Phase Precoding for Frequency-Selective Rayleigh and Rician Slowly Fading Channels

Weihua Zhuang, Member, IEEE, and W. Vincent Huang

Abstract—This paper presents a novel phase precoding (pre-equalization) technique to equalize frequency-selective Rayleigh and Rician slowly fading channels for personal communication systems using phase modulation. In order to achieve intersymbol interference (ISI)-free transmission, the precoding technique predistorts the signal transmitted from a base station to a portable unit. The novelty of the technique lies in using a spiral curve design: 1) to ensure the stability of the precoder even in equalizing a non-minimum-phase channel; 2) to obtain an ISI-free received signal; and 3) to keep a constant transmitted signal amplitude. Using the precoder can improve the bit-error-rate (BER) transmission performance without increasing the complexity of the portable unit receiver. The BER performance of coherent quadrature phase-shift-keying (QPSK) with the channel pre-equalization is analyzed theoretically for both Rayleigh and Rician fading channels. Analytical and simulation results demonstrate that coherent QPSK using the proposed channel precoder has a significantly lower BER than that using a conventional decision-feedback equalizer (DFE) because the precoder does not suffer from error propagation.

Index Terms—Channel precoding, equalization, frequency-selective fading, personal communications.

I. INTRODUCTION

In a typical indoor wireless environment, a transmitted signal often reaches a receiver via more than one path due to reflection, refraction, and scattering of radio waves by structures inside a building. This results in a phenomenon known as time-dispersive (or frequency-selective) multipath fading. The indoor wireless channel is usually characterized by amplitude fluctuation, carrier phase jitter, and propagation delay spread. In high-bit-rate transmission, the delay spread results in intersymbol interference (ISI), which dramatically increases the transmission bit error rate (BER) [1]. Several equalization techniques have been proposed for both transmitter and receiver using linear or nonlinear filters to achieve ISI cancellation [2]–[8]. Of all the receiver equalization techniques, the decision-feedback equalizer (DFE) is the most popular nonlinear equalizer for severe fading channels. However, when coded modulation is used to further improve system performance, the obvious combination of decoding with the DFE may not perform well because zero delay decisions in coded system are not sufficiently reliable for decision feedback. The problem can be solved at the expense of receiver complexity, such as using parallel decision feedback decoding (PDFD) [2] and [3]. The PDFD technique implements a DFE and a Viterbi decoder in one structure for coded modulation in order to mitigate error propagation effect of DFE and to achieve better transmission performance than the corresponding uncoded modulation with DFE. It uses a unique sequence of decisions for each state in the trellis, rather than using only one sequence of decisions in the feedback path of the DFE. These feedback sequences are based on the history of each state’s surviving path. As a result, instead of calculating metrics with one received sample per trellis stage, it is necessary to calculate metrics with a unique decision feedback sample for each state at that stage. Such approaches are in conflict with desirable characteristics of portable units of low-power personal communication systems. A key requirement for such systems is the availability of compact light hand-held portable voice and data units. Portable units are battery-powered and have stringent power and size constraints. On the other hand, the constraints are not so severe at the base stations. The consideration leads to the approach of moving channel equalization functions from the receivers of portable units to the transmitters of base stations, so that the portable units can be relieved from equalization tasks.

If the channel characteristics are known to the transmitter, channel precoding (or pre-equalization) at the base station becomes possible for the forward link (transmission from the base station to the portable unit). Unlike a DFE, in which all equalization takes place at the receiver, precoding is a transmitter technique. Ideally, the transfer function of the precoder is the inverse of the channel transfer function. When the fading channel is not minimum-phase, the inverse of the channel transfer function is unstable, which means that we cannot have a linear and stable channel precoder. Nonlinear operation is necessary to achieve the precoder stability. The main challenge in designing a nonlinear channel precoder is to, simultaneously: 1) ensure the stability of the precoder even in equalizing a nonminimum-phase fading channel; 2) achieve an ISI-free received signal; and 3) keep a constant transmitted signal amplitude in a system using phase modulation. Tomlinson–Harashima (TH) precoding is a technique very suitable for quadrature amplitude modulation (QAM) [4], [5]. This technique can be extended to trellis precoding, which combines equalization with trellis-coded QAM modulation [6]. The key point in TH precoding is the nonlinear modula-
This precoding technique cannot be directly applied to time-varying fading channels where amplitude fading causes serious errors in retrieving the original amplitude information at the receiver, due to the corresponding modulo-arithmetic reduction. Furthermore, the TH precoding predistorts the transmitted signal amplitude, which is undesirable for phase modulation with constant amplitude. The phase modulation is usually preferred in personal communication systems due to its frequency spectral efficiency and allowing the use of a power efficient nonlinear amplifier. Recently, an adaptive channel precoding technique has been proposed for phase modulation over a slowly fading channel, which predistorts both the carrier phase and amplitude of the transmitted signal by using feedforward and feedback linear filters and a nonlinear switchable delay device [7]. An automatic gain control (AGC) unit is used in the adaptive precoder to achieve a constant amplitude, which also introduces nonlinear distortion to the precoded transmission signal and results in channel equalization errors. In addition, since a feedforward transversal linear filter is used in the precoder to pre-equalize a nonminimum-phase channel, a large number of taps is necessary in order to combat ISI completely.

A new channel precoding approach has been introduced in [8] for phase modulation that can overcome the disadvantages of the adaptive precoder proposed in [7].

In this paper, we further develop and analyze the precoding technique initially presented in [8]. The technique uses a unique spiral curve design to precode only the carrier phase in order to achieve an ISI-free received signal. As a result, the predistorted signal to be transmitted still has a constant amplitude, which permits the use of a power efficient nonlinear amplifier without causing undue nonlinear distortion and maintains the spectral efficiency of phase modulation with a constant amplitude. The channel precoder is different from DFE in the sense that it is a transmitter equalization technique so that it does not suffer from the error propagation (which is inherent to the DFE). Since no decision feedback is necessary when using the channel precoding, powerful coding techniques with a Viterbi decoder [9] can be used in the system to further improve the transmission performance without significantly increasing the complexity of portable units. The remainder of this paper is organized as follows. Section II describes the system model where the channel precoder is used. Section III discusses the principle and operation of the nonlinear phase precoding. Section IV presents the theoretical BER performance analysis of coherent quadrature phase-shift keying (QPSK) using the channel precoder in Rayleigh and Rician fading channels. In Section V, numerical analysis and computer simulation results are given to demonstrate the BER performance improvement of coherent QPSK achieved by using the channel precoder over that using a conventional DFE. Conclusions of this work are given in Section VI.

