Adaptive H_∞ Channel Equalization for Wireless Personal Communications

Weihua Zhuang, *Member, IEEE*

***Abstract*—In this paper, a new adaptive H_∞ filtering algorithm is developed to recursively update the tap-coefficient vector of a decision feedback equalizer (DFE) in order to adaptively equalize the time-variant dispersive fading channel of a high-rate indoor wireless personal communication system. Different from conventional L_2 [such as the recursive least squares (RLS)] filtering algorithms which minimize the squared equalization error, the adaptive H_∞ filtering algorithm is a worst case optimization. It minimizes the effect of the worst disturbances (including input noise and modeling error) on the equalization error. Hence, the DFE with the adaptive H_∞ filtering algorithm is more robust to the disturbances than that with the RLS algorithm. Computer simulation demonstrates that better transmission performance can be achieved using the adaptive H_∞ algorithm when the received signal-to-noise ratio (SNR) is larger than 20 dB.**

***Index Terms*—Adaptive algorithm, decision feedback equalization, H_∞ filtering, wireless communications.**

I. Introduction

IN A TYPICAL indoor wireless environment, a transmitted signal often reaches a receiver via multiple propagation paths. In high-bit-rate transmission, the propagation delay spread of the time-dispersive (or frequency-selective) multipath fading channel results in intersymbol interference (ISI) which dramatically increases the transmission bit error rate (BER) [1]. Channel equalization is an efficient approach to combating ISI, and decision feedback equalizer (DFE) is the most popular nonlinear equalizer for severe fading channels. Adaptive filtering algorithms are used to adjust the tap coefficients of the DFE according to equalization errors in order to track a time-variant fading channel [2], [3]. All previous algorithms have been based on the minimization of the variance of the equalization error, i.e., the L_2 filtering such as Kalman and the recursive least squares (RLS) filtering approaches [3]–[5]. Using Kalman filtering algorithm, it is assumed that the receiver knows the statistical properties of additive Gaussian noise. In addition, the L_2 algorithms may not be robust against any uncertainty of channel fading models. For example, both feedforward and feedback filters of a DFE have a finite impulse response, while an ideal ISI cancellation may require that both filters have an infinite impulse response. The tails of the filter response are ignored because the components are relatively small if the tap number of the filters is large. However, unless a robust filtering algorithm is used, it is possible that the small modeling error may result in large equalization errors.

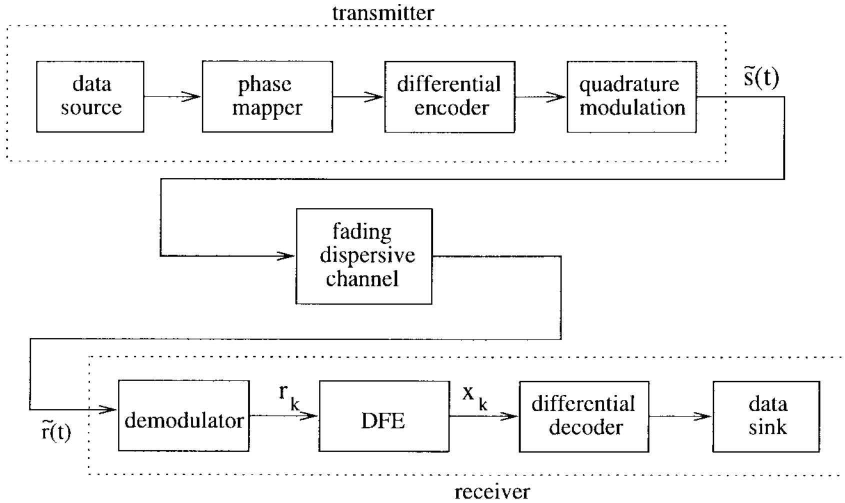
Recently, a new class of optimal filtering algorithm has been developed using H_∞ minimum spectrum error criterion [6]–[9]. In order to cope with disturbances of partially unknown statistics, the disturbances are modeled as deterministic square integrable signals, where only an upper bound of the signal power is assumed to be finite. This replaces the method of modeling the disturbance signal as a random process with a given spectral density, such as in the case of the L_2 filtering. The H_∞ filter is designed to ensure that the operator relating the noise signals and modeling errors to the resulting estimation error should possess an H_∞ norm less than a prescribed positive value. In other words, the H_∞ filtering is to minimize the worst possible amplification of the disturbance signal. In this paper, we investigate the application of H_∞ filtering to a DFE for high-rate data transmission of wireless personal communication systems. In particular, a new adaptive H_∞ filtering algorithm is developed in order to equalize time-variant fading channels. Theoretical formulation and mechanization procedures for the adaptive H_∞ filtering algorithm are given. For comparison, the transmission performance of the system using a DFE with the RLS algorithm is also studied. Computer simulation results show that the performance of the newly proposed adaptive H_∞ filter is better compared with that of the RLS filter in terms of equalization error and transmission bit error rate.

The remainder of this paper is organized as follows. In Section II, after a brief introduction of the communication system model under consideration, the application of the H_∞ filtering to the DFE is studied, and a new adaptive H_∞ filtering algorithm is derived for time-varying dispersive fading channels. Comparison between the new adaptive H_∞ filtering algorithm and the conventional RLS algorithm is discussed. Section III presents the equalization error and BER performance analysis using a DFE with the H_∞ and RLS filtering algorithms, respectively, where coherent quadrature phase-shift keying (QPSK) is used with differential encoding and decoding. Computer simulation results are given to demonstrate the BER performance improvement achieved by using the adaptive H_∞ algorithm over that using the RLS algorithm when the received signal-to-noise ratio (SNR) is larger than 20 dB. Conclusions of this work are given in Section IV.

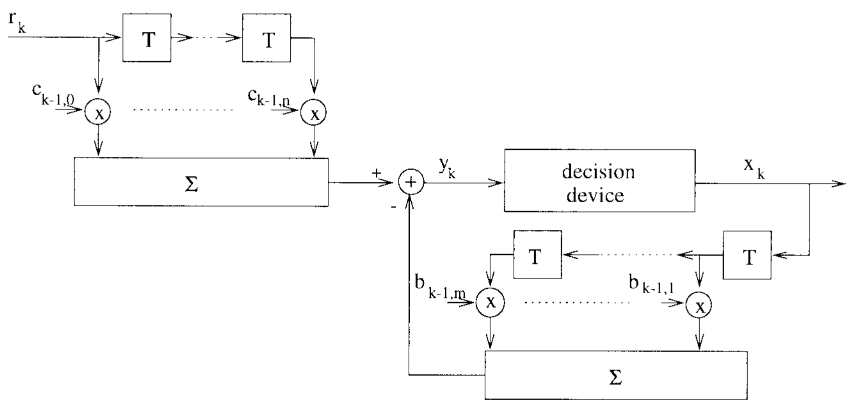
Manuscript received December 1, 1995; revised July 17, 1996. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada under Grant OGP0 155 131.

The author is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ont., N2L 3G1, Canada (e-mail: wzhuang@bbcr.uwaterloo.ca).

Publisher Item Identifier S 0018-9545(99)00604-0.



(a)



(b)

Fig. 1. Functional block diagram of the system model: (a) transmitter and receiver and (b) DFE.

II. Adaptive Filtering Algorithms for the DFE

In this section, a new adaptive H_∞ filtering algorithm is derived to optimize the tap coefficients of the DFE in order to achieve ISI-free transmission over time-varying dispersive fading channels. In order to distinguish the differences between the L_2 filtering and H_∞ filtering, we define L_2 norm of a discrete-time signal and H_∞ norm of a discrete-time system transfer function in the following.

