Least Squares Frequency Estimation in Frequency-Selective Channels and Its Application to Transmissions With Antenna Diversity

Eui-Rim Jeong, Sung-Kwon Jo, and Yong H. Lee, *Senior Member, IEEE*

***Abstract*—A new data-aided frequency estimator for frequency-selective fading channels is introduced. The proposed estimator is developed based on a least squares (LS) error criterion and can estimate frequency offsets without the need for channel information. Statistical analysis indicates that the resulting estimate is unbiased and tends to approach the Cramér–Rao lower bound (CRLB). Simulation shows that the proposed LS method is preferable to existing techniques in mobile communications. The application of the LS estimator to systems with transmitter antenna diversity is also considered. In particular, it is demonstrated that the LS method can be successfully applied to third-generation wireless communication systems [1].**

***Index Terms*—Frequency estimation, frequency-selective channels, least squares error criterion, transmitter antenna diversity.**

I. Introduction

RECENT communication systems tend to impose stringent requirements on the frequency stability of the transmitter and receiver oscillators. For example, third-generation partnership project (3GPP) recommendations for 3G wireless communication systems [1] demand that the modulated carrier frequency of the user equipment should be accurate within 0.1 ppm compared to the carrier frequency received from the base station. One way to relieve this need is to recover the carrier frequency at a receiver via signal processing.

Various techniques have been proposed for carrier frequency recovery [2]–[14]. Among these, data-aided techniques [4]–[14], which use a training sequence for frequency offset estimation, are popular because they can attain a good performance with a short training sequence. Data-aided techniques have been developed for additive white Gaussian noise (AWGN) channels [4]–[10] and then extended to flat fading [11], [12] and frequency-selective fading channels [13], [14]. When considering flat or frequency-selective fading channels, knowledge of the channel parameters is limited, if not unavailable, because channels are usually estimated after carrier frequency recovery. In the case of flat fading channels, the estimators in [11] and [12] need the Doppler bandwidth of the channel, which must be either assumed [11] or estimated [12]. The estimator in [13] requires statistics of the frequency-selective channels. Consequently, the performance of the estimators in [11]–[13] is degraded when the channel information is inaccurate. In [14], a maximum likelihood (ML) estimate is derived for frequency-selective channels. This estimator does not need explicit channel information and exhibits excellent behavior for a fixed channel; yet, its performance can be considerably degraded when the channel is time-varying and its implementation needs heavy computation.

In this paper, we develop an alternative approach to carrier frequency recovery. The proposed method is based on a least squares (LS) error criterion that does not need any prior knowledge about the channel. It will be shown that the proposed LS estimator can outperform those in [13] and [14] in time-varying channels and can be simpler to implement than the estimator in [14]. In addition, the proposed method is applied to the 3GPP code division multiple access (CDMA) systems with transmitter antenna diversity [21], [22].

The organization of this paper is as follows. The signal model is presented in Section II. The proposed estimator is developed in Section III and statistically analyzed in Section IV. Section V presents simulation results demonstrating the advantage of the proposed estimator over existing methods. In Section VI, the proposed estimator is applied to a system with transmitter diversity.

II. Communication System Model

The baseband system model considered in this paper is shown in Fig. 1. Here $d(j)$ denotes the transmitted M-ary PSK (or QAM) symbols, $h(t)$ is the baseband pulse shape, $n(t)$ is AWGN, θ is the initial random phase, and Δf represents the carrier frequency offset. In this model, all the pulse shaping is performed at the transmitter, and the receiver filter is a simple noise rejection filter. The output of the receiver filter sampled at $t = kT$ is

$$r(k) = e^{j(2\pi\Delta f kT+\theta)} \sum_{l=0}^{L} d(k-l) g_k(l) + \eta(k) \qquad (1)$$

Manuscript received February 1, 2001; revised June 14, 2001. This work was supported in part by the Korea Science and Engineering Foundation through the MICROS Center at KAIST, Korea. This paper was presented in part at the IEEE Vehicular Technology Conference (VTC), Rhodes, Greece, May 2001.

E.-R. Jeong and Y. H. Lee are with the Division of Electrical Engineering, KAIST, Yusong-gu, Taejon 305-701 Korea (e-mail: july@stein.kaist.ac.kr; yohlee@ee.kaist.ac.kr).