**II. SYSTEM DESCRIPTION**

Fig. 1 shows the proposed functional block diagram of an indoor wireless communication system, which consists of one base station and one portable unit for simplicity. The system operates in a time-division multiple access (TDMA) time-division duplex (TDD) scheme. In the reverse link from the portable unit to the base station, a DFE is used to reduce the ISI resulting from the time-dispersive fading channel. The channel estimator estimates the channel impulse response based on the tap coefficients of the DFE. In a TDD scheme, both forward link and reverse link share the same radio channel by using different time slots. If the channel impulse response is time-invariant, i.e., the channel can be modeled as a linear time-invariant system, then the RF links “are reciprocal” (in both directions) according to the reciprocity theorem applied to radiation patterns and linear systems [10]. In the case that the radio channel impulse response is time-variant, if the channel fades so slowly that the channel characteristics are approximately invariant over the time interval of two adjacent data frames for the forward and reverse links respectively, then the RF links are reciprocal over the time interval of the two adjacent data frames (one for transmission over the forward link and the other for the reverse link). As a result, the channel information estimated in one link can be used for precoding over the other link. Since the simplicity of the portable unit is a major concern in the system design, Fig. 1 shows that the channel information estimated in the reverse link is used in the precoder at the base station transmitter to pre-equalize the signal transmitted in the forward link. With the channel precoding, the signal received at the portable unit is ISI free. Since all the channel equalization functions are performed at the base station, the simplicity of the portable unit can be obtained.
As to the indoor radio fading channel, different propagation time delays of various propagation paths result in that two received signal components of different frequencies have independent statistical properties if the frequency separation is large enough. The maximum frequency difference for which the signals are still strongly correlated is called the coherence bandwidth of the radio fading channel [11]. In high-bit-rate transmission, when the transmitted signal bandwidth is larger than the coherence bandwidth of the fading channel, the channel exhibits frequency-selective fading. The fading channel can be described by its baseband complex impulse response [12]

\[ h(t) = \sum_{l=0}^{L} a_l(t) e^{j\omega_l t} \delta(t - \tau_l) \]  

(1)

where \( L + 1 \) is the total number of the propagation paths with distinguishable propagation delays; \( a_l(t) \) is Rayleigh distributed amplitude fading for the first path if there is no line-of-sight (LOS) component, Rician if there is; \( \psi_l(t) \) is the carrier phase jitter of the first path with a uniform distribution over \([-\pi, +\pi]\) for Rayleigh fading and a nonuniform for Rician fading; \( a_l(t)(l > 0) \) is uniformly distributed (over \([-\pi, +\pi]\)) carrier phase jitter for the \((l + 1)\)th path; and arrival time \( \tau_l(l = 0, 1, 2, \cdots, L) \) forms a Poisson process.

Phase-shift-keying (PSK) is considered as the modulation scheme for the system. The information signal to be transmitted over \( t \in [T, iT + T] \) (where \( T \) is one symbol interval) is given by

\[ s_i(t) = A \cos(\omega_c t + \beta_i) \]  

(2)

where \( A \) is the constant amplitude, \( \beta_i \) is the carrier phase carrying the transmitted information, and \( \omega_c \) is the radian radio frequency. The signal \( s_i(t) \) can be simplified to its complex-valued envelope denoted as

\[ s_i^e(t) = Ae^{j\beta_i}. \]  

(3)

After precoding, the actual transmitted signal with the same constant amplitude is

\[ \tilde{n}_i = Ae^{j\beta_i}. \]  

(4)

where \( \theta_i \) is the carrier phase after predistortion. The ISI due to the fading channel in the received signal, over \( t \in [iT, iT+T] \), is

\[ \tilde{u}_i = \sum_{l=1}^{L} a_l(t) e^{j\omega_l(t)} m_{i-l} \]

\[ = \sum_{l=1}^{L} a_l(t) e^{j\omega_l(t)} Ae^{j\beta_i-l} \]

\[ \triangleq a_T e^{j\psi_T} \]  

(5)

where \( a_T \) and \( \psi_T \) are the amplitude and phase of the ISI, respectively. In (5), the propagation delay differences among the multiple paths are assumed to be integer multiples of the symbol interval \( T \). However, the phase precoding discussed here can be extended to the situation where the delay differences are a fraction of a symbol interval. The ISI component \( \tilde{n}_i \) is known to the transmitter based on the channel characteristics and previously transmitted signals. If there is no fading on the first path, the received signal is

\[ \tilde{r}_i = \tilde{n}_i + \tilde{u}_i + \tilde{n}_i \]

\[ = Ae^{j\beta_i} + a_T e^{j\psi_T} + \tilde{n}_i \]

\[ \triangleq \tilde{r}_T e^{j\psi_T} \]  

(6)

III. NONLINEAR PHASE PRECODING

A. Minimum-Phase Channel

In the case that the ISI energy \( S_I \) is smaller than the transmitted signal energy \( S_T \), i.e., \( S_I < S_T \), the corresponding channel transfer function has all zeros inside the unit circle on the \( z \)-plane (minimum-phase condition). For the minimum-phase channel, one possible approach to cancel the ISI is illustrated in Fig. 2. The precoded signal is defined on the circle centered at the origin with radius \( A \) and \( \theta_i \) is the carrier phase after predistortion. The only received signal phase \( \phi_i \) is a major concern in detecting the signal. To achieve an ISI-free received signal, i.e., \( \phi_i = \beta_i \), we first offset \( \phi_i \) by \( \beta_i \). By finding the point where the circle centered at the origin with radius \( A \) intersects the line \( l_i \), we can obtain the transmitted signal \( \tilde{n}_i \) with constant amplitude \( A \) and, at the same time, achieve the ISI-free received signal \( \tilde{r}_i = \tilde{n}_i + \tilde{u}_i \).