Definition 1: The L_2 norm of a discrete-time signal $\{u_k\}$ is

$$||u||_2 \triangleq \sqrt{\sum_{k=0}^{\infty} |u_k|^2}. \tag{1}$$

Definition 2: If F is a transfer operator that maps a discrete-time input signal $\{u_k\}$ to a discrete-time output signal $\{v_k\}$, then the H_∞ norm of F is defined as

$$||F||_\infty \triangleq \sup_{u:0<||u||_2<\infty} \frac{||v||_2^2}{||u||_2^2} \tag{2}$$

where "sup" stands for supremum. The H_∞ norm can be viewed as the maximum energy gain from the input $\{u_k\}$ to the output $\{v_k\}$.

A. System Description

Fig. 1(a) shows the simplified functional block diagram of the digital communications system under consideration. The modulation scheme is coherent QPSK with differential encoding and decoding. The fading dispersive channel corrupts the transmitted signal $\tilde{s}(t)$ by introducing multiplicative envelope distortion, carrier phase jitter, and propagation delay spread. The transmitted signal is also corrupted by Gaussian noise with one-sided spectral density N_0. At the receiver, a DFE is implemented at baseband to jointly perform the equalization and carrier phase synchronization. The DFE has an $(n+1)$-tap feedforward filter and an m-tap feedback filter, as shown in Fig. 1(b), where the input to the feedforward filter is the received signal sequence $\{r_k\}$, and the input to the feedback filter is the decision on previous symbols $\{x_k\}$. The complex tap coefficients $\{c_{k-1,0}, c_{k-1,1}, \cdots, c_{k-1,n}, b_{k-1,1}, b_{k-1,2}, \cdots, b_{k-1,m}\}$ are jointly optimized with demodulation phase to minimize equalization and carrier phase synchronization error. For a slowly fading channel, the variation of the carrier phase jitter introduced by the channel over one symbol duration is very small $(\ll\pi)$, therefore, it is possible to combine the equalization with the phase demodulation (details are given in [10]). The differential encoding and decoding are necessary

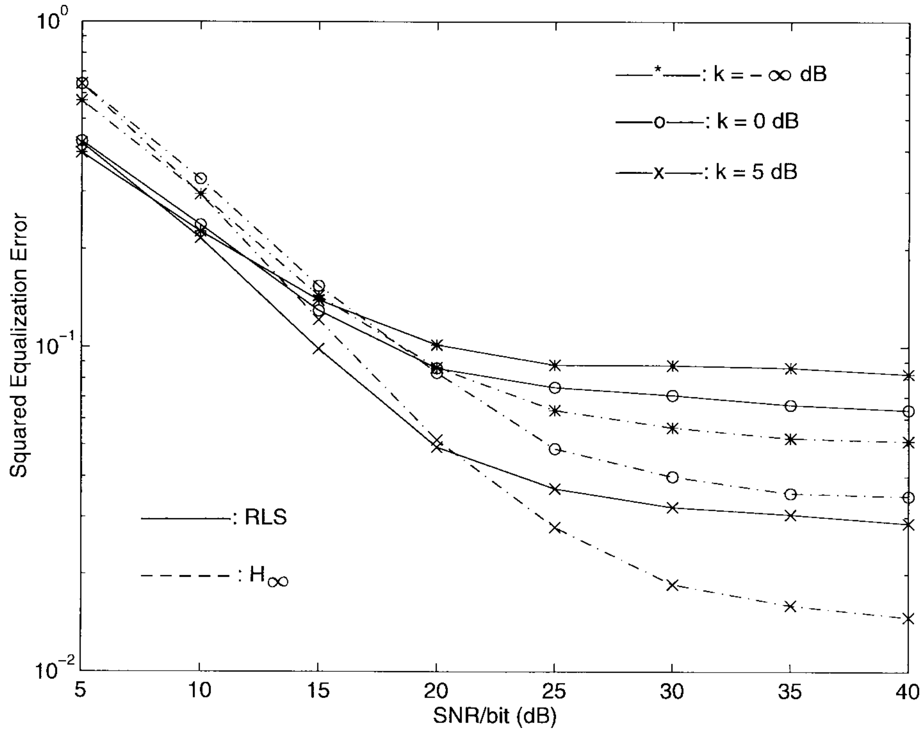


Fig. 2. Squared equalization error in the two-path Rician fading channel.

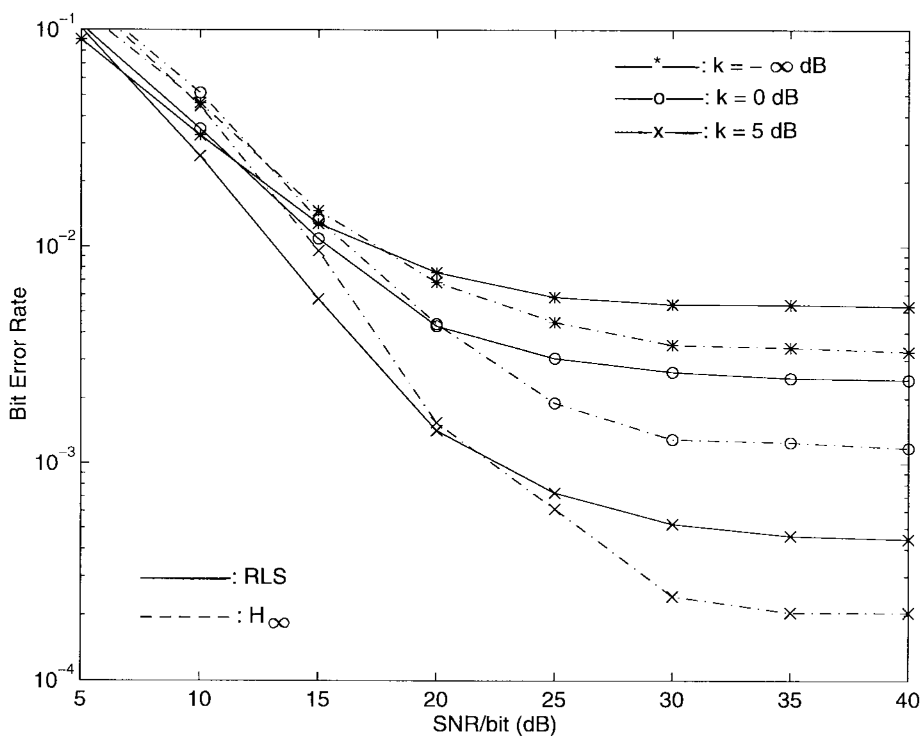


Fig. 3. BER performance in the two-path Rician fading channel.

in order to remove any possible phase ambiguity in the joint optimization.

The design of the DFE with the optimal tap-coefficient vector is considered in the following using: 1) the RLS algorithm, an L_2 minimum equalization error variance criterion, and 2) an H_∞ minimum equalization error spectrum criterion. The former is a modified Kalman filtering problem where a dynamical error weighting function is used. The latter is a minimaximization problem where the maximum "energy" of the estimation error over all disturbances is minimized.

