*Wireless Personal Communications* **8:** 15–29, 1998.


# RLS Algorithm with Variable Forgetting Factor for Decision Feedback Equalizer over Time-Variant Fading Channels

WEIHUA ZHUANG
*Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*
*E-mail: wzhuang@bbcr.uwaterloo.ca*

**Abstract.** In a high-rate indoor wireless personal communication system, the delay spread due to multipath propagation results in intersymbol interference (ISI) which can significantly increase the transmission bit error rate (BER). Decision feedback equalizer (DFE) is an efficient approach to combating the ISI. Recursive least squares (RLS) algorithm with a constant forgetting factor is often used to update the tap-coefficient vector of the DFE for ISI-free transmission. However, using a constant forgetting factor may not yield the optimal performance in a nonstationary environment. In this paper, an adaptive algorithm is developed to obtain a time-varying forgetting factor. The forgetting factor is used with the RLS algorithm in a DFE for calculating the tap-coefficient vector in order to minimize the squared equalization error due to input noise and due to channel dynamics. The algorithm is derived based on the argument that, for optimal filtering, the equalization errors should be uncorrelated. The adaptive forgetting factor can be obtained based on on-line equalization error measurements. Computer simulation results demonstrate that better transmission performance can be achieved by using the RLS algorithm with the adaptive forgetting factor than that with a constant forgetting factor previously proposed for optimal steady-state performance or a variable forgetting factor for a near deterministic system.

**Key words:** RLS algorithm, adaptive forgetting factor, channel equalization, wireless communications

## 1. Introduction

In an indoor personal wireless communication system, 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 multipath fading. The indoor wireless channel is usually characterized by amplitude fluctuation, carrier phase jitter and propagation delay spread. In high bit rate transmission where the transmitted signal bandwidth is larger than the channel coherence bandwidth, the delay spread 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 [2]–[3]. Normally, the recursive least squares (RLS) algorithm [4]–[5] is used to adaptively adjust the DFE tap-coefficient vector to track the dynamics of fading channels in order to minimize the squared equalization error. The RLS algorithm achieves the best steady-state performance in a stationary environment. In the case of a nonstationary environment, the algorithm uses a forgetting factor $\lambda \in (0, 1)$ to obtain only a finite memory in order to track slow statistical variations of the channel fading status. The forgetting factor gives a larger weight to more recent data in order to cope with the channel dynamics. The introduction of the forgetting factor for a nonstationary environment results in the RLS algorithm having two types of excess error (the error above the mean-squared error of the Wiener solution), i.e., the estimation noise and error due to lag effects [5]–[6]. If $\lambda = 1$, all the data are weighted equally, and the algorithm has an infinite memory length, which is optimal with respect to suppressing the estimation noise effect alone. On the other hand, with a smaller $\lambda$ value, the algorithm has a shorter memory length and is better adapted to channel dynamics. In other words, a larger $\lambda$ will reduce the estimation noise, and a smaller $\lambda$ will reduce the equalization error due to lag effects. Therefore, the optimal $\lambda$ value depends on channel fading dynamics and the extent of input noise effect on the equalization error.

Several algorithms for optimal forgetting factor have been investigated to considerably improve the performance of the RLS algorithm for a near deterministic system in nonstationary environments [6]–[8]. In [6]–[7], algorithms for time-invariant optimal forgetting factor are proposed, which minimizes the mean-squared error due to input noise and due to lag effects. In this case, the forgetting factor is optimized with respect to the steady-state performance, which may not be optimal with respect to instantaneous system dynamics. In [8], a similar algorithm is investigated for dynamically adjusting the forgetting factor to control the degree of the tradeoff between the RLS algorithm's dynamics tracking ability and its input noise suppression ability. The forgetting factor is chosen according to an inverse function of the residual power in order to achieve a constant weighted sum of the squares of the *a posteriori* errors. In other words, the amount of forgetting will at each step correspond to the amount of new information in the lastest measurement, thereby ensuring that the estimation is always based on the same amount of information. The basic concept of the time-varying optimal forgetting factor can be explained as follows. In the case of a near deterministic system, the *a posteriori* estimation error provides the information about the state of the estimator. When the initial value of $\lambda$ is set to unity and the error is small, it may be concluded that the estimator is sensitive enough to adjust to the variations of the system parameters and therefore to significantly reduce the estimation error. As a result, it is reasonable to choose a forgetting factor close to unity. However, when the estimation error is large, the estimator sensitivity should be increased by choosing a smaller forgetting factor. This will reduce the effective memory length of the estimator until the estimator parameters are readjusted and the estimation error becomes smaller. This type of algorithms improves the tracking capability of the RLS algorithm for a near deterministic system which has abrupt jumps in system parameters or states. However, all the previously developed algorithms are not suitable for equalizing time-variant frequency-selective fading channels, due to the fact that the channel parameters change with time continuously and dynamically and that the system is far away from being near deterministic. Most previously work on time-variant fading channel equalization uses a constant forgetting factor which is selected based on trial and error method. The process for choosing a suitable forgetting factor has a two-fold drawback:

(i) the process can be time-consuming and the selected value may be far away from the optimal one for the channel status;

(ii) more importantly, when the channel normalized fading rate is not a constant (such as in the case that the mobile unit does not travel at a constant speed), the optimal forgetting factor is time-variant, and therefore, using a constant forgetting factor is not an appropriate approach. As a result, it is necessary to develop new algorithms for an adaptive forgetting factor in order to obtain optimal performance of the RLS algorithm for a personal wireless communication system where channel parameters or states dynamically change with time.