B. Nonminimum-Phase Channel

When the ISI energy \( S_I \) is larger than the transmitted signal energy \( S_T \), i.e., \( S_I > S_T \), the corresponding channel
to combat the ISI completely. A new algorithm for detection of the transmitted information, which is a general approach to detecting PSK signals. In the case of a nonminimum-phase fading channel, a precoded signal does not exist if it is required, simultaneously, that: 1) the precoded signal has constant amplitude \( A \); 2) the transmitted information can be correctly detected by only the phase of the received signal; and 3) ISI-free transmission is achieved. In order to solve the problem, we first analyze the situation to find out if there is any additional information that we can make use of in the precoding or detection. Since the channel characteristics are unknown to the receiver of the portable unit, the only additional information available to the receiver is the amplitude of the received signal. In the above precoding scheme, the received signal is constrained on the one-dimensional (1-D) space (the line \( L_1 \) in Fig. 2) defined by the information carrier phase \( \beta_i \), because the receiver uses only the received signal carrier phase to detect the transmitted information. If the receiver can make use of both the amplitude and carrier phase of the received signal in detecting the transmitted information, then we should be able to relax the constraint on the received signal [or, equivalently, the requirement 2)]. In other words, we can extend the 1-D received signal space to a two-dimensional (2-D) received signal space. In this way, we should have more freedom in the design of the precoding scheme. In order for the receiver to be able to make use of both received signal amplitude and phase for detection, we need to establish a relationship between the amplitude and phase. One example of the 2-D received signal space is a spiral curve, which has exactly the required characteristic, as shown in Fig. 3. The spiral curve can be viewed as a generalization of the line \( L_1 \) in Fig. 2. Using polar coordinates \((r, \phi)\), the spiral curve can be described by

\[
r = A \left( 1 + \frac{\beta_i - \phi}{C \pi} \right). \tag{7}
\]

The spiral curve is characterized by two parameters \((\beta_i, C)\): the carrier phase \( \beta_i \) determines the angle of the point at \( r = A \) on the curve, and \( C \) determines the distance \( D_s(\equiv 2A/C) \) between any two points on the curve with a 2\( \pi \) phase difference. In order to transmit the signal with a constant amplitude \( A \), all the transmitted signal should be on the circle centered at origin with radius \( A \). In the absence of AWGN, all possibly received signals are on the mentioned circle offset by the ISI, \( a_0 e^{j\psi_0} \), denoted by

\[
r = a_1 \cos(\phi - \psi_1) \pm \sqrt{A^2 - a_1^2 \sin^2(\phi - \psi_1)}. \tag{8}
\]

This circle is called the ISI circle (as shown in Fig. 3). In the precoding, the precoded signal is determined in such a way that the received signal, \( r_1 e^{j\phi_1} \), is one root of (7) and (8), \( r_1 \) and \( \phi_1 \). The received signal is an intersection point of the ISI circle and the spiral curve, so that the relation between the amplitude and phase of the received signal is defined by (7). Once the received signal is determined, we can obtain the predistorted signal at the transmitter by subtracting \( a_0 e^{j\psi_0} \) from \( r_1 e^{j\phi_1} \):

\[
A e^{j\theta_i} = r_1 e^{j\phi_i} - a_0 e^{j\psi_0}, \tag{9}
\]

At the receiver, the spiral curve is determined from the amplitude \( r_1 \) and carrier phase \( \phi_i \) of the received signal, with a predetermined constant \( C \) (characterizing the spiral curve) known to the receiver. The carrier phase \( \beta_i \) can be detected according to

\[
\beta_i = \phi_i + \left( \frac{r_1}{A} - 1 \right) C \pi \tag{10}
\]

which is the angle of the point on the spiral curve where \( r = A \). All received signals are on the spiral curves with different initial angles depending on the value of \( \beta_i \).

If the first path signal experiences channel fading with a complex gain \( a_0 e^{j\psi_0} \), the effect of \( a_0 e^{j\psi_0} \) should be taken into account in obtaining the predistorted phase \( \theta_i \). The radius of the ISI circle should be changed from \( A \) to \( a_0 A \). Furthermore, the center of the ISI circle should be rotated by an angle \( -\psi_0 \) with respect to the spiral curve. This is equivalent to keeping the ISI circle unchanged but modifying the spiral curve at the precoder. In order to combat the carrier phase jitter \( \psi_0 \) of the first path fading, the spiral curve should be rotated by an angle \(-\psi_0\); because of the amplitude fading \( a_0 \), the distance \( D_s \) of the spiral curve should be scaled proportionally to \( D_s/a_0 \). As a result, the spiral curve (7) becomes

\[
r = \frac{A}{a_0} \left( 1 + \frac{\beta_i - \phi - \psi_0}{C \pi} \right). \tag{11}
\]

The algorithm for detection of the transmitted information is the same as that described in (10), because the first-path fading, \( a_0 e^{j\psi_0} \), has been taken into account in precoding the transmitted signal so that it should not affect the detection of the ISI-free received signal. The value of \( a_0 \) plays a very important role here. It expands \((a_0 < 1)\) or shrinks \((a_0 > 1)\) the distance \( D_s \). If the distance is too small, the receiver
C. Techniques to Reduce Precoding Errors

When there does not exist a root of (8) and (11) due to the first-path signal experiencing deep fading, i.e., \( \alpha_0 \ll 1 \), the following four techniques can be used, respectively, to determine the precoded signal.

1) **Small Amplitude Fluctuation**: A point on the spiral curve closest to the ISI circle is chosen to approximate a root of (8) and (11). This method can reduce the equalization error, but it can also result in a small amplitude fluctuation of the transmitted signal.

2) **Amplitude Modification**: In addition to applying technique 1), the amplitude of the transmitted signal is forced to be constant, which results in an equalization error as compared with that using technique 1).

3) **Deep-Fading Cutoff**: A predetermined minimum value (threshold) \( A_0 \) of the amplitude fading is chosen. For any first-path fading below this value, i.e., \( \alpha_0 < A_0 \), the threshold \( A_0 \) is applied to ensure the existence of a root or roots of (8) and (11).

4) **Using Adaptive \( C \) Values**: Since the distance \( D_s = 2A/\alpha_0 C \) if the value of \( C \) is chosen to be inversely proportional to \( \alpha_0 \), then the spiral curve has a constant value of \( D_s \) no matter what value \( \alpha_0 \) has. In this way, the equalization error due to \( \alpha_0 \ll 1 \) can be avoided, and the BER performance can be improved; however, it requires that the receiver know the time-varying value of \( \alpha_0 \).