S.-K. Jo is with Samsung Electronics Co., Ltd., Sungnam-si, Kyungki-do, Korea.

Publisher Item Identifier S 0733-8716(01)10314-8.

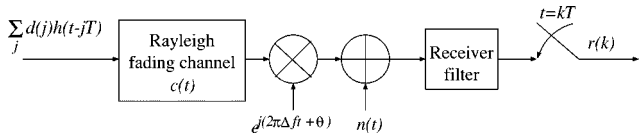

Fig. 1. Baseband system model.

where $g_k(l)$ is the impulse response of the equivalent channel at time k due to an impulse that is applied l time units earlier. It describes both $h(t)$ and $c(t)$ in the discrete time domain, and its duration is $L+1$. $\eta(k)$ is assumed to be AWGN with a variance of $2\sigma_\eta^2$.

III. Least Squares Frequency Offset Estimation

Suppose that K training symbols $\{d(k)|k=1, \ldots, K\}$ are available and that $\{g_k(l)\}$ represents a fixed ISI channel over the training period, i.e., $g_k(l) = g(l)$ for $k = 1, \ldots, K$. The problem considered in this section is stated as follows: given $\{d(k)|k=1, \ldots, K\}$ and $\{r(k)|k=1, \ldots, K\}$, estimate Δf without prior knowledge of $\{g(l)|l = 0, \ldots, L\}$. Estimating Δf directly from $\{r(k)\}$ is difficult, because this involves estimating the time-varying phase $2\pi\Delta fkT+\theta$ in (1). To overcome this difficulty, the product $r(k)r^*(k-m)$ which is denoted by $\gamma_m(k)$ is evaluated. This product is written as

$$\begin{aligned}\gamma_m(k) &\triangleq r(k)r^*(k-m)\\ &= \left(\sum_{l=0}^{L}\sum_{i=0}^{L} d(k-l)d^*(k-m-i)g(l)g^*(i)\right)\\ &\quad\cdot e^{j2\pi\Delta fmT} + n_m(k)\\ &= \mathbf{d}_m^T(k)\mathbf{g}e^{j2\pi\Delta fmT} + n_m(k) \end{aligned} \tag{2}$$

where $n_m(k)$ represents all the noise terms due to $\eta(k)$. Both $\mathbf{d}_m(k)$ and $\mathbf{g}$ are $(L+1)^2$ dimensional vectors given by

$$\begin{aligned}\mathbf{d}_m(k) \triangleq\ & [d(k)d^*(k-m),\ d(k)d^*(k-m-1),\ \ldots,\\ & d(k-1)d^*(k-m),\ d(k-1)d^*(k-m-1),\ \ldots,\\ & d(k-L)d^*(k-m-L)]^T \end{aligned} \tag{3}$$

$$\begin{aligned}\mathbf{g} \triangleq\ & [|g(0)|^2,\ g(0)g^*(1),\ \ldots,\ g(1)g^*(0),\ |g(1)|^2,\\ & g(1)g^*(2),\ \ldots,\ |g(L)|^2]^T. \end{aligned} \tag{4}$$

Note that $2\pi\Delta fmT$ in (2) represents a fixed phase shift. The symbols $\{d(k)\}$ that appear in forming $\mathbf{d}_m(k)$ in (3) should be the training symbols. This leads to the following range for k:

$$m+L+1 \le k \le K. \tag{5}$$

Using vector notation, $\{\gamma_m(k)|m+L+1 \le k \le K\}$ can be represented as

$$\begin{aligned}\boldsymbol{\gamma}_m &= \mathbf{D}_m\mathbf{g}e^{j2\pi\Delta fmT} + \mathbf{n}_m\\ &= \mathbf{D}_m\mathbf{p}_m + \mathbf{n}_m \end{aligned} \tag{6}$$

where $\boldsymbol{\gamma}_m \triangleq [\gamma_m(m+L+1), \gamma_m(m+L+2), \ldots, \gamma_m(K)]^T$, $\mathbf{n}_m \triangleq [n_m(m+L+1), n_m(m+L+2), \ldots, n_m(K)]^T$, $\mathbf{D}_m$ is a $(K-m-L)\times(L+1)^2$ dimensional matrix defined as $\mathbf{D}_m \triangleq [\mathbf{d}_m(m+L+1), \mathbf{d}_m(m+L+2), \ldots, \mathbf{d}_m(K)]^T$ and $\mathbf{p}_m \triangleq \mathbf{g}e^{j2\pi\Delta fmT}$. Before proceeding further, an example illustrating the structures of the matrix and vectors in (6) is presented.