In this paper, an adaptive algorithm is proposed to calculate the time-variant forgetting factor for the RLS algorithm for a DFE to equalize time-varying frequency-selective fading channels. The algorithm takes account of the effect of both channel dynamics and input

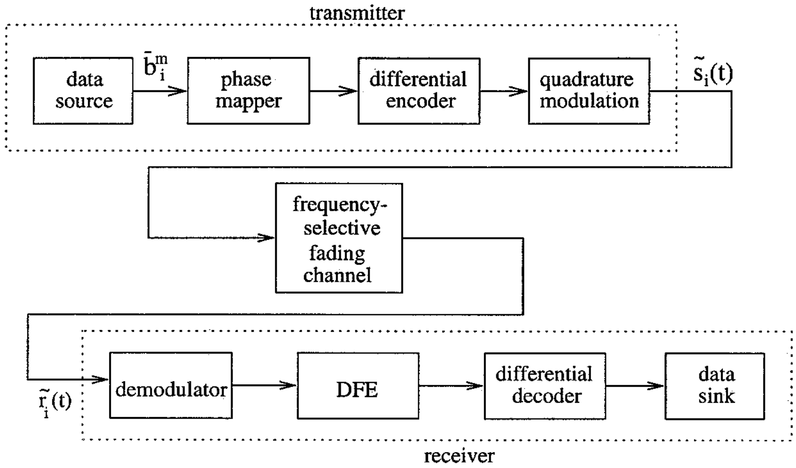



*Figure 1.* Functional block diagram of the system model.

noise on the equalization error. The new algorithm is developed according to the optimality condition of the Kalman filtering theory. Computer simulation results demonstrate that better transmission performance can be achieved by using the RLS algorithm with the adaptive forgetting factor than that with a constant forgetting factor previously proposed for optimal steady-state performance or a variable forgetting factor for a near deterministic system. The remainder of this paper is organized as follows. Section 2 describes the system model where the DFE with the RLS filtering algorithm is used. In Section 3, the algorithm for a variable forgetting factor is derived and the procedure of calculating the adaptive forgetting factor based on on-line measurement data is discussed. Section 4 presents the BER performance analysis of differential quadrature phase-shift keying (DQPSK) using the DFE, where the RLS algorithm uses constant and time-varying forgetting factors respectively. Conclusions of this work are given in Section 5.

## 2. System Description

The simplified functional block diagram of a personal wireless communication system is shown in Figure 1. Although differentially encoded $M$-ary phase-shift-keying (DMPSK) is considered here, the algorithm proposed in this paper can be directly applied to other modulation schemes.

### 2.1. Transmitter

The transmitter consists of a data source, a phase mapper, a differential encoder, and a quadrature modulator. The binary sequence is encoded, and modulated into the transmitted signal $\tilde{s}_i(t)$. The data source generates equiprobable, independent $m$-bit information words $\bar{b}_i^m = [b_i^1, b_i^2, \ldots, b_i^m]$, where $b_i^l \in \{0, 1\}$ for $1 \leq l \leq m$. The phase mapper converts the uncoded binary sequence $\bar{b}_i^m$ into an $M$-ary PSK symbol, where $M$ is the number of different symbols ($M = 2^m$). The mapping rule can be described in two steps. First, $\bar{b}_i^m$'s are converted into $\triangle\Phi_i = 2\pi(i-1)/M$ ($i = 1, 2, \ldots, M$) using Gray encoding. Then, a complex valued sequence $z_i$ is obtained by letting $z_i = \mathrm{e}^{j\triangle\Phi_i}$. The differential encoding applied to $z_i$ is also a complex operation, and can be described by $\mathrm{e}^{j\phi_i} = z_i\,\mathrm{e}^{j\phi_{i-1}}$, where $\phi_i$ represents the phase of the transmitted symbol. The relation between $\phi_i$ and $\triangle\Phi_i$ is given by

$$\phi_i = \triangle\Phi_i \oplus \phi_{i-1}, \tag{1}$$

where $\oplus$ represents modulo-$2\pi$ addition. The information signal to be transmitted over $t \in [iT, iT + T]$ (where $T$ is one symbol interval) is given by

$$\tilde{s}_i(t) = \sqrt{2}\,\mathrm{Re}\{\sqrt{P} s_i(t) \exp[j(2\pi f_c t + \phi_0)]\}, \tag{2}$$

where $s_i(t) = \exp(j\phi_i)s(t - iT)$ is the complex envelope of the signal, $P$ is the transmitted signal power, $f_c$ is the carrier frequency, $\phi_0$ is the carrier phase at $t = 0$, and $s(t)$ is the baseband pulse waveform.

### 2.2. FREQUENCY-SELECTIVE FADING CHANNEL

In the wireless mobile communication system, 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 [9]. 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 [10]–[11]

$$h(t) = \sum_{l=0}^{L} a_l(t)\,\mathrm{e}^{j\theta_l(t)}\delta(t - \tau_l), \tag{3}$$

where $L + 1$ is the total number of the distinguishable propagation paths; $a_l(t)$ $(l = 0, 1, 2, \ldots, L)$ is Rayleigh distributed amplitude fading of the $l$-path if there is no line-of-sight (LOS) component, Rician distributed otherwise; $\theta_l(t)$ is the carrier phase jitter of the $l$-path with a uniform distribution over $[-\pi, +\pi]$ for Rayleigh fading and a non-uniform distribution for Rician fading; $\tau_l$ is the propagation delay of the $l$-path; and $\delta(\cdot)$ is the Dirac delta function. The normalized power spectral density of the multipath diffusive components of each propagation path is given by

$$A(f) = \begin{cases} \frac{1}{\pi f_D \sqrt{1-(f/f_D)^2}}, & |f| \le f_D \\ 0, & |f| > f_D \end{cases}, \tag{4}$$

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.