Among the first three techniques, only the second and third still keep a constant transmitted signal amplitude. It is expected that the second technique is better since it has a “soft” cutoff rather than a “hard” cutoff as the third technique does. On the other hand, the system BER performance using the first technique should be better than that using the second technique, with the tradeoff of a nonconstant transmitted signal amplitude. Among the three techniques, the second is most preferred if a constant transmitted signal amplitude is essential to the system design. In the case that the receiver has the information of the \( \alpha_0 \) value, the last technique can be combined with the second technique to give the best performance. Details of the effect of \( C \) value and how to adapt the \( C \) value to the channel are discussed in Section V. The transmission performance of coherent QPSK using the channel precoder with the second technique is analyzed theoretically in Section IV. Computer simulation results of the BER performance using the precoder with each of the four techniques are presented in Section V.

Fig. 4 shows the structure of the phase precoder by using the above amplitude modification method. The inputs of the precoder are information signal and channel information estimated at the channel estimator (as depicted in Fig. 1). If \( r_{l+1} - r_l = T \) for \( l = 0, 1, \ldots, L - 1 \), then \( d_l(t) = a_l(t) e^{j \phi_l(t)} \) for \( l = 0, 1, \ldots, L \). The precoding process can be summarized as in the following.

1) Define the spiral curve for the current transmitted symbol based on \( \beta_l \) and \( \alpha_0 e^{j \phi_0} \) according to (11).
2) Calculate the ISI component \( a_l e^{j \phi_l} \) based on previously transmitted symbols and the estimated channel impulse response defined in (1).
3) Determine the received signal by calculating one root of (8) and (11). If a root does not exist, the precoder chooses a point on the spiral curve that is closest to the ISI circle.
4) Compute the predistorted signal \( A_0 e^{j \phi_i} \) according to (9).

D. Discussion

So far, we have described the nonlinear phase precoding scheme with constant transmitted signal amplitude for phase modulation. Three issues concerning the precoding scheme need to be addressed.

1) The maximum amplitude fluctuation of the precoded signal without amplitude modification is \( D_s/2 \). The fluctuation is smaller than the distance between the point chosen on the spiral curve and the circle centered at origin with radius \( A \). Hence, the amplitude fluctuation is reduced as compared with that when using the precoding method proposed in [7]. The AGC unit is to achieve a constant amplitude of the transmitted signal in the case that there does not exist a root of (8) and (11). The nonlinear equalization error due to the use of an AGC unit is mitigated by the precoding method proposed here because of the smaller amplitude fluctuation as compared with that using the method of [7]. The amplitude fluctuation using the precoding can be further reduced if the parameter \( C \) is adapted to the channel fading (to be discussed in Section V).

2) In the above discussion, the first-path signal is considered to be the desired signal. In other words, only post-cursors are considered in precoding. In general, the desired path should be the one that has the largest average power among all the propagation paths. In the case where the first path does not have the maximum average power, channel equalization should combat both precursors and post-cursors in the channel impulse response in order to achieve optimum transmission performance. In the DFE at the base station receiver, the effect of precursors is mitigated by a linear feedforward filter, and
the post-cursors (tails) are removed by a feedback filter with previously detected decisions as input. Similarly, if precursors should be considered in the precoding scheme, a linear feedforward filter can be used with the same tap number and tap gain coefficients, respectively, as those of the feedforward filter in the DFE. The filter can be inserted either after the precoder at the transmitter or before the signal detector at the receiver. Since inserting the filter after the transmitter would result in a nonconstant transmitted signal amplitude, it is preferred that the filter be inserted at the receiver front end. There are two ways for the receiver to determine the tap coefficients of the filter: 1) the base station transmits the information to the receiver, which achieves the receiver simplicity at the expense of the transmission overhead, and 2) the receiver linearly equalizes the precursors by continuously adapting the coefficients to maintain the overall channel response that the precoder expects [6], [13]. Although using the linear feedforward filter at the receiver increases the complexity of the portable units, the complexity does not increase when coded modulation and a Viterbi decoder is used because no decision feedback is required by the linear equalizer. The complexity of a portable unit using the linear equalizer with a precoder at the base station is significantly reduced as compared with that using a DFE because the DFE requires zero-delay decision feedback.

3) Although only PSK is considered here, the precoding technique can be directly applied to other phase modulation schemes such as π/4-shifted differential QPSK (the modulation used in the North American digital cellular system standard IS 54). In addition, the precoding technique can also be extended to precode continuous-phase frequency-shift keying (CPFSK) signals.

IV. PERFORMANCE ANALYSIS OF COHERENT QPSK

In this section, the BER performance of coherent QPSK using the channel precoder with the amplitude modification method is derived for Rayleigh and Rician fading channels. Define the phase detection error as

\[ \Delta \beta_i \triangleq (\beta_i - \beta_k) \mod 2\pi \]

where \( \beta_k \) is the detected signal phase and \( \beta_i \) the correct signal phase. To make a correct decision, it requires

\[ -\frac{\pi}{4} < \Delta \beta_k < \frac{\pi}{4}. \]  

(12)

The complex input white Gaussian noise accompanying the \( \beta \)th received signal symbol at baseband can be represented as

\[ \eta_i = n_{ci} + j n_{si} = n\exp(j\alpha_i) \]  

(13)

where \( n_{ci} \) and \( n_{si} \) are independent Gaussian random variables with zero mean and variance \( \sigma^2 \), \( n_z = \sqrt{n_{ci}^2 + n_{si}^2} \) has a Rayleigh distribution, and \( \alpha_i = \tan^{-1}(n_{si}/n_{ci}) \) has a uniform distribution over \([ -\pi, +\pi ]\). The received signal plus noise can be described as

\[ r_i e^{j\beta_i} + n_i e^{j\alpha_i} = (r_i + n_{ri}) e^{j(\beta_i + n_{ci})} \]  