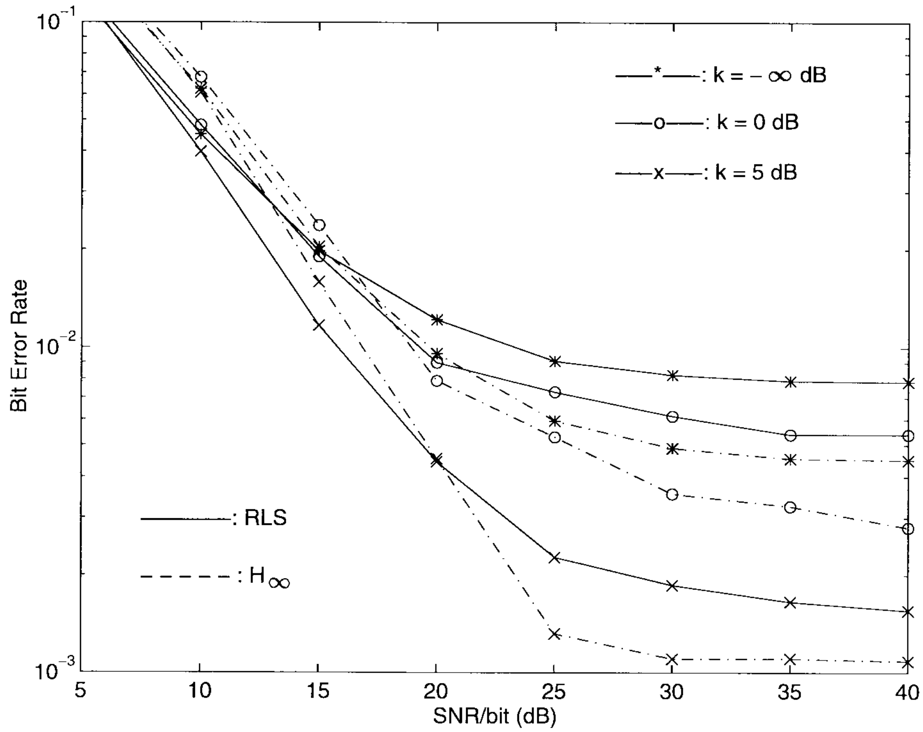


Fig. 4. BER performance in the three-path Rician fading channel.

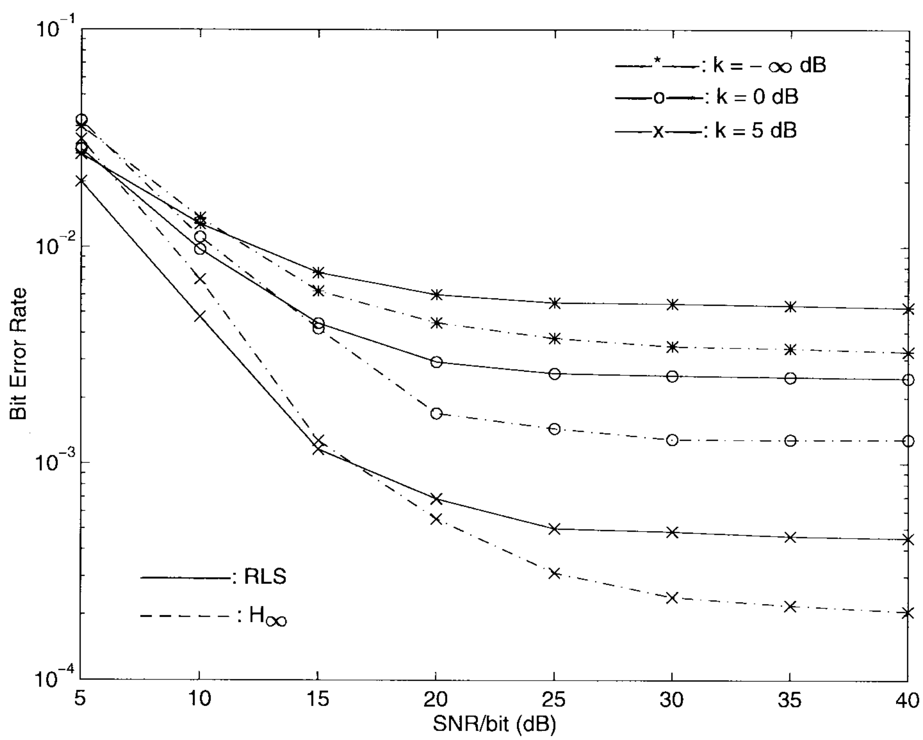


Fig. 5. BER performance in the two-path Rician fading channel with colored Gaussian noise.

B. The L_2 Filtering (RLS) Algorithm

In Fig. 1(b), the signal applied to the decision device is given by

$$y_k = \sum_{i=0}^{n} \hat{c}_{k-1-i} r_{k-i} - \sum_{i=1}^{m} \hat{b}_{k-1-i} x_{k-i} = U_k^H \hat{C}_{k-1} \quad (3)$$

where

$$U_k = [r_k^*, r_{k-1}^*, \cdots, r_{k-n}^*, -x_{k-1}^*, -x_{k-2}^*, \cdots, -x_{k-m}^*]^T$$
$$\hat{C}_{k-1} = [\hat{c}_{k-1,0}, \hat{c}_{k-1,1}, \cdots, \hat{c}_{k-1,n}, \hat{b}_{k-1,1}, \hat{b}_{k-1,2}, \cdots, \hat{b}_{k-1,m}]^T$$

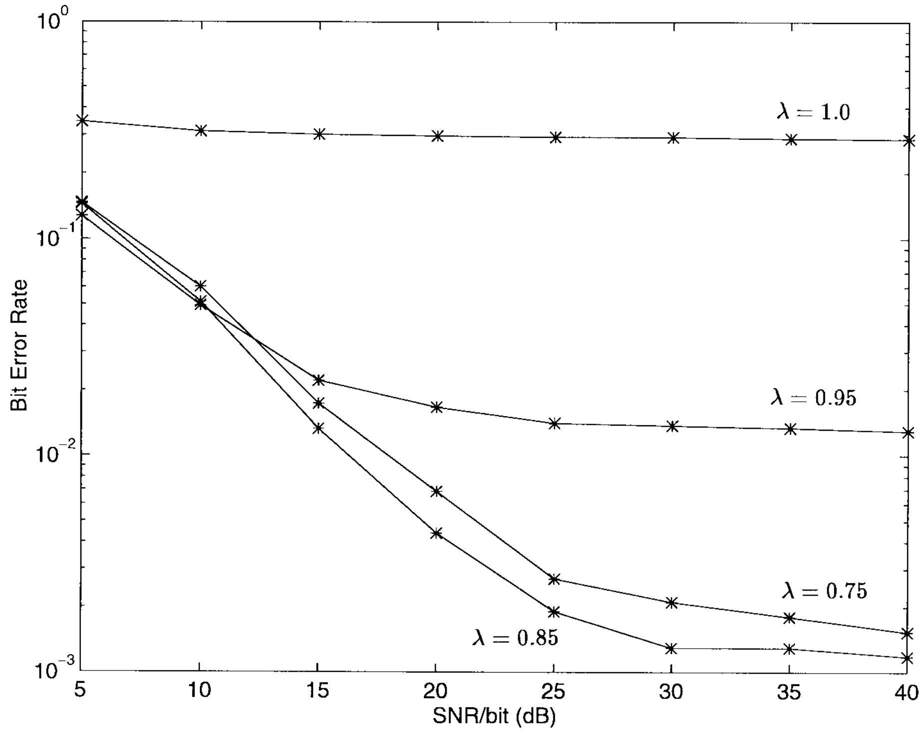


Fig. 6. BER performance in the two-path Rician fading channel using adaptive H_∞ algorithm with $k = 0$ dB and $f_DT = 0.005$.