The frequency-selective fading channel corrupts the transmitted signal $\tilde{s}_i(t)$ by introducing the multiplicative envelope distortion $a_l(t)$, the carrier phase disturbance $\theta_l(t)$, and propagation delay $\tau_l$ over the $l$th path. The multiple propagation paths result in time dispersion of the transmitted signal. The transmitted signal is also corrupted by Gaussian noise with one-sided spectral density $N_0$. For slowly fading channels, $f_D T \ll 1.0$, the channel amplitude fading

$a_l(t)$ and phase jitter $\theta_l(t)$ are approximately invariant over the period of a number of symbol intervals. Thus, the received signal over $t \in [iT, iT+T]$ can be represented as

$$\tilde{r}_i(t) = \sqrt{2}\,\mathrm{Re}\{r_i(t)\exp[j(2\pi f_c t + \phi_0)]\}, \tag{5}$$

where

$$r_i(t) = \sqrt{P}\sum_{l=0}^{L} a_l(iT)\,\mathrm{e}^{j\theta_l(iT)} s(t-\tau_l) + n(t) \tag{6}$$

is the complex envelope of the received signal and $n(t)$ is the complex envelope of the additive Gaussian noise.

### 2.3. Receiver

The basic units in the receiver include a demodulator, a DFE, a differential decoder, and a data sink. Here, the demodulator consists of an in-phase and a quadrature matched filters with a free-running oscillator which removes the carrier component $\exp(j2\pi f_c t)$ from the received signal. Assume that (i) $\tau_l = lT$ and (ii) the channel fades slowly, the output signal of the demodulator is discrete complex envelope of the received signal

$$r_i = \sqrt{P}\sum_{l=0}^{L} a_l(iT)\,\mathrm{e}^{j\theta_l(iT)} s_{i-l} + n_i, \tag{7}$$

where $s_{i-l} = \exp(j\phi_{i-l})$, $n_i$ is the equivalent Gaussian noise at baseband with zero mean and variance $\sigma_n^2$. Here, the equalization and carrier phase synchronization are jointly performed by the DFE at baseband. The DFE has an $(m_1+1)$-tap feedforward filter and an $m_2$-tap feedback filter, as shown in Figure 2, where the input to the feedforward filter is the received signal sequence $\{r_i\}$, and to the feedback filter is the decision on previous symbols $\{x_i\}$. The complex tap-coefficient vector $\{c_{i-1,0}, c_{i-1,1}, \ldots, c_{i-1,m_1}, d_{i-1,1}, d_{i-1,2}, \ldots, d_{i-1,m_2}\}$ is jointly optimized with demodulation phase to minimize equalization and carrier phase synchronization error. For the slowly fading channel, the variation of carrier phase jitter, $\theta_l$ $(l = 0, 1, \ldots, L)$, 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 [12]). The differential decoder is to counteract the differential encoder of the transmitter for recovering the original information data. The differential encoding and decoding are necessary in order to remove any possible phase ambiguity in the joint optimization.

## 3. Adaptive Forgetting Factor for RLS Algorithm

In this section, a new adaptive algorithm for time-variant forgetting factor is derived to optimize the tap coefficients of the DFE in order to achieve ISI-free transmission over time-varying fading channels.

### 3.1. The RLS algorithm for DFE

In a stationary environment, the DFE can be described by the following system model equations

$$C_{i+1} = C_i, \tag{8}$$

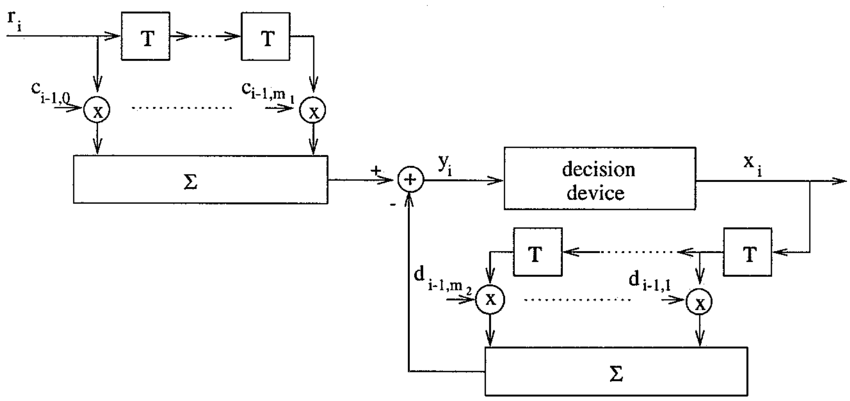



*Figure 2.* Structure of the DFE.