(14)

where

\[ n_{ri} \approx n_i \cos(\phi_i - \alpha_i), \]

\[ n_{ci} \approx \frac{n_i}{r_i} \sin(\phi_i - \alpha_i) \]

in the case of a large SNR value of the input signal. It can be proven that \( n_{ri} \) and \( n_{ci} \) are independent Gaussian random variables with zero mean and variance \( \sigma^2 \) and \( \sigma^2/r_i^2 \) (given \( r_i \)), respectively. Since the variance of \( n_{ci} \) is inversely proportional to the SNR value of the input signal, its effect on BER performance is negligible compared with that of \( n_{ri} \) when the input signal has a large SNR value. In the following analysis, the effect of \( n_{ci} \) on BER is neglected. If ISI-free transmission is achieved, \( \Delta \beta_i \) results only from the Gaussian noise \( \eta_i \). From (10), it can be derived that (12) is equivalent to

\[ -\frac{\pi}{4} < \left\{ \frac{C\pi}{A} n_{ri} \mod 2\pi \right\} < \frac{\pi}{4} \]  

(15)

where \( (C\pi/A)n_{ri} \) is a Gaussian random variable with zero mean and variance \( (C\pi/A)^2 \sigma^2 \). Using Gray code for a QPSK signal, a 1-b error occurs if \( \Delta \beta_i \in (\pi/4, 3\pi/4) \) or if \( \Delta \beta_i \in (-3\pi/4, -\pi/4) \) and a 2-b error occurs if \( \Delta \beta_i \in (3\pi/4, 5\pi/4) \). As a result, the BER is

\[ T(\pi/4) + T(3\pi/4) = 2T(\pi) \]

in the case of perfect ISI cancellation, with the function \( T(\cdot) \) defined as

\[ T(x) = \sum_{i=0}^{\infty} Q \left( \frac{2n\pi + x}{\sigma C\pi} \right) \]  

(16)

where

\[ Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp \left( -\frac{t^2}{2} \right) dt. \]

On the spiral curve, the distance increment \( D_\theta \) for each \( 2\pi \) phase increment is \( 2A/\alpha_0 C \) according to (11). When \( D_\theta \leq 2A, i.e., \alpha_0 \geq 1/C \), there always exists a root or roots of (8) and (11), which corresponds to ISI-free transmission. Thus, the conditional probability of bit error, given \( \alpha_0 \geq 1/C \), is

\[ p_{e|\alpha_0} = T(\pi/4) + T(3\pi/4) - 2T(\pi). \]  

(17)

When \( D_\theta > 2A \), i.e., 0 < \( \alpha_0 < 1/C \), there may not exist a root of (8) and (11) in some cases, as shown in Fig. 5, which gives an example of the relative positions of the ISI circle and the lines of the spiral curve passing points \( a \) and \( b \). The shortest distance between the center of the ISI circle and any of the two lines is defined as \( |m| \), where \( m \) is positive if the circle is on the left side of the line closest to it, and negative if on the right. If \( |m| \leq A \), the ISI circle crosses the lines and an ISI-free signal is received. The conditional probability of bit error, given \( 0 \leq |m| \leq A \) and \( 0 < \alpha_0 < 1/C \), is

\[ p_{e|2A \leq |m| < A} = T(\pi/4) + T(3\pi/4) - 2T(\pi). \]  

(18)
which actually is not a function of $m$. Otherwise, if $A < |m| \leq D_s/2$, the ISI circle lies totally between $a$ and $b$. In this case, the distance $m$ increases phase detection error in addition to that caused by input noise. From (10), it can be obtained that the phase detection error is

\[
\Delta \beta_i = \begin{cases} 
\frac{C\pi}{A} (m - A + \alpha_i), & \text{if } m \in (A, +D_s/2) \\
\frac{C\pi}{A} (m + A + \alpha_i), & \text{if } m \in (-D_s/2, -A) 
\end{cases}
\]

(19)

which is a Gaussian random variable with mean $\mu_+ = C\pi(m - A)/A$ if $m \in (+A, +D_s/2)$, $\mu_- = C\pi(m + A)/A$ if $m \in (-D_s/2, -A)$, and variance $C^2\pi^2\alpha^2/2A$. Hence, the conditional probability of bit error, given $A < |m| \leq D_s/2$ and $0 < \alpha_0 < 1/C$, can be derived as

\[
p_e|_{A, \alpha_0, m} = 0.5 \left[ T \left( \frac{\pi}{4} - \mu \right) - T \left( -\frac{\pi}{4} - \mu \right) \right] + 0.5 \left[ T \left( \frac{3\pi}{4} - \mu \right) - T \left( -\frac{3\pi}{4} - \mu \right) \right] - [T(\pi - \mu) - T(\pi)]
\]

(20)

where $\mu = \mu_-$ if $m \in (-D_s/2, -A)$ and $\mu = \mu_+$ if $m \in (+A, +D_s/2)$.

Based on the above discussion, the probability of bit error can be calculated by

\[
p_e = \int_0^{1/C} f_2(\alpha_0) \, d\alpha_0 \int_{-A/\alpha_0 C}^A f_1(m) \, dm
\]

\[
\cdot \left\{ 0.5 \left[ T \left( \frac{\pi}{4} - \mu_- \right) - T \left( -\frac{\pi}{4} - \mu_- \right) \right] + 0.5 \left[ T \left( \frac{3\pi}{4} - \mu_- \right) - T \left( -\frac{3\pi}{4} - \mu_- \right) \right] - [T(\pi - \mu_-) - T(\pi - \pi_-)] \right\} f_1(m) \, dm
\]

\[+ \int_0^{1/C} f_2(\alpha_0) \, d\alpha_0 \int_{-A/\alpha_0 C}^A f_1(m) \, dm\]

(21)

where $f_1(\cdot)$ and $f_2(\cdot)$ are the probability density functions (pdf’s) of the distance $m$ and the first path amplitude fading $\alpha_0$, respectively. It is shown in Appendix that if the ISI has the phase component uniformly distributed over $[-\pi, +\pi]$, the distance $m$ has a uniform distribution over $[-A/\alpha_0 C, +A/\alpha_0 C]$ given $\alpha_0$. No matter whether $\alpha_0$ has a Rayleigh distribution or a Rician distribution.