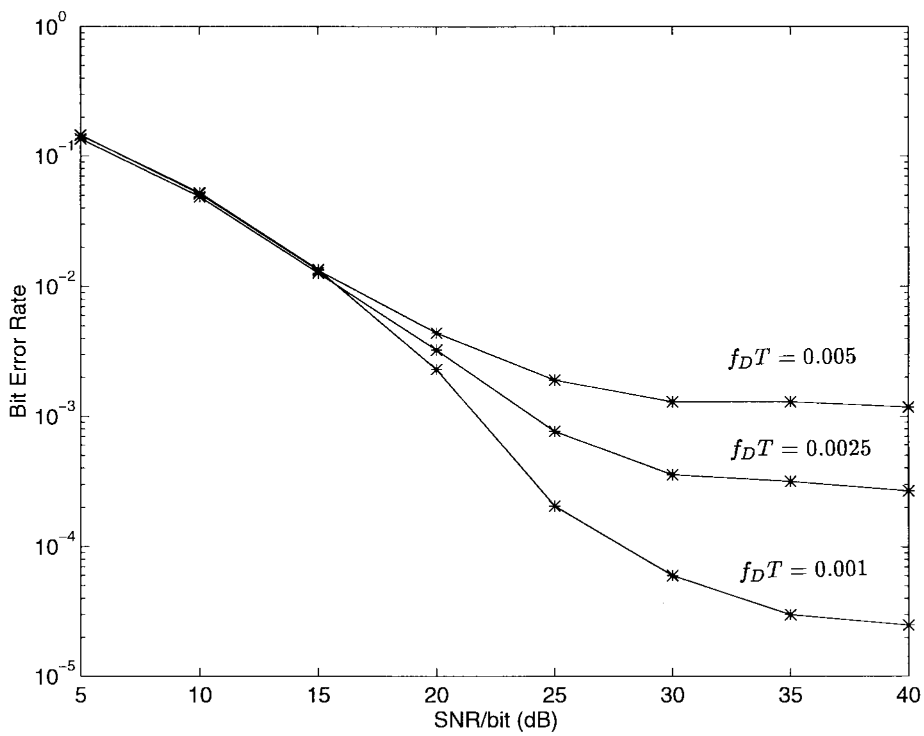


Fig. 7. BER performance in the two-path Rician fading channel using adaptive H_∞ algorithm with $k = 0$ dB and $\lambda = 0.85$.

are complex vectors, the superscript "$*$" denotes complex conjugation, "T" transposition, and "H" Hermitian transposition. $\hat{C}_{k-1}$ is an estimate of the optimal tap-coefficient vector C_{k-1} at $t = (k-1)T$, where C_{k-1} is time invariant over a number of symbol intervals (a reasonable assumption for a slowly fading channel). The estimate $\hat{C}_{k-1}$ is computed based on the received signals up to $t = (k-1)T$. The equalization error at $t = kT$ is defined as

$$\epsilon_k \triangleq d_k - y_k = d_k - U_k^H \hat{C}_{k-1} \tag{4}$$

where d_k is the corresponding desired signal. The design criterion of the RLS algorithm is to adaptively estimate the tap-

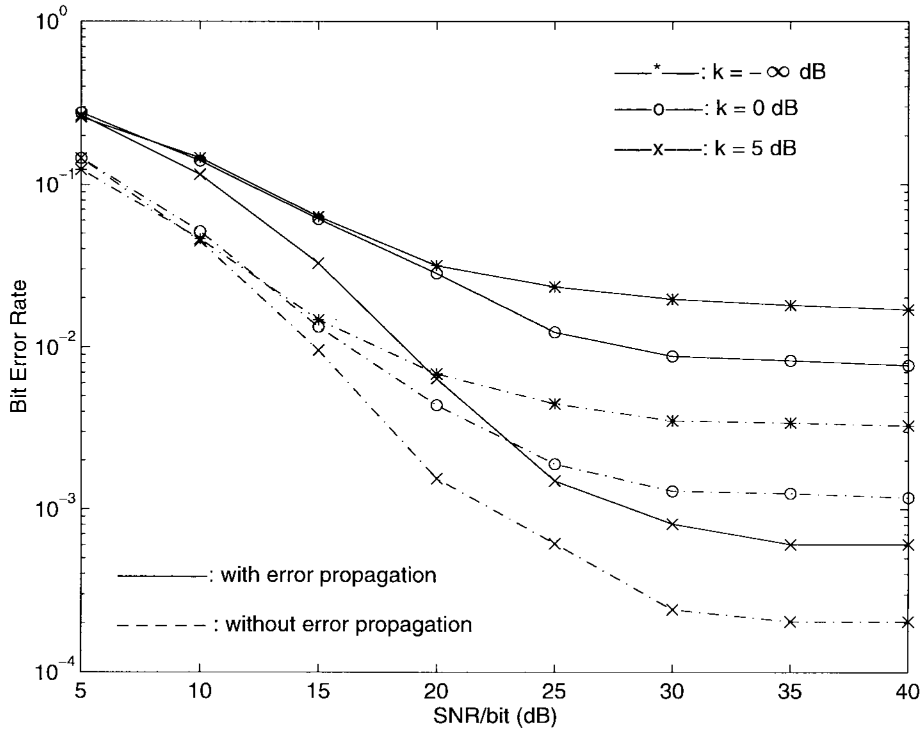


Fig. 8. BER performance in the two-path Rician fading channel with and without error propagation.

coefficient vector such that the weighted squared error (cost function) at $t = kT$, defined as

$$J = \sum_{i=0}^{k} \lambda^{k-i} |\epsilon_i|^2 = \sum_{i=0}^{k} \lambda^{k-i} \left| d_i - U_i^H \hat{C}_{i-1} \right|^2 \tag{5}$$

is minimized. In (5), λ^{k-i} is an exponential forgetting factor taking into account that the channel impulse response changes with time. If $\lambda = 1$, then all data are to be treated equally; if $\lambda < 1$, then the data obtained at earlier times are to have a smaller influence than more recent data. The RLS algorithm for updating the estimate of the tap-coefficient vector $\hat{C}_k$ can be summarized as [4], [5]

$$\hat{C}_k = \hat{C}_{k-1} + K_k \left(d_k - U_k^H \hat{C}_{k-1} \right)^* \tag{6}$$

$$K_k = P_k U_k \left(\lambda + U_k^H P_k U_k \right)^{-1} \tag{7}$$

$$P_{k+1} = \left(P_k - K_k U_k^H P_k \right) / \lambda \tag{8}$$

where the $(n+m+1)$-by-$(n+m+1)$ matrix P_k is defined as $P_k \triangleq [\sum_{i=0}^{k} \lambda^{k-i} U_i U_i^H]^{-1}$. The initial values of $\hat{C}_k$ and P_k can be chosen as $\hat{C}_0 = \vec{0}, P_0 = \delta I$ for a soft-constrained initialization, where $\delta \gg 1$ is a large positive constant and I is the identity matrix of $(n+m+1)$ dimension.

C. The H_∞ Filtering Algorithm

In general, the DFE can be described by the following state and measurement equations:

$$C_{k+1} = A_k C_k + v_k \tag{9}$$

$$d_k = U_k^H C_k + w_k \tag{10}$$

where C_k is the optimal tap-coefficient vector under the constraint of a finite tap number, A_k is the state transition matrix of $(n+m+1)$ dimension, v_k is a zero-mean random process representing system noise vector, and w_k is an unknown disturbance resulting from the input additive noise and the modeling error of the DFE (due to the finiteness of the tap numbers). We shall not make any assumptions on the disturbances v_k and w_k. The estimation error at $t_k = kT$ is defined as

$$e_k \triangleq U_k^H C_k - U_k^H \hat{C}_k. \tag{11}$$

In the L_2 filtering, we are interested in the optimal estimation of C_k, however, in the H_∞ filtering, we are interested in the estimation of $z_k = U_k^H C_k$, the linear combination of C_k's elements. As a result, $e_k = z_k - \hat{z}_k$. The measure of performance is defined as the transfer operator which transforms the disturbances (v_k and w_k) and the uncertainty of the initial tap-coefficient vector value (C_0) to the estimation error (e_k) [7]

$$J = \frac{\sum_{i=0}^{k} |e_i|_{Q_i}^2}{|C_0 - \hat{C}_0|_{P_0^{-1}}^2 + \sum_{i=0}^{k} |v_i|_{V_i^{-1}}^2 + \sum_{i=0}^{k} |w_i|_{W_i^{-1}}^2} \tag{12}$$

where the notation $|x|_G^2$ is defined as the square of the weighted (by G) L_2 norm of x, i.e., $|x|_G^2 \triangleq x^H G x$, $Q_k \geq 0$ and $W_k > 0$ are the weighting variables, and $V_k > 0$ and $p_0 > 0$ (i.e., positive definite) are the weighting matrices with $(n+m+1)$ dimension. The weighting variables and matrices are left to the choice of the designer and depend on performance requirement. The optimal estimate z_k among all possible $\hat{z}_k$ (i.e., the worst case performance measure) should satisfy

$$||J||_\infty = \sup_{(v_k, w_k, C_0)} J \leq \gamma^2 \tag{13}$$

where $\gamma^2 > 0$ is a prescribed level of disturbance attenuation and the supremum is taken with respect to the worst case noise, disturbance, and parameter uncertainty (v_k, w_k, C_0). It is assumed that the L_2 norms of $\{v_k\}$ and $\{w_k\}$ exist.