$$v_i = U_i^H C_i + w_i, \tag{9}$$

where $C_i$ is the optimal tap-coefficient vector under the constraint of a finite tap number, $v_i$ is the ideal signal (ISI free and noise free) at the ouput of the DFE, $w_i$ is the measurement error with zero mean and variance $\sigma_w^2$ resulting from the input additive noise and modeling error of the DFE (due to the finiteness of the tap numbers), and $U_i$ is the input vector of the DFE

$$U_i = [r_i^*, r_{i-1}^*, \ldots, r_{i-m_1}^*, -x_{i-1}^*, -x_{i-2}^*, \ldots, -x_{i-m_2}^*]^T$$

the superscript "$*$" denotes complex conjugation, "$T$" transposition, and "$H$" Hermitian transposition (i.e., the operation of transposition combined with complex conjugation). Equation (8) is adequate over a short duration of time (a number of symbol intervals) for a slowly fading channel. However, the model is inadequate over a long time interval, which is to be overcome by introducing an exponential forgetting factor to the filtering algorithm. In Figure 2, the signal applied to the decision device is given by

$$y_i = \sum_{k=0}^{m_1} \hat{c}_{i-1-k} r_{i-k} - \sum_{k=1}^{m_2} \hat{d}_{i-1-k} x_{i-k} = U_i^H \hat{C}_{i-1}, \tag{10}$$

where

$$\hat{C}_{i-1} = [\hat{c}_{i-1,0}, \hat{c}_{i-1,1}, \ldots, \hat{c}_{i-1,m_1}, \hat{d}_{i-1,1}, \hat{d}_{i-1,2}, \ldots, \hat{d}_{i-1,m_2}]^T$$

is an estimate of the optimal tap-coefficient vector $C_{i-1}$ at $t = (i-1)T$, under the assumption that $C_{i-1}$ is time-invariant over a number of symbol intervals (a reasonable assumption for a slowly fading channel). The estimate $\hat{C}_{i-1}$ is computed based on the received signals up to $t = (i-1)T$. The optimal complex tap coefficient vector $C_{i-1}$ for the feedforward and feedback transversal filters should provide the carrier phase compensation and, at the same time, achieve ISI-free transmission. The received signal from the path with the maximum average power is to be considered as the desired signal component. The received signal components from all other paths are either postcursor or precursor, and are to be eliminated by the DFE. The ideal signal applied to the decision device is $v_i = s_{i-l_1-l_2}$, where $0 \le l_1 \le L$

is an integer due to propagation delay, and $0 \leq l_2 \leq m_1$ is an integer due to the delay units in the feedforward filter. The equalization error at $t = iT$ is defined as

$$\varepsilon_i \stackrel{\triangle}{=} v_i - y_i = v_i - U_i^H \hat{C}_{i-1}. \tag{11}$$

In the RLS algorithm, the equalization error sequence $\{\varepsilon_i\}$ is considered to be a deterministic process. We start the computation with an initial estimate $\hat{C}_0$ and use the information contained in new data samples to update the old estimates. Therefore, the length of observable data is variable. The design criterion is to adaptively estimate the tap-coefficient vector such that the weighted squared error (cost function) at $t = iT$, defined as

$$J_i = \sum_{k=0}^{i} \lambda^{i-k} |\varepsilon_k|^2 = \sum_{k=0}^{i} \lambda^{i-k} |v_k - U_k^H \hat{C}_{k-1}|^2 \tag{12}$$

is minimized. In (12), $\lambda^{i-k}$ 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 with a constant $\lambda$ for updating the estimate of the tap-coefficient vector, $\hat{C}_i$, can be summarized as [4]–[5]

$$\hat{C}_i = \hat{C}_{i-1} + K_i (v_i - U_i^H \hat{C}_{i-1})^*, \tag{13}$$

$$K_i = P'_{i-1} U_i (1 + U_i^H P'_{i-1} U_i)^{-1}, \tag{14}$$

$$P_i = (P_{i-1} - K_i U_i^H P_{i-1})/\lambda, \tag{15}$$

where the $(m_1+m_2+1)$-by-$(m_1+m_2+1)$ matrix $P_i$ is defined as $P_i \stackrel{\triangle}{=} [\Sigma_{k=1}^{i} \lambda^{i-k} U_k U_k^H]^{-1}$ and $P'_{i-1} = P_{i-1}/\lambda$. The initial values of $\hat{C}_i$ and $P_i$ can be chosen as

$$\hat{C}_0 = \mathbf{0}, \qquad P_0 = \delta I,$$

for a soft-constrained initialization, where $\delta \gg 1$ is a large positive constant, and $I$ is the identity matrix of $(m_1 + m_2 + 1)$ dimension.

In applying the RLS algorithm to the DFE, it is assumed that the channel impulse response is approximately time-invariant over a small time duration. That is, the optimal tap coefficients of the DFE are assumed to be fixed during the observation interval $1 \leq i \leq m_1$. In order to overcome the error of the channel modeling (8) over a long time interval, the exponential forgetting factor with a constant $\lambda$ is used in the filtering algorithm (13)–(15). Usually, trial and error method can be used to search for "good" $\lambda$ values, which may not be optimal and the process can be very burdensome. Furthermore, the optimal value of $\lambda$ depends on the instantaneous dynamics of the channel and the noise component in the input signal.

### 3.2. Adaptive Algorithm for Optimal Forgetting Factor

For the RLS algorithm to be able to track the dynamics of a time-variant fading channel and at the same time to suppress the effect of the received noise on the equalization error, based on the Kalman filtering theory, the optimal filtering gain $K_i$ should be chosen in such a way that the equalization error $\{\varepsilon_i\}$ is an uncorrelated noise sequence. With the optimal filtering, the equalization error has zero mean, $E[\varepsilon_i] = 0$, and it can be derived that the variance of the error is

$$E(\varepsilon_i \varepsilon_i^*) = \sigma_w^2 S_i, \tag{16}$$

where

$$S_i = 1 + U_i^H P_{i-1}' U_i \tag{17}$$

is the normalized variance with respect to $\sigma_w^2$. Furthermore, it can be derived that the autocorrelation function of the equalization error is

$$\begin{aligned} R(j) &= E(\varepsilon_i \varepsilon_{i+j}^*) \\ &= \sigma_w^2 U_{i+j}^H \left[ \prod_{k=0}^{j-1} (1 - K_{i+k} U_{i+k}^H) \right] (\lambda_i^{-1} P_{i-1} U_i - K_i S_i), \quad j = 1, 2, \ldots . \end{aligned} \tag{18}$$