V. RESULTS AND DISCUSSION

The BER performance of coherent QPSK using the channel precoder with the four precoding error reduction methods and using a conventional DFE is evaluated by numerical analysis and computer simulations. Time-variant Rayleigh and Rician fading channels with two independent propagation paths are considered, with propagation delay between the two paths equal to one symbol interval $T$. The ratio of the transmitted signal amplitude $A$ to the spiral curve parameter $C$ is chosen to be 0.25. The normalized power spectral density of the multipath diffusive components of both propagation paths is given by

\[
S(f) = \begin{cases} 
\frac{1}{\pi f_D \sqrt{1 - (f/f_D)^2}}, & |f| \leq f_D \\
0, & |f| > f_D 
\end{cases}
\]

(22)

where $f_D$ is the maximum Doppler frequency shift. Generally, the value of the normalized fading rate, $f_D T$, determines the degree of signal fading. When $f_D T \ll 1.0$, the channel exhibits slowly fading. Here, the normalized fading rate is chosen to be 0.005 and the channel information is updated over
Fig. 6. The BER using precoder with \textit{gain control} in Rician-fading channels.

each symbol interval at the precoder. In a practical system, for example, if each data frame contains 1000 symbols and the channel information is updated over each data frame, then the equivalent normalized fading rate correspondingly is $0.5 \times 10^{-5}$, which characterizes a slowly fading channel.

SNR/bit is defined as the ratio of the ensemble average of the received signal power per bit from both paths to the Gaussian noise variance $\sigma^2$. The ratio of the average power of the first path diffusive component to that of the second path diffusive component is denoted by $\rho$. The Rayleigh fading channel with the classical Doppler spectrum is simulated using the “sum of sine waves” method [11]. The Rician fading of the first path is also characterized by the $k$-factor, which is defined as the ratio of the received signal power from the LOS path to the power of the diffusive component. The value of $k$ is selected as $-10$, $-5$, $0$, and $5$ dB in the following analysis. The Rician fading of the first path is simulated by adding a time-invariant LOS component (determined by the $k$-factor) to a Rayleigh-distributed diffusive component. In the simulations, it is assumed that the precoder has accurate channel information. For comparison, simulation results of the transmission performance using a conventional DFE are also obtained, assuming that the receiver using a DFE can accurately track the carrier phase jitter of the first-path channel fading. In the simulations, the number of information bits transmitted through the system for each SNR/bit value is chosen in such a way that the simulated BER performance results are within the 90% confidence interval $[0.5\bar{e}_c, 2.0\bar{e}_c]$. The BER confidence interval depends on the number of bit-error occurrences as well as the number of fading cycles that the system experiences.

Fig. 6 shows both numerical analysis and simulation results of the bit error rates using \textit{amplitude modification with gain control} method for a Rician fading channel with $\rho = 0$ dB and $k = -10$, $-5$, $0$, $5$ dB, respectively. The solid lines represent numerical analysis results, and dotted lines simulation results. It is observed: 1) the BER floor decreases as the value of $k$ increases. A larger $k$ value means a stronger deterministic LOS path, therefore, there is less chance for the first-path to experience deep fading so that the equalization error [due to nonexisting root of (8) and (11)] is reduced. As a result, the BER floor is significantly reduced as $k$ increases; and 2) at low SNR/bit values, an increase of $\rho$ value results in a degradation of the BER performance. With the same SNR/bit value, the value of $\sigma^2$ increases as $k$ increases. Since the amplitude of the received signal is used for detection, when SNR/bit is low, Gaussian noise has a dominant effect on BER performance deterioration compared with the equalization error, which results in an increased BER as $\sigma^2$ (or $k$) increases.

Fig. 7 illustrates the BER using the precoder in a Rayleigh fading channel with $\rho = 0$ and $10$ dB, respectively, using the \textit{amplitude modification with gain control} method. The BER floors in the Rayleigh fading channels are higher than those of the Rician fading channels, due to the lack of the LOS signal component and more frequently occurred deep fading of the first path. It is also observed from Fig. 7 that, with an increase of $\rho$ value, the BER performance is improved. This is because: 1) as $\rho$ increases, the average signal power from the first path is increased and from the second path is decreased, since the sum of the average signal powers from both paths is normalized to one, and 2) with a larger signal energy from the first path, there is a larger probability for the channel to be minimum phase, resulting in a smaller equalization error. In both Figs. 6 and 7, the probabilities of bit error from numerical analysis based on (21) agree very well with those obtained from computer simulations.
Figs. 8 and 9 give the BER simulation results over a Rician fading channel using small amplitude fluctuation method and using deep fading cutoff method, respectively, with $\rho = 0\,\text{dB}$ and $k = -10$, $-5$, $0$, $5\,\text{dB}$. Comparing Fig. 6 with Fig. 9, we can see that the performance has been improved significantly by the amplitude modification method as expected, due to the fact that it is still possible to cancel ISI completely with a deeply faded first-path signal using this method. From Figs. 6, 8, and 9, it is concluded that: 1) when there is a strong first-path signal all the time (i.e., in the case of $k = 5\,\text{dB}$), perfect ISI cancellation is possible to a great extent no matter which of the three methods is used, therefore, the corresponding transmission performance is almost the same in all the three figures, and 2) when it is possible that the first-path signal goes through deep fading (i.e., in the case of $k \leq 0\,\text{dB}$), there exist precoding errors, and the small amplitude fluctuation method provides the best transmission performance among the three methods. However, we sacrifice the constant amplitude as a trade-off. Taking into account of all aspects, the amplitude modification method should be the best choice among the first three techniques, since it still keeps a constant amplitude of the transmitted signal and has a lower BER value than that of the deep-fading cutoff method.

For comparison, Fig. 10 gives the BER simulation results of coherent QPSK using a DFE with $\rho = 0$ and $k = -\infty$, $-10$, $-5$, $0$, $5\,\text{dB}$, respectively. When $k$ is small, the channel approaches to a two-path Rayleigh fading channel and the first path often experiences deep fading, resulting in poor BER performance. With an increase of $k$ value, the first-path signal becomes stronger and the effect of ISI is reduced, thus, the BER performance is improved. The inherent error propagation of a DFE results in the high BER floors. Comparing the BER performance using the precoder with that using a DFE, it is observed that the former has lower BER floors. Since the precoder is not subject to error propagation, coherent QPSK using the precoder has a much lower BER floor than that using a DFE under the same channel condition.