Example 1: Suppose that there are 14 training symbols ($K = 14$) and that the channel has two taps ($L = 1$). When $\mathrm{m} = 8$, $\boldsymbol{\gamma}_m$ has five elements and $\mathbf{D}_m$ is a 5×4 matrix. Equation (6) is written as shown at the bottom of the page. ■

The proposed estimator is derived using (6). The procedure for its derivation consists of two steps: an LS estimate of $\mathbf{p}_m$, say $\hat{\mathbf{p}}_m$ is obtained first, then Δf is estimated from $\hat{\mathbf{p}}_m$. In the LS method, $\mathbf{p}_m$ is chosen so as to minimize the following sum of the error squares:

$$J(\mathbf{p}_m) \triangleq \sum_{k=m+L+1}^{K} |\gamma_m(k) - \mathbf{d}_m^T(k)\mathbf{p}_m|^2. \tag{7}$$

The LS estimate $\hat{\mathbf{p}}_m$ minimizing $J(\mathbf{p}_m)$ is expressed as

$$\hat{\mathbf{p}}_m = (\mathbf{D}_m^H\mathbf{D}_m)^{-1}\mathbf{D}_m^H\boldsymbol{\gamma}_m \tag{8}$$

if $\mathbf{D}_m^H\mathbf{D}_m$ is nonsingular (or equivalently, if the columns of $\mathbf{D}_m$ are linearly independent) [15]. This solution holds for determined and overdetermined cases, that is, when the row dimension of $\mathbf{D}_m$ is greater than or equal to its column dimension.[1] The range of m that meets this condition is

$$1 \le m \le K-L-(L+1)^2. \tag{9}$$

To obtain $\hat{\mathbf{p}}_m$, a proper training sequence that leads to a nonsingular $\mathbf{D}_m^H\mathbf{D}_m$ is needed. The frequency offset

[1]The underdetermined case is ignored to simplify the derivation and reduce the computational cost for implementing the proposed estimator.

$$\begin{bmatrix} r(10)r^*(2)\\ r(11)r^*(3)\\ r(12)r^*(4)\\ r(13)r^*(5)\\ r(14)r^*(6)\end{bmatrix} = \begin{bmatrix} d(10)d^*(2) & d(10)d^*(1) & d(9)d^*(2) & d(9)d^*(1)\\ d(11)d^*(3) & d(11)d^*(2) & d(10)d^*(3) & d(10)d^*(2)\\ d(12)d^*(4) & d(12)d^*(3) & d(11)d^*(4) & d(11)d^*(3)\\ d(13)d^*(5) & d(13)d^*(4) & d(12)d^*(5) & d(12)d^*(4)\\ d(14)d^*(6) & d(14)d^*(5) & d(13)d^*(6) & d(13)d^*(5)\end{bmatrix} \begin{bmatrix} |g(0)|^2\\ g(0)g^*(1)\\ g^*(0)g(1)\\ |g(1)|^2\end{bmatrix} e^{j2\pi\Delta f8T} + \mathbf{n}_8$$

Δf can be estimated by examining the elements of $\hat{\mathbf{p}}_m$. Let $\hat{\mathbf{p}}_m = [\hat{p}_m(1), \hat{p}_m(2), \ldots, \hat{p}_m((L+1)^2)]^T$. Since $\mathbf{p}_m = \mathbf{g}e^{j2\pi\Delta f mT}$, then $\hat{\mathbf{p}}_m = \hat{\mathbf{g}}e^{j2\pi\Delta\hat{f}mT}$, where $\hat{\mathbf{g}}$ and $\Delta\hat{f}$ are estimates of $\mathbf{g}$ and Δf, respectively. Then elements of $\hat{\mathbf{p}}_m$ are compared with the corresponding elements of $\hat{\mathbf{g}}e^{j2\pi\Delta\hat{f}mT}$. For the first element from (4)

$$\hat{p}_m(1) = |\hat{g}(0)|^2 e^{j2\pi\Delta\hat{f}mT}. \tag{10}$$

The key to the derivation of the proposed estimate is the observation that only the magnitude of the channel estimate $\hat{g}(0)$ appears in (10). Owing to this fact, $2\pi\Delta\hat{f}mT$ is equal to the phase of $\hat{p}_m(1)$, and $\Delta\hat{f}$ is represented as