Substituting (14) and (17) into (18), we have $R(j) = 0$ for $j = 1, 2, \ldots$. This verifies that the equalization error sequence $\{\varepsilon_i\}$ is uncorrelated when the optimal gain $K_i$ is used. In practical situations, the normalized variance of the error, $S_i$, is different from the theoretical one given in (17) due to errors in the system model (8) in a long time interval. Thus, the actual measurement of $R(j)$ may not be zero. From (18), we see that if the forgetting factor can be chosen in such a way that the last term of $R(j)$, which is the only common term for all $j = 1, 2, \ldots$, be zero

$$\lambda_i^{-1} P_{i-1} U_i - K_i S_i = 0 \tag{19}$$

then $K_i$ is optimal. In other words, if the gain is optimal, then (19) holds. From (19) and (14)–(15), the corresponding forgetting factor is

$$\lambda_i = \frac{\lambda_{i-1}(1 + U_i^H P_{i-1}' U_i)}{S_i}. \tag{20}$$

It should be mentioned that for a time-varying fading channel, $S_i$ in (20) is calculated based on equalization error $\varepsilon_i$ (16) instead of (17). With the forgetting factor $\lambda_i$ (20) changing with time, (15) should be modified to

$$P_i = (P_{i-1} - K_i U_i^H P_{i-1}) / \lambda_i \tag{21}$$

and $P_i' = P_i / \lambda_i$. In the case of a time-invariant channel, the system model (8) is accurate, the filtering gain $K_i$ is optimal and (17) indeed gives the normalized variance of the equalization error. As a result, from (20), $\lambda_i$ can be chosen to be equal to unity, which is consistent with the performance analysis of the RLS algorithm [5].

### 3.3. On-Line Implementation Issues

An unbiased estimate of $S_i$ can be obtained from the measurement data according to the following equation

$$\hat{S}_i = \frac{1}{\sigma_w^2} \frac{1}{i-1} \sum_{k=1}^{i} \varepsilon_k \varepsilon_k^*, \quad i = 1, 2, \ldots . \tag{22}$$

Taking into account of fading channel dynamics, we can introduce the forgetting factor $\lambda_i$ to the estimate in order to give a larger weight to more recent equalization error. Equation (22) can be modified to

$$\hat{S}_i = \frac{1}{\sigma_w^2} \frac{\sum_{k=1}^{i-1} [\prod_{m=k}^{i-1} \lambda_m] \varepsilon_k \varepsilon_k^* + \varepsilon_i \varepsilon_i^*}{1 + \sum_{k=1}^{i-1} \prod_{m=k}^{i-1} \lambda_m}. \tag{23}$$

The estimate can be calculated recursively as

$$\hat{S}_i = \frac{1}{\sigma_w^2} \frac{\lambda_{i-1} \hat{S}_{i-1} + \varepsilon_i \varepsilon_i^*}{\gamma_i}, \tag{24}$$

where

$$\gamma_i = 1 + \lambda_{i-1} \gamma_{i-1}, \tag{25}$$

with initial condition $\lambda_0 = 1$, $\hat{S}_0 = 0$ and $\gamma_0 = 0$.

Due to effects of input noise and channel dynamics, there may exist errors in the $S_i$ estimation. As a result, (20) may give a $\lambda_i$ value which is larger than 1. Indeed, none of the previously proposed algorithms for variable forgetting factor ensures $\lambda_i \in (0, 1]$ [8]. Thus, in practice, it is necessary to place a reasonable limits on $\lambda_i$ value.

In the case that the received signal has a low signal-to- noise ratio (SNR) value, the input noise can significantly deteriorate the accuracy of the $S_i$ estimation (24). As a result, the on-line estimate of $\lambda_i$ value obtained from (20) will fluctuate in the neighborhood of its optimal value. For a slowly fading channel, the optimal value of $\lambda_i$ changes slowly as $i$ increases. As a result, in order to reduce the undue fluctuations of the on-line $\lambda_i$ estimate, a low-pass (LP) filter can be used to mitigate the effect of input noise. The LP filter can be implemented by limiting the variation of $\lambda_i$ over each sampling interval.

In summary, the RLS algorithm with the adaptive forgetting factor for the tap coefficient vector can be described by the flowchart shown in Figure 3. It includes the following steps:

(i) setting the initial values — $\hat{C}_0 = \mathbf{0}$, $U_0 = \mathbf{0}$, $P_0 = \delta I$, $\lambda_0 = 1$, $\hat{S}_0 = 0$, and $\gamma_0 = 0$;

(ii) calculating the filtering gain $K_i$ based on $P_{i-1}$ and the current input signal $U_i$ according to (14), and updating the estimate of the tap-coefficient vector $\hat{C}_i$ (13) based on its previous value $\hat{C}_{i-1}$, current gain $K_i$, input $U_i$ and desired signal $v_i$;

(iii) updating $\gamma_i$ (25) based on $\lambda_{i-1}$ and $\gamma_{i-1}$, and updating $\hat{S}_i$ (24) based on current $\gamma_i$ and equalization error $\varepsilon_i$, previous $\lambda_{i-1}$ and $\hat{S}_{i-1}$;