The spiral curve design parameter $C$ plays a very important role in the detection algorithm. The dependence of the BER performance on the precoder design parameter $C$ is studied in the following. Fig. 11 shows the BER as a function of $C$. With a larger $C$ value, the distance $D_s$ is smaller; therefore, it is more likely for (8) and (11) to have a root or roots. That is, the precoder has a better capability of handling the first path fading. On the other hand, $\beta$ is proportional to $C\tau_i$ according to (10). With a small $C$ value, the precoder is less sensitive to the AWGN component $\tilde{n}_k$. In summary, a small $C$ value should be used when the SNR/bit of the received signal is small, and the $C$ value should be increased as the SNR/bit value increases. The transmission performance can be improved by using a $C$ value adapted to the SNR/bit value. One example of choosing a $C$ value based on the SNR/bit value is shown in Table I. The BER performance using a precoder with the amplitude modification method and the $C$ values is shown in Fig. 12 for the case of $\rho = 0$ and $k = 0$. The value of $C$ can also be determined by the first-path amplitude fading $\alpha_0$ (as mentioned in Section III). From (11), it is observed that if the value of $C$ is inversely proportional to the value of $\alpha_0$, then the distance $D_s$ is a constant independent of the first-path amplitude fading $\alpha_0$. As a result, the effect of a small $\alpha_0$ value (when the first path experiences deep fading) on the equalization error is suppressed, and the BER performance is expected to be significantly improved. This is verified by the BER performance with $C = 1/\alpha_0$ as shown in Fig. 12. With $C = 1/\alpha_0$ (i.e., $D_s = 2A$), there always exists at least one root for (8) and (11), which means no amplitude.
modification is necessary and ISI-free transmission is always possible.

It should be pointed out that: 1) the optimum $C$ value depends on both channel fading statistics and instantaneous SNR/bit value. The system using the precoder with the amplitude fluctuation method and a $C$ value adapted to both SNR/bit and $a_0$ should have better BER performance than that shown in Fig. 12. The performance improvement is achieved at the expense of the increased system complexity for finding the optimum $C$ value, and 2) if the $C$ value is adapted to $a_0$, then the receiver needs to know the value of $a_0$ either by the information sent from the transmitter or by estimating the value locally. In fact, it is much easier for the receiver to estimate $a_0$ alone than to estimate the whole channel impulse response $h(t)$. One way to estimate the $a_0$ value at the receiver is to transmit a preamble sequence of a few symbols. The
preamble sequence can be designed in such a way that the signals from delayed paths cancel each other, which makes it possible to estimate the first-path amplitude fading $\alpha_0$. The estimation error depends on the value of SNR/bit. For example, with the two-path channel, we can choose a preamble sequence of five symbols with pattern \{1, 1, -1, -1, 1\}. At the receiver, the received signal sequence corresponding to the preamble is \{r_1, r_2, r_3, r_4, r_5\} contaminated by additive white Gaussian noise. The first path fading $\alpha_0 e^{j\phi_0}$ can be estimated by $(r_1 + r_2 - r_3 - r_4 + r_5)/5$, which is a complex Gaussian random variable with mean $\alpha_0 e^{j\phi_0}$ and variance $\sigma^2/5$. The accuracy of the estimation can be increased by increasing the symbol number in the preamble. In general, we can always design a preamble sequence to estimate $\alpha_0$ for an indoor channel consisting of $L + 1$ paths (where $L \geq 1$). The accuracy of the estimation is better if a longer preamble

Fig. 10. The BER using DFE in Rayleigh and Rician fading channels.

Fig. 11. The BER using precoder with gain control in a Rician fading channel with $\rho = 0$ dB, $k = 0$ dB.
sequence is used or if there exists fewer multiple paths. The BER performance using the precoder with the parameter $C$ adapted to an estimate $a_0$ (with the 5-symbol preample) and the amplitude modification method is shown in Fig. 12. It is observed that the BER performance is not very sensitive to the error in estimating $a_0$ due to noise. For comparison purpose, the BER performance of the system using a DFE (with error propagation) instead of the precoder is also shown in Fig. 12 under the same channel condition. It is clearly observed that the precoder can improve the transmission performance much more effectively than a DFE. The reason for the precoder to have better performance is that the precoder has no inherent error propagation (as the DFE does). On the other hand, a DFE with accurate channel information and correct feedback decisions can outperform the precoder.

In practice, the channel impulse response information used in the precoder is estimated in the reverse link using a channel information estimator, as shown in Fig. 2. The estimator makes use of the tap coefficients of the DFE at the base station receiver. The error propagation of the DFE may result in channel information estimation error. However, for slowly fading channels, powerful tracking algorithms can be applied to suppress fast (false) variation components in the channel fading status due to decision errors in the DFE and due to additive Gaussian noise, so that the estimation error can be greatly reduced. With an accurate estimate of channel impulse response, the transmission performance with a practical precoder is expected to be very close to that shown in this section.

### VI. CONCLUSIONS

A new constant-amplitude channel precoder for equalizing slowly fading channels has been developed and analyzed for phase modulation. The transmitted signal carrier phase is precoded in a unique way (by using the spiral curve design) to achieve an ISI-free received signal. The transmitted information is detected at the receiver according to the established relation between the received signal amplitude and carrier phase. It has been shown that the new channel precoder can improve the system BER performance significantly better than a DFE does because of no inherent error propagation. The precoder is especially useful for portable communications, where ISI due to multipath fading channels can severely deteriorate the BER transmission performance and where the simplicity of portable units is a vital characteristic of the system.