The design objective of the H_∞ filter is unique as compared with that of the L_2 filtering. It is to process the received signal samples (at baseband) to produce an estimate of the tap-coefficient vector $\hat{C}_k$ and, at the same time, to guarantee that the ∞ norm of the system error transfer function $\|J\|_\infty$ is less than or equal to a prescribed positive value γ^2. The discrete H_∞ filtering can be interpreted as a minimax problem where the estimator strategy $U_k^H \hat{C}_k$ plays against the exogenous inputs v_k and w_k and the uncertainty of the initial vector C_0. The performance criterion can be equivalently represented as

$$\min_{e_k} \max_{(v_k, w_k, C_0)} J = -\gamma^2 |C_0 - \hat{C}_0|^2_{p_0^{-1}} + \sum_{i=0}^{k} \left[|e_i|^2_{Q_i} - \gamma^2 \left(|w_i|^2_{W_i^{-1}} + |v_i|^2_{V_i^{-1}} \right) \right] \tag{14}$$

where "min" and "max" stand for minimization and maximization, respectively. The minimaximization problem can be solved by using a game theory approach. Following the results given in [8] and [9], it can be proved that for a given γ^2 value, there exists an H_∞ filter for z_k if and only if there exists a stabilizing symmetric solution $P_k > 0$ $(k = 1, 2, \cdots)$ to the following discrete-time Riccati equation:

$$\begin{aligned} P_{k+1} = {} & A_k P_k A_k^H - A_k P_k U_k \left(W_k + U_k^H P_k U_k \right)^{-1} U_k^H P_k A_k^H \\ & + V_k + \gamma^{-2} P_{k+1} U_{k+1} \\ & \cdot \left(Q_{k+1}^{-1} + \gamma^{-2} U_{k+1}^H P_{k+1} U_{k+1} \right)^{-1} U_{k+1}^H P_{k+1} \end{aligned} \tag{15}$$

$$P_0 = \left(p_0^{-1} - \gamma^{-2} U_0 Q_0 U_0^H \right)^{-1}. \tag{16}$$

If the solution P_k exists, then the H_∞ filtering algorithm can be described as

$$\hat{C}_k = A_{k-1} \hat{C}_{k-1} + K_k \left(d_k - U_k^H \hat{C}_{k-1} \right)^* \tag{17}$$

$$K_k = A_k P_k U_k \left(W_k + U_k^H P_k U_k \right)^{-1} \tag{18}$$

where the initial condition can be selected as $\hat{C}_0 = \vec{0}$ and $P_0 = \delta I$, with $\delta \gg 1$ being a large positive constant.

Comparing (12) and (13) with (5), we observe that: 1) in the H_∞ and RLS filtering, both equalization error sequence $\{\epsilon_i\}$ and estimation error sequence $\{e_i\}$ are treated as a deterministic process; 2) the RLS algorithm is to minimize the average squared equalization error (the L_2 norm of equalization error), therefore, it belongs to L_2 filtering; and 3) the H_∞ filtering is a worst case optimization because it is equivalent to minimize the effect of the worst disturbance on the estimation error.

D. Adaptive H_∞ Filtering Algorithm

In practice, it is very difficult (if not impossible) to obtain the channel state transition matrix A_k of (9) because the channel impulse response of a dispersive fading channel is a random process. Therefore, in reality it is not easy to implement the H_∞ algorithm described in (15)–(18), and an adaptive H_∞ algorithm is necessary. For a slowly fading channel, the channel impulse response is time invariant over a number of data symbol intervals. As a result, the optimal tap-coefficient vector can be assumed to be unchanged over the time interval for adjacent two updates of the vector. The assumption is reasonable as long as the normalized fading rate $f_D T_u \ll 1$ (where f_D is the maximum Doppler frequency shift of the fading channel and T_u is the time interval for each update of the vector). In this case, the state equation (9) can be approximated by

$$C_{k+1} = C_k \tag{19}$$

i.e., $A_k \approx I$ and $v_k \approx 0$. The approximation is adequate over a short duration of time (a number of symbol intervals), but inadequate over a long time interval. By introducing a forgetting factor to the filtering algorithm, we can overcome the limitation of the approximation over a long time interval. As a result, the approximation (19) is used in the following analysis.

In the case that the fading channel has a time-variant impulse response, the H_∞ filtering algorithm needs to be modified in order for the tap-coefficient vector C_k to change adaptively to the channel status. One approach is to introduce an exponential forgetting factor to the design criterion as in the case of the RLS algorithm, where the design criterion [(12) and (13)] can be modified to

$$\|F_1\|_\infty \triangleq \sup_{(w_k, C_0)} \frac{\sum_{i=0}^{k} \lambda^{k-i} |e_i|^2_{Q_i}}{\lambda^k |C_0 - \hat{C}_0|^2_{p_0^{-1}} + \sum_{i=0}^{k} \lambda^{k-i} |w_i|^2_{W_i^{-1}}} \leq \gamma^2. \tag{20}$$

In (20), the exponential term λ^{k-i} gives a larger weight to the more recent data (including the estimation error and disturbances). Equation (20) can be equivalently represented as

$$\|F_2\|_\infty \triangleq \sup_{(w_k, C_0)} \frac{\sum_{i=0}^{k} \lambda^{-i} |e_i|^2_{Q_i}}{|C_0 - \hat{C}_0|^2_{p_0^{-1}} + \sum_{i=0}^{k} \lambda^{-i} |w_i|^2_{W_i^{-1}}} \leq \gamma^2. \tag{21}$$

In order to derive an adaptive H_∞ filtering algorithm based on the design criterion (21), we introduce the following transformations so that we can represent $\|F_2\|_\infty$ in the same format as $\|J\|_\infty$ corresponding to (12) and (13). Let

$$\begin{aligned} &\bar{C}_k = \lambda^{-k/2} C_k \quad \hat{\bar{C}}_k = \lambda^{-k/2} \hat{C}_k \\ &\bar{e}_k = \lambda^{-k/2} e_k = U_i^H \bar{C}_k - U_k^H \hat{\bar{C}}_k \\ &\bar{d}_k = \lambda^{-k/2} d_k \quad \bar{w}_k = \lambda^{-k/2} w_k = \bar{d}_k - U_k^H \bar{C}_k \end{aligned} \tag{22}$$

then the system model can be described as

$$\bar{C}_{k+1} = A_k \bar{C}_k \tag{23}$$

$$\bar{d}_k = U_k^H \bar{C}_k + \bar{w}_k \tag{24}$$

where $A_k = \lambda^{-1/2} I$. The design criterion (21) can be correspondingly transformed to

$$\|\bar{F}\|_\infty \triangleq \sup_{(w_k, C_0)} \frac{\sum_{i=0}^{k} |\bar{e}_i|^2_{Q_i}}{|\bar{C}_0 - \hat{\bar{C}}_0|^2_{p_0^{-1}} + \sum_{i=0}^{k} |\bar{w}_i|^2_{W_i^{-1}}} \leq \gamma^2. \tag{25}$$