$$\Delta\hat{f} = \frac{1}{2\pi mT}\arg\{\hat{p}_m(1)\}. \tag{11}$$

For the second element of $\hat{\mathbf{p}}_m$

$$\hat{p}_m(2) = \hat{g}(0)\hat{g}^*(1)e^{j2\pi\Delta\hat{f}mT}. \tag{12}$$

In (12), unfortunately, the channel estimates $\hat{g}(0)$ and $\hat{g}^*(1)$ appear in complex form. Since the channel information is unavailable, $\Delta\hat{f}$ cannot be evaluated from $\hat{p}_m(2)$. Continuing in this manner, elements of $\hat{\mathbf{p}}_m$ are selected from which $\Delta\hat{f}$ is obtained. There are $(L+1)$ such elements which are expressed as

$$\hat{p}_m((L+2)i+1) = |\hat{g}(i)|^2 e^{j2\pi\Delta\hat{f}mT} \tag{13}$$

for $0 \le i \le L$. Since $\sum_{i=0}^{L}\hat{p}_m((L+2)i+1) = (\sum_{i=0}^{L}|\hat{g}(i)|^2)e^{j2\pi\Delta\hat{f}mT}$

$$\Delta\hat{f} = \frac{1}{2\pi mT}\arg\left\{\sum_{i=0}^{L}\hat{p}_m((L+2)i+1)\right\}. \tag{14}$$

This represents an estimate of Δf for a given m. Its acquisition range is given by

$$|\Delta f T| \le \frac{1}{2m} \tag{15}$$

because $|2\pi\Delta f mT| \le \pi$ in (13). The proposed estimate is an average of the estimates in (14) for all possible values of m and is written as

$$\Delta\hat{f} = \frac{1}{N}\sum_{m=1}^{N}\frac{1}{2\pi mT}\arg\left\{\sum_{i=0}^{L}\hat{p}_m((L+2)i+1)\right\} \tag{16}$$

where the parameter N represents the maximum value of m. This estimates the product of the channel information $(\sum_{i=0}^{L}|g(i)|^2)$ and the frequency offset term $(e^{j2\pi\Delta f mT})$ in an LS sense and will be referred to as the *LS estimator* (LSE). In many practical applications, the maximum value of the frequency offset, say $\Delta f_{\max}$, is known. In such cases, from (15), N is upper bounded by

$$N \le \frac{1}{2\Delta f_{\max}T}. \tag{17}$$

TABLE I
SOME TRAINING SEQUENCES

sequence	d(k)
GSM (K=16)	1, -j, 1, j, 1, -j, -1, -j, -1, j, -1, -j, -1, j, -1, -j
IS-136 (K=14)	$e^{j3\pi/4}, e^{j\pi/2}, e^{j\pi/4}, e^{j\pi}, e^{-j3\pi/4}, e^{j0}, e^{-j3\pi/4}, e^{j0}, e^{-j3\pi/4}, e^{j\pi}, e^{-j\pi/4}, e^{j0}, e^{-j\pi/4}, e^{j0}$
Baker code (K=11)	1, -1, 1, 1, -1, 1, 1, 1, -1, -1, -1

For a given N, the lower bound of the training period K is derived from (9)

$$K \ge N + L + (L+1)^2. \tag{18}$$

Therefore, K should be increased in proportion to L^2.

Due to the fact that $\hat{\mathbf{p}}_m$ in (8) only requires $\{d(k)\}$ and $\{r(k)\}$, the LSE can estimate Δf without knowledge of the channel parameters. This is designed for a fixed ISI channel and should be suitable for slow or moderate fading channels. Its implementation is also reasonably simple, because $(\mathbf{D}_m^H\mathbf{D}_m)^{-1}\mathbf{D}_m^H$ in (8) can be *precalculated* for a given training sequence. The accuracy of the estimate can be enhanced by increasing either N or the training period K. Of course, it is not always possible to increase these values as desired because an increase in N causes a reduction of the acquisition range, whereas an increase in K tends to increase the transmission overhead.

