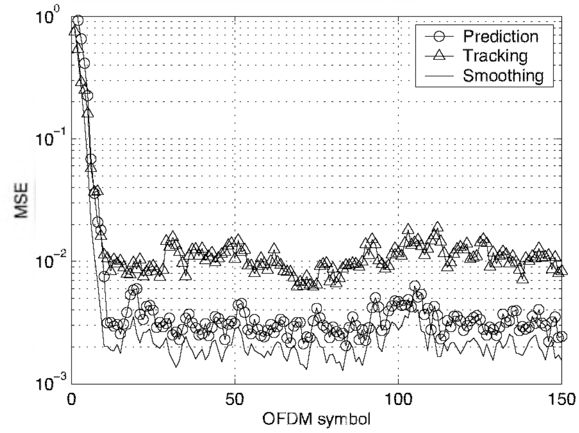

Fig. 2. The channel estimation MSE in each of the three stages of the proposed PE-based algorithm.

V. Simulation Results

We assume an OFDM system which employs a 128 point IFFT. The middle 113 subcarriers are used for data transmission and the rest are left as guard bands. The channel duration is assumed to be of 16 symbol-spaced taps, exponential delay power profile, with an RMS value of 4 taps. Each tap is an independent random process changing according to the Jakes' model with the specified fading rate f_d. Each OFDM symbol is assigned a cyclic-prefix of 16 samples. The input symbols are selected from a QPSK constellation with Gray encoding. For all two dimensional filtering operations, we use a 5-by-5 filter. In the case of PE scheme, we choose $\rho = 0.1$, i.e., 10% of transmit power is allotted to pilots. In the case of PI scheme, 10% of symbol slots is allocated to pilots.

In order to get an insight into how the proposed PE-based algorithm performs, plots of the channel estimation MSE at the output of the three stages of the algorithm, in a typical run, is presented in Fig. 2. The mean here refers to the average of the squared estimation errors along all the subcarriers in a given OFDM symbol. We note that the MSE at the output of the predictor is much smaller than that of the tracking stage. This is because, in the tracking stage, we use the (low-power) pilot symbols to obtain the channel estimates. These estimates, in spite of being noisy, are unbiased and hence reliable in the sense that they never deviate away from their true values, even when channel is in deep fade and many decision errors may occur. The third stage improves on channel estimates and results in a MSE that is lower than that of the previous two estimates.

Another aspect of the PE-based algorithm, that is observed from the results of Fig. 2, is the ability of the algorithm to rapidly converge to the neighborhood of the true channel values, starting from no a priori knowledge. We started this run with random values for the initial data symbols, and did our predicting, tracking, and smoothing using them. This observation, which we have seen consistently in simulation, strongly supports our claim that the embedded pilots help in anchoring the channel estimates to their true values.

In Section IV, we compared the capacity of the two pilot based schemes and found that at higher SNRs, the PE scheme has the potential to transmit at a higher rate. We now support this claim by giving some simulation results.

We first consider a PE scheme with a rate 9/10 convolutional code. This code is derived from the rate 1/2 convolutional code with the octal polynomials (133, 171) by appropriate puncturing [11]. The encoder has a memory of 6 bits and hence a tail of 6 zeros are needed to clear its memory every time. The receiver uses a soft-decision Viterbi decoder with a truncated memory length of 30 bits [19]. We follow the interleaving procedure recommended in IEEE 802.11a standard [20] which suggest a bit-by-bit block interleaving within each OFDM symbol. Since the interleaving is done individually within each OFDM symbol, there is no processing delay. The coded and interleaved bits are mapped to QPSK symbols, which are used to modulate the subcarriers. For the PI scheme, along alternate OFDM symbols, every fifth subcarrier transmits a pilot symbol. Hence, ten percent of the bandwidth is occupied by the pilots. There is no coding employed by the PI case in this example. We now have two systems that transmit data at the same rate – the PI scheme allots 10% of the subcarriers for pilots, where as the PE scheme uses them for coding. In Fig. 3, we plot the BER of the two schemes under the above setup and find that the performance of the two schemes are almost identical. Note that the horizontal axis is the bit power per noise noise power, E_b/N_o. This accommodates the SNR loss due to the pilots power.

As a second example, we now compare the packet error rate (PER) of the two schemes assuming BCH coding [11], where

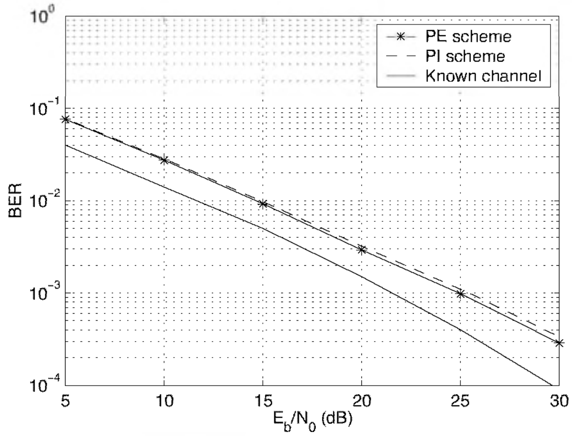

Fig. 3. Comparison of BER of the two schemes under equal data transmission rate at a fading rate of $f_d T_s = 0.001$.