By using the newly defined variables (22), we have obtained the new system model [(23) and (24)] and design criterion (25) which correspond to the one with $\lambda = 1$ described by (9) and (10) and (12) and (13), respectively. Therefore, we can make use of the H_∞ filtering algorithm (15)–(18) for the new system model, as follows:

$$\hat{\bar{C}}_k = A_{k-1}\hat{\bar{C}}_{k-1} + \bar{K}_k\big(\bar{d}_k - U_k^H\hat{\bar{C}}_{k-1}\big)^* \quad (26)$$

$$\bar{K}_k = A_k\bar{P}_kU_k\big(W_k + U_k^H\bar{P}_kU_k\big)^{-1} \quad (27)$$

$$\begin{aligned}\bar{P}_{k+1} &= A_k\bar{P}_kA_k^H - A_k\bar{P}_kU_k\big(W_k + U_k^H\bar{P}_kU_k\big)^{-1}U_k^H\bar{P}_kA_k^H \\ &\quad + \gamma^{-2}\bar{P}_{k+1}U_{k+1}\big(Q_{k+1}^{-1} + \gamma^{-2}U_{k+1}^H\bar{P}_{k+1}U_{k+1}\big)^{-1} \\ &\quad \times U_{k+1}^H\bar{P}_{k+1}\end{aligned} \quad (28)$$

with the initial condition

$$\bar{P}_0 = \big(p_0^{-1} - \gamma^{-2}U_0Q_0U_0^H\big)^{-1}. \quad (29)$$

If we use (22) to retrieve the original variables, then from (26) and (27) we obtain

$$\hat{C}_k = \hat{C}_{k-1} + \lambda^{1/2}\bar{K}_k\big(\lambda^{-1/2}d_k - U_k^H\hat{C}_{k-1}\big)^* \quad (30)$$

$$\bar{K}_k = \lambda^{-1/2}\bar{P}_kU_k\big(W_k + U_k^H\bar{P}_kU_k\big)^{-1}. \quad (31)$$

By defining $K_k \triangleq \lambda^{1/2}\bar{K}_k$ and $P_k \triangleq \lambda\bar{P}_k$, we can rewrite (30) and (31) as

$$\hat{C}_k = \hat{C}_{k-1} + K_k\big(\lambda^{-1/2}d_k - U_k^H\hat{C}_{k-1}\big)^* \quad (32)$$

$$K_k = P_kU_k\big(\lambda W_k + U_k^HP_kU_k\big)^{-1}. \quad (33)$$

The matrix P_k can be calculated according to (28) and (29). Substituting $A_k = \lambda^{-1/2}I$ and $\bar{P}_k = \lambda^{-1}P_k$ into (28) and (29), we have

$$\begin{aligned}P_{k+1} &= \lambda^{-1}P_k - \lambda^{-1}P_kU_k\big(\lambda W_k + U_k^HP_kU_k\big)^{-1}U_k^HP_k \\ &\quad + \gamma^{-2}P_{k+1}U_{k+1}\big(\lambda Q_{k+1}^{-1} + \gamma^{-2}U_{k+1}^HP_{k+1}U_{k+1}\big)^{-1} \\ &\quad \times U_{k+1}^HP_{k+1}\end{aligned} \quad (34)$$

$$P_0 = \big(\lambda^{-1}p_0^{-1} - \lambda^{-1}\gamma^{-2}U_0Q_0U_0^H\big)^{-1}. \quad (35)$$

If the weighting variables and matrix are chosen to be $Q_k = 1$, $W_k = 1$, and $p_0 = \delta I$ and let $\tilde{\gamma} = \sqrt{\lambda}\gamma$, (34) can be written as

$$\begin{aligned}P_{k+1} &= (\lambda^{-1}P_k) - (\lambda^{-1}P_k)U_k\big[1 + U_k^H(\lambda^{-1}P_k)U_k\big]^{-1} \\ &\quad \times U_k^H(\lambda^{-1}P_k) + P_{k+1}U_{k+1}\big(\tilde{\gamma}^2 + U_{k+1}^H \\ &\quad \times P_{k+1}U_{k+1}\big)^{-1}U_{k+1}^HP_{k+1}\end{aligned} \quad (36)$$

which is equivalent to

$$\begin{aligned}P_{k+1} &- P_{k+1}U_{k+1}\big(\tilde{\gamma}^2 + U_{k+1}^HP_{k+1}U_{k+1}\big)^{-1}U_{k+1}^HP_{k+1} \\ &= (\lambda^{-1}P_k) - (\lambda^{-1}P_k)U_k\big[1 + U_k^H(\lambda^{-1}P_k)U_k\big]^{-1} \\ &\quad \times U_k^H(\lambda^{-1}P_k).\end{aligned} \quad (37)$$

Using the matrix inversion lemma [5]

$$B - BC(D + C^HBC)^{-1}C^HB = (B^{-1} + CD^{-1}C^H)^{-1}$$

we have

$$\big(P_{k+1}^{-1} + \tilde{\gamma}^{-2}U_{k+1}U_{k+1}^H\big)^{-1} = \big(\lambda P_k^{-1} + U_kU_k^H\big)^{-1}. \quad (38)$$

As a result, (34) and (35) can be simplified to

$$P_{k+1}^{-1} = \lambda P_k^{-1} + U_kU_k^H - \tilde{\gamma}^{-2}U_{k+1}U_{k+1}^H \quad (39)$$

$$P_0^{-1} = (\lambda\delta)^{-1}I - \tilde{\gamma}^{-2}U_0U_0^H. \quad (40)$$

In summary, the newly developed adaptive H_∞ filtering algorithm for the DFE can be described as

$$\begin{cases}\hat{C}_k = \hat{C}_{k-1} + K_k\big(\lambda^{-1/2}d_k - U_k^H\hat{C}_{k-1}\big)^* \\ K_k = P_kU_k\big(\lambda + U_k^HP_kU_k\big)^{-1} \\ P_{k+1}^{-1} = \lambda P_k^{-1} + U_kU_k^H - \tilde{\gamma}^{-2}U_{k+1}U_{k+1}^H\end{cases}$$

with the initial condition $\hat{C}_0 = \vec{0}$ and $P_0^{-1} = (\lambda\delta)^{-1}I - \tilde{\gamma}^{-2}U_0U_0^H$.

E. Recursive Algorithm for the Optimal $\tilde{\gamma}$

The matrix Riccati-type equation (39) with the initial condition (40) always has positive definite solutions as long as the parameter $\tilde{\gamma}$ is large enough. The adaptive H_∞ filtering problem can have many solutions (each for a proper $\tilde{\gamma}$ value), which is different from the L_2 filtering with only one solution. In fact, the equation reduces to the Riccati equation associated with the L_2 filter in the limit where $\tilde{\gamma}$ approaches infinity. In the design criterion (20), it is observed that the smaller the $\tilde{\gamma}$ value, the less effect the interference has on the equalization error. On the other hand, we require that the Riccati equation (39) to have a positive definite solution P_{k+1} $(k = 0, 1, 2, \cdots)$. As a result, the minimum $\tilde{\gamma}$ value at $t = (k+1)T$, $\tilde{\gamma}_{m,k+1}$, depends on λ, P_k, U_k, and U_{k+1}. In the following, we propose an algorithm for adaptively adjusting $\tilde{\gamma}$ value to its minimum at each iteration for $\hat{C}_{k+1}$.