(iv) updating the forgetting factor $\lambda_i$ (20) based on current input $U_i$ and the estimate $\hat{S}_i$, previous $P_{i-1}$ and $\lambda_{i-1}$, and limiting the variation of $\lambda_i$ according to the following rule

$$\lambda_i = \lambda_{i-1} + \mu \cdot \mathrm{sgn}(\lambda_i - \lambda_{i-1}), \tag{26}$$

where $\mathrm{sgn}(\cdot)$ is the signum function and $\mu$ is the step size whose value should be proportional to the channel fading rate and inversely proportional to the SNR value of the received signal;

(v) updating $P_i$ (21) based on its previous value $P_{i-1}$, current forgetting factor $\lambda_i$, input $U_i$ and filtering gain $K_i$, and updating $P_i'$ accordingly based on $P_i$ and $\lambda_i$. It should be mentioned that using the LP filtering algorithm (26) for the adaptive forgetting factor, estimation of $\sigma_w^2$ is not necessary for the on-line implementation of the algorithm. In fact, $\sigma_w^2$ can be set as a constant in computer simulations according to the SNR range of targeted applications.

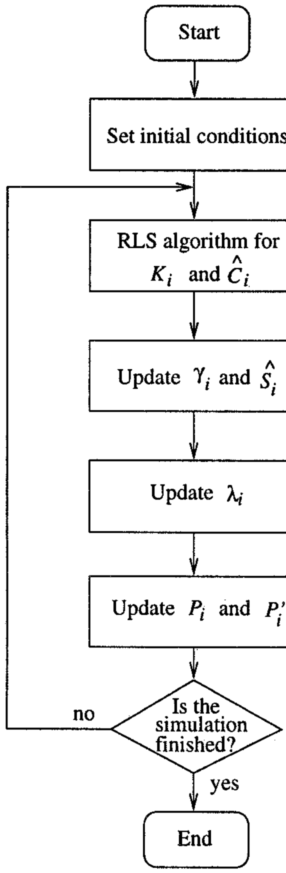



*Figure 3.* Flowchart of the RLS algorithm with adaptive forgetting factor.

## 4. Simulation Results and Discussion

The BER performance of the communications system shown in Figure 1 with DQPSK modulation is evaluated by computer simulations. The RLS algorithm with the adaptive forgetting factor is used to update the tap coefficient vector of the DFE according to channel status. Two-path time-variant Rayleigh and Rician fading channels are considered, with propagation delay between the two paths equal to one symbol interval $T$. It is assumed that the two paths fade independently and only the first path may have a LOS component and that both paths have the same average power of the diffusive 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 $k$-factor is selected as 5 dB, 0 dB and $-\infty$ dB (i.e., Rayleigh fading) in the following analysis. SNR/bit is defined as the ratio of the ensemble average of the received signal power (per bit) from both paths to the variance of the received (additive) white Gaussian noise at baseband. The normalized fading rate $f_D T$ is chosen to be 0.005, characterizing a slowly fading channel. The diffusive component (having Rayleigh distributed amplitude fading) with the classical Doppler spectrum (4) is simulated using the “sum of sine waves" method [9]. 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. The DFE is assumed to have a 3-tap feedforward filter and a 2-tap feedback filter. The sampling interval is equal to one symbol interval $T$, and the tap-coefficient vector estimate $\hat{C}_i$ (13) is updated once over each sampling interval. 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 number of bit-error occurrences is at least 100 and the number of fading cycles that the system experiences is at least 50, in order to guarantee the accuracy of the simulation results. The step size $\mu$ in (26) is chosen to be 0.001; however, it is observed in the simulations that the adaptive forgetting factor is not sensitive to the value of $\mu$ in the neighborhood of the chosen value.

Figures 4–6 show the BER performance of the system using a DFE with a constant forgetting factor ($\lambda = 0.95$), a variable forgetting factor [8], and the proposed adaptive forgetting factor, respectively, in the two-path Rayleigh/Rician fading channels. It is observed that: (i) The algorithms for constant forgetting factor previously proposed in [6]–[7] for optimal steady-state performance are not applicable to equalizing the time-variant fading channels discussed here. The algorithms are to compensate the steady-state misadjustment, which give a constant $\lambda$ value in the range of $[0.9, 1.0]$. Using the optimal constant $\lambda$ values, the BER performance of the system is in the neighborhood of that when $\lambda = 0.95$ is used. In addition, the algorithms require to estimate $\sigma_w^2$ and other statistics of the received signal sequence and estimation error sequence; (ii) The algorithm of variable forgetting factor given in [8] was derived for near deterministic systems with sudden parameter jumps. The time-varying forgetting factor does not provide satisfactory performance for a system with continuously time-varying parameters. Indeed, the variable forgetting factor gives the worst system BER performance; (iii) The effectiveness of the algorithm for the adaptive forgetting factor developed in this paper is clearly demonstrated by the simulation results. The forgetting factor gives the RLS algorithm a good balance between the input noise suppression capability and system dynamics tracking capability. The system using the adaptive forgetting factor achieves the best BER performance; (iv) Since the previous decisions are used in the feedback filter of the DFE, the decision device is likely to make error decisions once an error occurs. The feedback errors cause the DFE to adjust itself away from its optimum settings and produce an ill-equalized output which is likely to cause further decision errors. The inherent error propagation of the DFE is the major reason for the BER floors observed in the figures; (v) As the value of the $k$-factor increases, the BER floor decreases no matter which $\lambda$ value is used, because the increase of the $k$-factor is equivalent to reduce the ISI (the delayed signal) components. With a less severe ISI, the equalization accuracy is increased, so that the BER performance is improved.