### APPENDIX:

**DERIVATION OF THE PDF OF $m$**

We consider a two-path fading channel, with the delayed path having Rayleigh distributed amplitude fading $a_1$ and uniformly distributed carrier phase jitter $\psi_1$ over $[-\pi, +\pi]$. From (5), the ISI component has Rayleigh-distributed amplitude $a_2A$.
and uniformly distributed phase $\psi Y \overset{\text{i.i.d.}}{\sim} \psi Y + \theta_{n-1}$ over $[-\pi, +\pi]$ for any $\theta_{n-1}$ value. On the plane using the polar coordinates $(r, \phi)$, $\phi = \psi Y$ is the line connecting the origin and the center of the ISI circle. From (11), the spiral curve intersects the line at the following coordinates:

$$r_n = \frac{A}{\alpha_0} \left( 1 + \frac{\beta_n}{C \pi} - \psi Y \right)$$

$$= \frac{A}{\alpha_0} + \frac{2nA}{\alpha_0 C} + X, \quad n = 1, 2, \ldots$$  \hspace{1cm} (23)

where

$$X = \frac{A}{\alpha_0} \left[ \frac{(\beta_i - \psi Y - \psi_0)}{\alpha_0 C} \right] \quad \text{Modulo} \quad 2\pi$$

is a random variable uniformly distributed over $[-A/\alpha_0 C, +A/\alpha_0 C]$ because $[(\beta_i - \psi Y - \psi_0) \text{Modulo} \quad 2\pi]$ has a uniform distribution over $[-\pi, +\pi]$ for any $\beta_i$ and $\psi_0$. The distance $m$ can be represented as

$$m \overset{\text{i.i.d.}}{\sim} (a_I - r_n) \quad \text{Modulo} \quad \frac{2A}{\alpha_0 C}$$

$$= \left[ (a_I - X) - \frac{A}{\alpha_0 C} \right] \quad \text{Modulo} \quad \frac{2A}{\alpha_0 C} \quad \text{(24)}$$

where $m \in [-A/\alpha_0 C, +A/\alpha_0 C]$. First, let us derive the pdf of random variable $Y \overset{\text{i.i.d.}}{\sim} a_I - X$, which is defined on $[-A/\alpha_0 C, +A/\alpha_0 C]$. For $\forall y \in [-A/\alpha_0 C, +A/\alpha_0 C]$, the cumulative distribution function (cdf) of $Y$ is

$$P(Y \leq y) = \int_{-A/\alpha_0 C}^{A/\alpha_0 C} \frac{1}{2A} f_X(u) f_{a_I}(v) \, du$$

$$= \frac{A}{2A} \int_{-A/\alpha_0 C}^{A/\alpha_0 C} f_X(u) f_{a_I}(v) \, du$$

$$= \frac{A}{2A} \int_{-A/\alpha_0 C}^{A/\alpha_0 C} [F_{a_I}(y + u) - F_{a_I}(0)] \, du \quad \text{(25)}$$

where $f_X(u) = a_0 C/2A$ is the pdf of $X$, $f_{a_I}(\cdot)$ is the pdf of $a_I$, and $F_{a_I}(x) \overset{\text{i.i.d.}}{\sim} F_{a_I}(y + u)$ is the cdf of $a_I$. The pdf of $Y$ can be obtained by differentiating its cdf

$$f_Y(y) = \frac{d}{dy} P(Y \leq y)$$

$$= \frac{A}{2A} \frac{d}{dy} \int_{-A/\alpha_0 C}^{A/\alpha_0 C} [F_{a_I}(y + u) - F_{a_I}(0)] \, du$$

$$= \frac{A}{2A} [F_{a_I}(\frac{A}{\alpha_0 C} + y) - F_{a_I}(\frac{-A}{\alpha_0 C})] \quad \text{(26)}$$

For $\forall y \in [A/\alpha_0 C, \infty)$, the cdf of $Y$ is

$$P(Y \leq y) = P\left( \frac{-A}{\alpha_0 C} \leq Y \leq \frac{A}{\alpha_0 C} \right)$$

$$+ P\left( \frac{A}{\alpha_0 C} < Y \leq y \right) \quad \text{(27)}$$

Similarly, it can be derived that the pdf of $Y$ is

$$f_Y(y) = \frac{A}{2A} \left[ F_{a_I}\left( y + \frac{A}{\alpha_0 C}\right) - F_{a_I}\left( y - \frac{A}{\alpha_0 C}\right) \right] \quad \text{(28)}$$

Substituting the cdf of the Rayleigh distributed random variable $a_I$ into (26) and (28), we have (29), shown at the bottom of the page, where $\sigma^2_\phi \overset{\text{i.i.d.}}{\sim} \alpha_E [\alpha_\phi^2]/2$. As a result, the random variable $Z \overset{\text{i.i.d.}}{\sim} Y - A/\alpha_0$ has the pdf shown in (30), shown at the bottom of the page. Finally, the pdf of $m \overset{\text{i.i.d.}}{\sim} (Z \text{Modulo} \quad 2A/\alpha_0 C)$ is

$$f_m(t) = \sum_{r=-\infty}^{\infty} f_Z\left( t + \frac{2nA}{\alpha_0 C} \right)$$

$$= \frac{a_0 C}{2A}, \quad t \in [-A/\alpha_0 C, +A/\alpha_0 C] \quad \text{(31)}$$

That is, $m$ is uniformly distributed over $[-A/\alpha_0 C, +A/\alpha_0 C]$ given the first-path amplitude fading $\alpha_0$. We conclude that the pdf of $m$ is independent of the distribution of $\psi_0$, due to the fact that $[(\beta_i - \psi Y - \psi_0) \text{Modulo} \quad 2\pi]$ is uniformly distributed over $[-\pi, +\pi]$ for any $\beta_i$ and $\psi_0$ as long as $\psi Y$ is uniform over $[-\pi, +\pi]$.
ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their thorough and meticulous reviews, which significantly improve the quality and presentation of this paper.

REFERENCES


Weihua Zhuang (M’93) received the B.Sc. (1982) and M.Sc. (1985) degrees from Dalian Marine University, China, and the Ph.D. degree (1993) from the University of New Brunswick, Canada, all in electrical engineering.

From January 1992 to September 1993, she was a Post Doctoral Fellow, first at the University of Ottawa and then at Telecommunications Research Labs (TRLabs, Edmonton), working on land mobile satellite communications and indoor wireless communications. Since October 1993, she has been with the Department of Electrical and Computer Engineering, University of Waterloo, where she is an Assistant Professor. Her current research interests include digital transmission over fading channels, wireless networking, and radio positioning.

Dr. Zhuang is a Licensed Professional Engineer in the Province of Ontario, Canada.

W. Vincent Huang received the M.S. degree in electrical engineering from Chalmers University of Technology (Sweden) in 1994. He is currently a Ph.D. candidate in medical electronics at the same University.

In 1993 and 1994, he completed the M.Sc. program in the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include wireless communications, simulation of neuromuscular systems, closed-loop control, adaptive control, and their applications to human motor control.