For the adaptive H_∞ algorithm, from (39), in order for P_{k+1} to be positive definite, it requires

$$\begin{aligned}&\lambda P_k^{-1} + U_kU_k^H - \tilde{\gamma}_{k+1}^{-2}U_{k+1}U_{k+1}^H > 0 \\ &\quad\Rightarrow \lambda P_k^{-1} + U_kU_k^H > \tilde{\gamma}_{k+1}^{-2}U_{k+1}U_{k+1}^H \\ &\quad\Rightarrow \tilde{\gamma}_{k+1}^2 I > U_{k+1}U_{k+1}^H\big(\lambda P_k^{-1} + U_kU_k^H\big)^{-1} \\ &\quad\Rightarrow \tilde{\gamma}_{k+1}^2 I > \lambda^{-1}U_{k+1}U_{k+1}^H\big(I - K_kU_k^H\big)P_k \\ &\quad\Rightarrow \tilde{\gamma}_{k+1}^2 > \max\big\{\mathrm{eig}\big[\lambda^{-1}U_{k+1}U_{k+1}^H\big(I - K_kU_k^H\big)P_k\big]\big\} \\ &\quad\Rightarrow \tilde{\gamma}_{m,k+1} = \xi\max\big\{\mathrm{eig}\big[\lambda^{-1}U_{k+1}U_{k+1}^H \\ &\qquad \times \big(I - K_kU_k^H\big)P_k\big]\big\}^{0.5}\end{aligned} \quad (41)$$

where $\max\{\mathrm{eig}(A)\}$ denotes the maximum eigen value of the matrix A, and ξ is a constant close to one, but larger than one to ensure that $\tilde{\gamma}$ is always greater than the maximum square root eigen value of the matrix $\lambda^{-1}U_{k+1}U_{k+1}^H(I - K_kU_k^H)P_k$, but is close to it. It should be mentioned that the $\tilde{\gamma}_{m,k+1}$ value chosen to be too close to the maximum square root eigen value may result in that the matrix P_{k+1} is close to a singular matrix. This may cause an ill-conditioned numerical problem in solving the matrix Riccati-type equation. Singular perturbation methods can be used to overcome the problem [11]. On the other hand, the ξ value can be increased accordingly for $\tilde{\gamma}_{m,i} (i > k+1)$ to avoid the singularity. In fact, a $\tilde{\gamma}_{k+1}$ value larger than the minimum value may be desirable if the SNR value of the received signal is relatively low (i.e., if the equalization error mainly results from the input noise instead of the modeling error), which will be discussed in Section III.

Comparing the adaptive L_2 filtering algorithm with the adaptive H_∞ filtering algorithm, we observe the following: 1) the adaptive H_∞ algorithm [(32) and (33)] has an observer structure the same as that of the RLS algorithm [(6) and (7)] and 2) each filtering algorithm has its unique recursive algorithm for updating P_k. However, when $\tilde{\gamma}_k$ approaches infinity for all k, the adaptive H_∞ filtering (39) reduces to the RLS filtering (8); 3) the L_2 filter usually provides a good estimate of the tap-coefficient vector with minimum least squares error. On the other hand, since the adaptive H_∞ filter is designed based on an upper bound of the estimation error, it is more robust in the practical equalization of time-variant fading channels. The algorithm minimizes the H_∞ norm of the mapping from exogenous inputs (disturbance including noise and modeling error) to the estimation error. The strength of the adaptive H_∞ filter lies in its superior performance at the peak error range and in its impressive robustness to parameter uncertainty and modeling error; and 4) compared with the RLS algorithm, the adaptive H_∞ algorithm achieves better performance at the expense of an increased complexity if the optimal $\tilde{\gamma}$ value is evaluated and used at each iteration. According to (41), the evaluation involves the computations of products and summations of matrices and the maximum square root eigen value of the matrix.

III. Simulation Results and Discussion

The BER performance of coherent QPSK with differential encoding and decoding using a DFE with the adaptive H_∞ filtering algorithm and RLS algorithm, respectively, is evaluated by computer simulations. Time-variant dispersive Rayleigh and Rician fading channels are considered, with propagation delay between any two adjacent paths equal to one symbol interval T. It is assumed that only the first path may have a line-of-sight (LOS) component. The k-factor of the Rician amplitude fading is defined as the ratio of the average power of the LOS component to the average power of the diffusive component of the first path. The value of the k-factor is selected as 5, 0, and $-\infty$ dB (i.e., Rayleigh fading) in the following analysis. ρ_1 and ρ_2 are defined as the ratios of the average power of the first-path diffusive component to that of the second-path diffusive component and to that of the third-path diffusive component, respectively. SNR/bit is defined as the ratio of the ensemble average of the received signal power (per bit) from all the paths to the variance of the received additive white Gaussian noise except for Fig. 5 where colored noise is used. Two channel models are considered: 1) two-path channel with $\rho_1 = 0$ dB (i.e., both paths having the same average power of the diffusive component) and 2) three-path channel, with $\rho_1 = 1.0$ dB and $\rho_2 = 2.5$ dB. The normalized fading rate $f_D T$ is chosen to be 0.005 characterizing a slowly fading channel and the forgetting factor λ is chosen to be 0.85 except some specified situations. The DFE is assumed to have a three-tap feedforward filter and a two-tap feedback filter. The sampling interval is equal to one symbol interval T, and the tap-coefficient vector estimate $\hat{C}_k$ is updated once over each sampling interval (i.e., $T_u = T$).

A. Equalization Error

Fig. 2 shows the squared equalization error averaged over the first 5000 data samples versus SNR/bit in the two-path Rician fading channel. It is observed that: 1) the averaged equalization error decreases as the value of the k-factor increases no matter which algorithm is used. With a larger k-factor, the strength of the first-path signal increases and the degree of the first-path fading decreases, so that the ISI component (from the second path) is relatively reduced. As a result, the equalization error is also reduced; 2) when SNR/bit <20 dB, input noise is a dominant factor responsible for the equalization error. Since the channel is time variant and the forgetting factor $\lambda < 1$ is introduced in the filtering algorithms, the initial tap-coefficient vector uncertainty $(C_0 - \hat{C}_0)$ does not contributed much to the error averaged over 5000 data samples. In this case, we see that the L_2 has a smaller averaged equalization error than the H_∞ algorithm. In other words, the L_2 filtering algorithm is superior to the H_∞ algorithm in handling input white Gaussian noise; 3) when SNR/bit ≥ 20 dB, the equalization error corresponding to the adaptive H_∞ filtering algorithm is much smaller than that corresponding to the RLS algorithm, especially when the SNR/bit has a large value. The reason is that, as the SNR/bit value increases the effect of the input noise decreases and the equalization error mainly results from the modeling error. In this case, compared with the L_2 algorithm, the adaptive H_∞ filtering algorithm is much more powerful in suppressing the effect of the modeling error on the equalization error; and 4) when the SNR/bit increases to a very large value such as 30–40 dB, equalization error floors exist no matter which filtering algorithm is used, due to the residual ISI, especially when the desired path experiences deep fading.

B. BER Performance

Figs. 3–5 show the BER performance of the system using the DFE in the two-path Rician fading channel with additive white and colored Gaussian noise and three-path Rician fading channels with additive white Gaussian noise, respectively. The additive colored Gaussian noise is generated by applying a white Gaussian noise sequence $\{n_k\}$ to a linear filter with transfer function $H(z) = (1 - 0.9z)^{-1}$. It is observed that: 1) as the value of the k-factor increases, the BER floor decreases no matter whether the RLS algorithm or the adaptive H_∞ algorithm is used because the increase of the k-factor is equivalent to reduce the ISI (delayed signal) components. With a less severe ISI level, the equalization accuracy is increased, so that the BER performance is improved; 2) in the case that SNR ≤ 20 dB, the noise component accompanying the input signal is relatively large. The input noise is a key player for the equalization error and the effect of modeling error becomes negligible. Therefore, it is further verified that the L_2 algorithm is more capable of suppressing the effect of the input noise on the equalization error than the adaptive H_∞ filtering algorithm; 3) the DFE using the adaptive H_∞ filtering algorithm has a much better BER performance than that using the RLS algorithm when SNR/bit > 20 dB, especially when the k-factor has a large value. With a large SNR/bit value and a large k-factor, the equalization error primarily comes from the modeling error. As a result, it proves that the adaptive H_∞ filtering algorithm has a much better capability of combating the modeling error than the L_2 algorithm; 4) the BER performance over the three-path channel is worse than that over the two-path channel for the same k-factor and the same filtering algorithm. From the delay-power profiles of the channels, we know that the ISI component over the three-path channel is much more severe than that over the two-path channel. Moreover, since the DFE tap numbers are kept the same in both cases, the modeling error over the three-path channel is larger than that over the two-path channel. With a larger ISI component and a larger modeling error, the BER performance over the three-path channel is inferior, as expected; 5) comparing Fig. 3 with Fig. 5, the BER floor (for SNR/bit > 30 dB) does not change when the input additive Gaussian noise is changed from white to the colored. This is because the effect of input noise on the equalization error is negligible in the SNR/bit domain. The major error source is modeling error, which is the same as long as the channel fading statistics do not change; and 6) when the input noise component is relatively strong with SNR/bit < 20 dB, the BER performance over the fading channels with colored noise is better than the corresponding one with white noise. The colored noise has a strong correlation among noise samples over a number of symbol intervals. With the reduction of the randomness of the input colored noise, both L_2 and H_∞ filtering algorithms have a better ability to mitigate the effect on the noise on equalization performance.