Figure 7 shows the average value of the adaptive forgetting factor and the variable forgetting factor, respectively, as a function of the SNR/bit for the Rician fading channel with $k$-factor equal to $-\infty$ dB, 0 dB and 5 dB respectively. It is observed that: as the SNR/bit increases the value of both adaptive and variable forgetting factor decreases, so that the RLS algorithm can more effectively handle the equalization error due to channel dynamics as the error component becomes a dominant factor responsible for transmission errors. However, The variable forgetting factor decreases at a much slower rate than the adaptive forgetting factor, because the former is to target a near deterministic system where a $\lambda$ value close to unity is appropriate. It is also observed that the adaptive forgetting factor increases as the $k$-factor increases. This can be explained from the following two points of view: (i) In the two-path fading channel model, both propagation paths have the same average power of the diffusive signal component due to multipath. With a larger $k$-factor, the constant LOS component is increased, resulting in a relatively decreased time-varying diffusive component. As a result, the

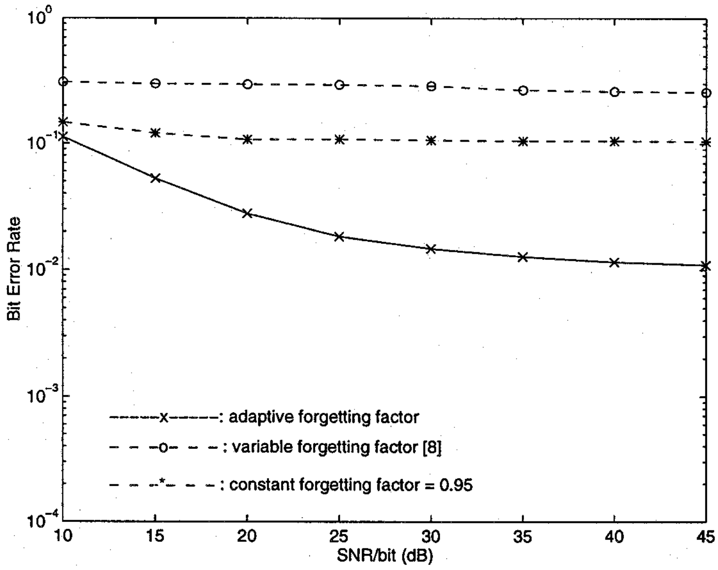



*Figure 4.* BER performance of DQPSK in the two-path Rayleigh fading channel.

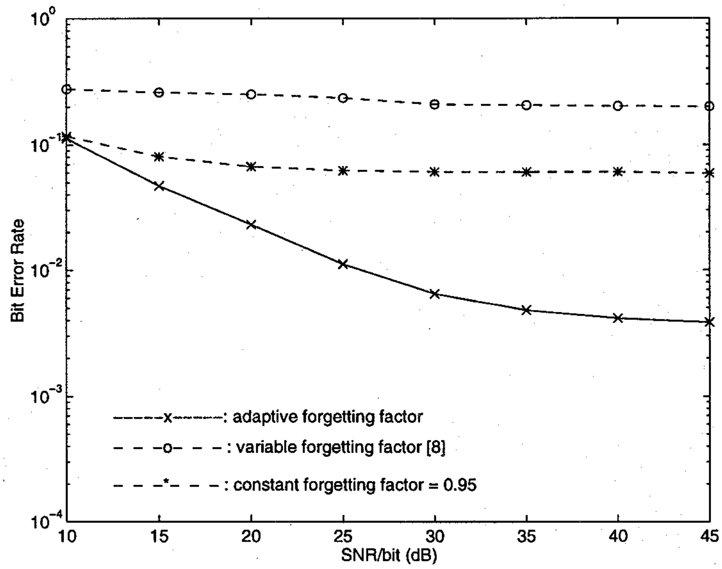



*Figure 5.* BER performance of DQPSK in the two-path Rician fading channel $(k = 0\,\text{dB})$.

dynamics of the fading channel is relatively reduced as compared with the case of a smaller $k$-factor, and the adaptive forgetting factor increases correspondingly; (ii) With a larger $k$-factor, the average received signal power is increased. Therefore, for the same received SNR/bit value, the actual input noise power increases as the $k$-factor increases. This makes the input noise have a larger contribution to the equalization error as compared with that for a smaller $k$-factor. In order to cope with the increased effect of the input noise, the adaptive algorithm increases the value of $\lambda$ accordingly. This demonstrates that the algorithm for the adaptive forgetting factor is able to distinguish the equalization error component due to input noise and

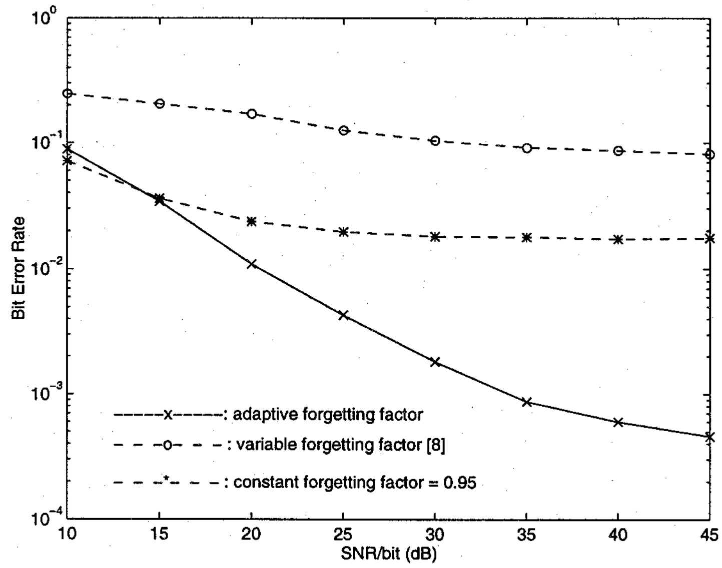



*Figure 6*. BER performance of DQPSK in the two-path Rician fading channel $(k = 5\,\mathrm{dB})$.