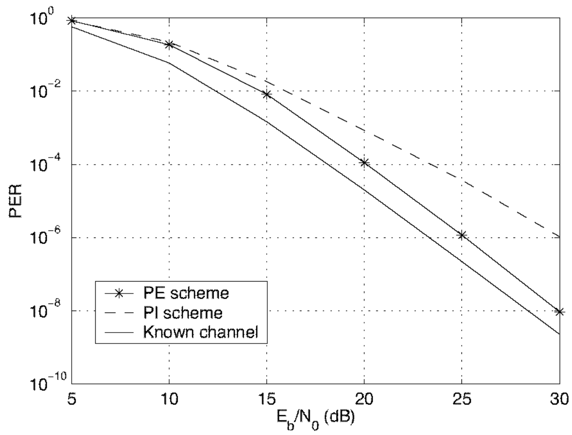

Fig. 4. Comparison of PER of the two schemes under equal data transmission rate with BCH coding. The result are for the case $f_d T_s = 0.001$. The results confirm that at high SNRs the PE scheme can have superior performance.

the term packet here refers to each block of BCH code. For the PE scheme, we employ a 45/63 BCH code that can correct up to three bits. The PI scheme employs the same pilot pattern as that of the previous example. Hence, to make the transmission rates of the two schemes the same, we apply a 51/63 BCH code for the PI scheme, which can correct up to two bit errors. Assuming perfect interleaving, we can assume that the bit errors are uniformly distributed. Hence, from the uncoded BER, one can calculate the packet error rate. These curves are plotted in Fig. 4. We note that at higher SNRs, the PE scheme has a superior performance. We also note that PER curves of the known channel case and the PE case differ by a constant horizontal shift of about 1.5 dB. This shift arises because of 0.5 dB loss due to pilot power and another 1 dB due to channel estimation error.

VI. Conclusion

We studied the possibility of adding pilots to data for the purpose of channel estimation and tracking in OFDM systems. This scheme, which we called pilot embedding (PE), was compared against the conventional pilot assisted scheme where pilot symbols are inserted in between the data symbols. It was shown that the PE scheme has the potential of offering a higher capacity when compared with the pilot insertion (PI) scheme. Computer simulations that show equal or superior behavior of the PE scheme against PI scheme when both systems transmit at the same rate were also presented.

References

[1] J. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communications Magazine*, vol. 28, pp. 5–14, May 1990.

[2] "Radio broadcasting systems: Digital audio broadcasting (DAB) to mobile, portable and fixed receivers," *ETS 300 401, European Telecommunications Standards Institute*, May 1997.

[3] "Digital terrestrial broadcasting systems," *ETS 300 744, European Telecommunications Standards Institute*, February 1997.

[4] J. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

[5] P. Frenger and N. Svensson, "Decision-directed coherent detection in multicarrier systems on Rayleigh fading channels," *IEEE Transactions on Vehicular Technology*, vol. 48, pp. 490–497, March 1999.

[6] J. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Transactions on Vehicular Technology*, vol. 40, pp. 686–693, November 1991.

[7] P. Hoeher, S. Kaiser, and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering," *ICASSP-97*, vol. 3, pp. 1845–1848, April 1997.

[8] B. Farhang-Boroujeny, "Pilot-based channel identification: proposal for semiblind identification of communication channels," *Electronic Letters*, vol. 31, pp. 1044–1046, June 1995.

[9] C. Ho, B. Farhang-Boroujeny, and F. Chin, "Added pilot semi-blind estimation scheme for OFDM in fading channels," *GlobeCom 2001, IEEE*, vol. 5, pp. 3075–3079, November 2001.

[10] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Artech House, 2000.

[11] S. Wilson, *Digital Modulation and Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[12] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[13] G. Fishman, *Monte Carlo: Concepts, Algorithms and Applications*. Springer Verlag, 1996.

[14] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*. John Wiley and Sons, 1998.

[15] A. Aghamohammadi and H. Meyr, "On the error probability of linearly modulated signals on Rayleigh frequency-flat fading channels," *IEEE Transactions on Communications*, vol. 38, pp. 1966–1970, November 1990.

[16] S. Balasubramanian, B. Farhang-Boroujeny, and V. J. Mathews, "Pilot embedding for channel estimation and tracking in ofdm systems," *IEEE Trans. Communications*, vol. submitted.

[17] R. Nilsson, O. Edfors, M. Sandell, and P. Borjesson, "An analysis of two-dimensional pilot-symbol assisted modulation for OFDM," *IEEE International Conference on Personal Wireless Communications*, pp. 71–74, December 1997.

[18] M. Sandell and O. Edfors, "A comparative study of pilot-based channel estimators for wireless OFDM," *Research Report, TULEA 1996:19, Division of Signal Processing, Lulea University of Technology*, September 1996.

[19] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Transactions on Communications*, vol. 40, pp. 873–884, May 1992.

[20] "IEEE standard 802.11a-2000," *IEEE*, 1999.