From the above analysis of Figs. 3–5, we see that a compromise between the L_2 algorithm and the adaptive H_∞ algorithm can be made to obtain the best BER performance over a broad range of the SNR/bit value. The most desirable tradeoff would be one where the best estimator in the L_2 norm sense can be obtained among the set of filters that satisfies a given H_∞ error bound. However, so far no theoretical solution is known to achieve the tradeoff. According to the analysis in Section II, this objective can be achieved suboptimally by choosing $\tilde{\gamma}$ between $\tilde{\gamma}_{m,k}$, $k = 0, 1, 2, \cdots$, (corresponding to the optimal H_∞ filter) and ∞ (corresponding to the RLS filter). When the SNR/bit has a small value, a large $\tilde{\gamma}$ should be used to combat the input noise, otherwise, a $\tilde{\gamma}$ value close to that of $\tilde{\gamma}_{m,k}$ should be chosen to combat the modeling error.

C. Effect of the Forgetting Factor and Fading Rate

Fig. 6 illustrates the effect of the forgetting factor on the BER performance using the adaptive H_∞ filtering algorithm. When $\lambda = 1$, all data samples are treated equally in the calculation of the most current estimate of the tap-coefficient vector, no matter how long it has elapsed since each data was sampled. In other words, the filtering algorithm tries to find an estimate $\hat{C}_k$ that converges to a constant optimal value $C_k = C$. Since the channel impulse response is time variant, the optimal value C_k changes with time, and the estimate $\hat{C}_k$ diverges from its optimal value C_k as time elapses. As a result, a high BER floor is observed when $\lambda = 1$ due to a large modeling error. An increase of the SNR/bit value does not reduce the probability of transmission error. The BER performance with the DFE can be even worse than that without using the DFE. Therefore, it is necessary to have a λ value smaller than one, so that the filtering algorithm can adapt to channel fading dynamics. The value of λ determines how alert the adaptive H_∞ algorithm is able to equalize the time-variant fading channel. A small λ means that DFE can quickly adapt to a new channel status, and, at the same time, it is sensitive to input noise. The best choice of λ value is thus a tradeoff between alertness to the changes of the channel status and noise sensitivity. If the channel fading rate is fixed, then there exists an optimal λ value. This is verified by the simulation results shown in the figure. Using $\lambda = 0.85$ results in a smaller probability of transmission error than that using either $\lambda = 0.95$ or $\lambda = 0.75$, due to the fact that both input noise and system dynamic tracking error simultaneously deteriorate the transmission performance. In summary, the optimum λ value depends on how fast the channel status changes compared to the noise level in the input signal.

Fig. 7 shows the effect of the channel fading rate $f_D T$ on the BER performance using the adaptive H_∞ filtering algorithm, over the two-path channel with $k = 0$ dB, $\lambda = 0.85$. With a smaller value of $f_D T$, the channel fades more slowly, so that the estimate $\hat{C}_k$ converges to the optimal C_k relatively faster. Therefore, with a smaller equalization error, the BER performance is improved. As the channel fading rate $f_D T$ increases, the BER performance degrades. This can be mitigated in two ways: 1) by reducing the value of λ correspondingly (as discussed), so that the DFE can better track the variations of the channel and 2) by updating $\hat{C}_k$ more frequently (i.e., over a time interval less than a symbol interval), so that the variation of the channel status over the (smaller) interval is reduced. During the simulations, we have observed that the adaptive H_∞ filtering is less sensitive not only to the modeling error, but also to the exact knowledge of the system dynamics.

D. Effect of Error Propagation

In the above simulations, correct decisions are used in the decision feedback of the DFE, i.e., no error propagation effect is taken into account in the analysis. However, error propagation is inherent in the DFE. Once a decision error is made, it will propagate down the feedback filter. Fig. 8 demonstrates the effect of error propagation on the BER performance using the adaptive H_∞ filtering algorithm. As the value of the k-factor increases, the probability of bit error reduces, which also reduces the probability of error propagation. For example, as k increases from 0 to 5 dB, the error floor is reduced from 1.2 $\times$ 10^{-3} to 2 $\times$ 10^{-4} in the absence of error propagation and from 7.7 $\times$ 10^{-3} to 6.0 $\times$ 10^{-4} in the presence of error propagation. The improvement in the case of error propagation is sightly better than that of no error propagation. However, the pattern of the BER performance improvement as k increases with error propagation is very similar to that without error propagation.

IV. Conclusion

A new adaptive H_∞ filtering algorithm has been developed for the DFE in order to equalize indoor dispersive fading radio channels. The H_∞ filtering algorithm minimizes the effect of worst disturbance (including both input noise and modeling error) on the estimation error. A forgetting factor is introduced to the H_∞ algorithm in order for the DFE to track the variations of the channel status as time goes by. It has been shown that as $\tilde{\gamma}$ approaches infinity, the adaptive H_∞ filtering algorithm converges to the RLS algorithm. Computer simulation results have demonstrated that the adaptive H_∞ filtering algorithm has superior performance in suppressing the effect of modeling error on the equalization error and bit error rate and is less sensitive to the exact knowledge of channel dynamics, however, in the case of a large-input noise component, a compromise between the H_∞ and L_2 algorithms should be made in order to achieve better BER performance.

Acknowledgment

The author wishes to thank the anonymous reviewers for their helpful reviews and suggestions.

References

[1] W. Zhuang, W. A. Krzymien, and P. A. Goud, "Trellis-coded CPFSK and soft-decision feedback equalization for micro-cellular wireless applications," *Wireless Personal Commun.*, vol. 1, no. 4, pp. 271–285, 1995.
[2] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 74, pp. 1349–1387, Sept. 1985.
[3] D. Godard, "Channel equalization using a Kalman filter for fast data transmission," *IBM J. Res. Develop.*, pp. 267–273, May 1974.
[4] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*. Cambridge, MA: MIT Press, 1983, ch. 2.
[5] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: 1991, ch. 13.
[6] M. J. Grimble and A. Elsayed, "Solution of the H_∞ optimal linear filtering problem for discrete-time systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1092–1104, 1990.
[7] U. Shaked and Y. Theodor, "H_∞-optimal estimation: A tutorial," in *Proc. 31st IEEE CDC*, Tucson, AZ, Dec. 1992, pp. 2278–2286.
[8] X. Shen and L. Deng, "Discrete H_∞ filter design with application to speech enhancement," in *Proc. ICASSP'95*, Detroit, MI, May 1995, pp. 1504–1507.
[9] I. Yaesh and U. Shaked, "Game theory approach to state estimation of linear discrete-time processes and its relation to H_∞—Optimal estimation," *Int. J. Control*, vol. 55, pp. 1443–1452, 1992.
[10] D. D. Falconer, "Jointly adaptive equalization and carrier recovery in two-dimensional digital communication systems," *Bell Syst. Tech. J.*, vol. 55, no. 3, pp. 317–334, Mar. 1976.
[11] P. V. Kokotovic, H. K. Khalil, and J. O'Reilly, *Singular Perturbation Methods in Control: Analysis and Design*. New York: Academic, 1986.

Weihua Zhuang (M'93), for a photograph and biography, see this issue, p. 114.