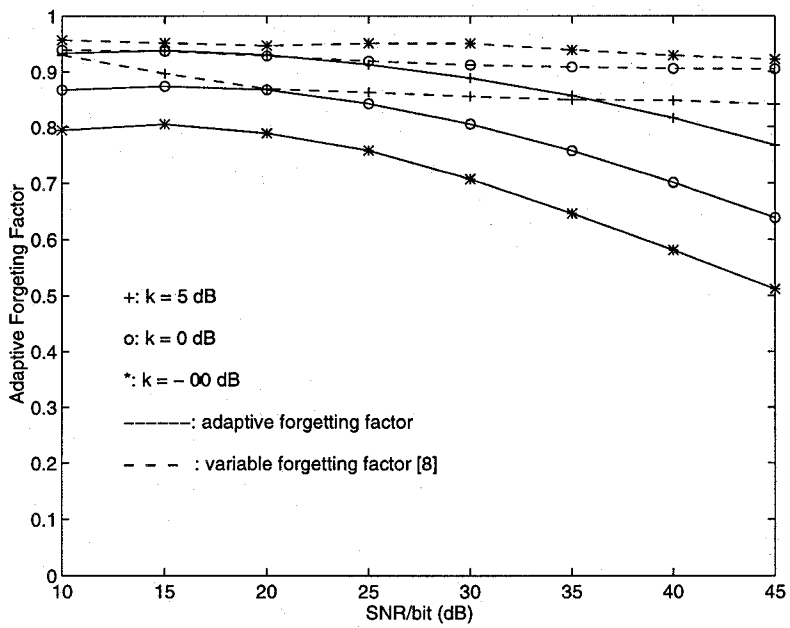



*Figure 7*. Average forgetting factor for the two-path Rayleigh/Rician fading channels.

that due to channel dynamics and to deal with the two types of error efficiently. On the other hand, the variable forgetting factor decreases as the $k$-factor increases because the equalization error increases as the noise component in the input signal increases. The algorithm developed in [8] reduces the forgetting factor as long as the equalization error increases. It does not have the capability of handling the equalization error due to input noise differently from that due to channel dynamics since the algorithm is developed for a near deterministic system.

From the above analysis, we conclude that a compromise between the equalization error due to input noise and that due to channel dynamics should be made to obtain the best BER performance over a broad range of the SNR/bit value. The compromise can be implemented by choosing a proper time-variant $\lambda$ value. The on-line implementation of the adaptive $\lambda$ value gives a suboptimal solution, which is a promising approach.

## 5. Conclusions

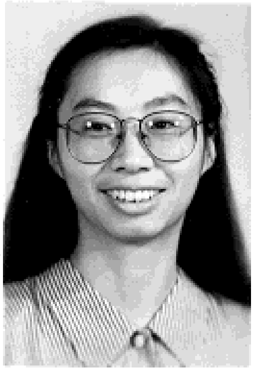
A new adaptive algorithm for time-variant forgetting factor has been developed for the RLS algorithm to be used in a DFE in order to equalize an indoor frequency-selective fading channel. The forgetting factor is introduced in order for the DFE to track the variations of the channel status as time goes by. The RLS filtering algorithm with the adaptive forgetting factor minimizes the effect of both channel dynamics and input noise on the equalization error. Computer simulation results have demonstrated that, the RLS filtering algorithm with the adaptive forgetting factor is superior to that using the previously proposed algorithms and to that using a constant forgetting factor in transmission performance and in the simplicity of finding a forgetting factor suitable for channel fading statistics.

## Acknowledgements

This work is supported by the research grant (OGP0155131) from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

## 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. Stein, “Fading channel issues in system engineering”, *IEEE J. Select. Areas Commun.*, Vol. SAC-5, No. 2, pp. 68–89, Feb. 1987.
3. S.U.H. Qureshi, “Adaptive equalization”, *Proc. IEEE*, Vol. 74, No. 9, pp. 1349–1387, Sept. 1985.
4. L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, 1983. Chapter 2.
5. S. Haykin, *Adaptive Filter Theory*, 2nd Edition, Prentice Hall: Englewood Cliffs, 1991. Chapter 13.
6. E. Eleftheriou and D.D. Falconer, “Tracking properties and steady-state performance of RLS adaptive filter algorithms”, *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 34, No. 5, pp. 1097–1109, Oct. 1986.
7. S.D. Peters and A. Antoniou, “A parallel adaptation algorithm for recursive-least-squares adaptive filters in nonstationary environments”, *IEEE Trans. Signal Processing*, Vol. 43, No. 11, pp. 2484–2495, Nov. 1995.
8. T.R. Fortescue, L.S. Kershenbaum, and B.E. Ydstie, “Implementation of self-tuning regulators with variable forgetting factors”, *Automatica*, Vol. 17, No. 6, pp. 831–835, 1981.
9. W.C. Jakes (ed.), *Microwave mobile communications*, New York: Wiely, 1974.
10. A.A.M. Saleh and R.A. Valenzuela, “A statistical model for indoor multipath propagation”, *IEEE J. Select. Areas Commun.*, Vol. SAT-5, No. 2, pp. 128–137, Feb. 1987.
11. H. Hashemi, “The indoor radio propagation channel”, *Proc. IEEE*, Vol. 81, No. 7, pp. 943–967, July 1993.
12. 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.



**Weihua Zhuang** 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. 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.