As mentioned earlier, the proposed estimator requires a training sequence which guarantees the nonsingularity of $\mathbf{D}_m^H\mathbf{D}_m$. Next some training sequences are presented and the singularities of the corresponding $\mathbf{D}_m^H\mathbf{D}_m$ examined.

Example 2: Table I shows the midamble of the GSM system [16], the preamble of the IS-136 system [17], and a Barker code of length 11. The singularity of the corresponding $\mathbf{D}_m^H\mathbf{D}_m$ is examined under the assumption that $L = 1$ and the results are summarized in Table II. In the case of the IS-136 preamble, $\mathbf{D}_m^H\mathbf{D}_m$ was nonsingular for all possible values of m. However, for the GSM midamble and Barker code, $\mathbf{D}_m^H\mathbf{D}_m$ was singular for certain values of m. ■

IV. PERFORMANCE ANALYSIS

The first part of this section derives approximate expressions for the mean and mean square error (MSE) of the proposed LSE in (16). Then the analytical results are confirmed through computer simulation.

A. Derivation of Mean and MSE

The mean and MSE of the LSE are derived under the assumption of a high signal to noise ratio (SNR). Following the procedure outlined in the Appendix, (16) is approximated as

$$\Delta\hat{f} \approx \Delta f + \frac{1}{N}\sum_{m=1}^{N}\frac{1}{2\pi mT} \cdot \mathrm{Im}\left[e^{-j2\pi\Delta f mT}\sum_{i=0}^{L}\alpha_m((L+2)i+1)\right] \tag{19}$$

TABLE II
SINGULARITY OF $\mathbf{D}_m^H\mathbf{D}_m$ FOR SEQUENCES IN TABLE I

sequence	singularity of $\mathbf{D}_m^H\mathbf{D}_m$ (S: singular, NS: nonsingular)										
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$	$m=10$	$m=11$
GSM	NS	NS	NS	NS	NS	S	NS	NS	NS	S	NS
IS-136	NS	NS	NS	NS	NS	NS	NS	NS	NS	·	·
Barker	NS	NS	S	NS	S	S	·	·	·	·	·

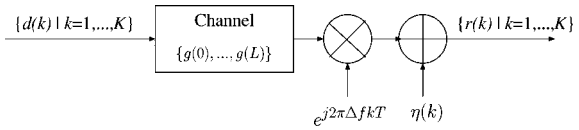

Fig. 2. System model used for simulation.

where $\mathrm{Im}[\cdot]$ represents the imaginary part and $\alpha_m(i)$ is the ith element of $\mathbf{D}_m^+\mathbf{n}_m$ where $\mathbf{D}_m^+ = (\mathbf{D}_m^H\mathbf{D}_m)^{-1}\mathbf{D}_m^H$. After some calculation, it can be shown that

$$\sum_{i=0}^{L}\alpha_m((L+2)i+1) = \sum_{k=1}^{K-m-L}\sum_{i=0}^{L}D_m^+((L+2)i+1,k)n_m(k+m+L) \quad (20)$$

where $D_m^+(i,j)$ is the (i,j)th entry of $\mathbf{D}_m^+$. Since $\mathrm{E}[n_m] \approx 0$ (see the Appendix), (20) indicates that

$$E\left[\sum_{i=0}^{L}\alpha_m((L+2)i+1)\right] \approx 0. \quad (21)$$

The approximately unbiased nature of $\Delta\hat{f}$ in (19) comes from (21). The MSE of the estimate is also obtained from (19). The MSE can be approximately expressed as

$$\sigma_e^2 = \mathrm{E}[(\Delta\hat{f}-\Delta f)^2] \approx \frac{\sigma_\eta^2}{N^2}\sum_{m=1}^{N}\frac{1}{4\pi^2m^2T^2}\sum_{k=1}^{K-m-L}(q_{k+m+L}+q_{k+L}) \cdot\left|\sum_{i=0}^{L}D_m^+((L+2)i+1,k)\right|^2 \quad (22)$$

where $q_k = |\sum_{l=0}^{L}d(k-l)g(l)|^2$ (see the Appendix for the derivation).

B. Comparison Between Numerical and Simulation Results

Fig. 2 shows the system model used for simulation. The pilot $\{d(k)\}$ was the IS-136 preamble shown in Table I. The channel parameters were assumed to be fixed at $\{g(0), g(1)\} = \{1/\sqrt{5}, 2/\sqrt{5}\}$ ($L = 1$). In simulation, the mean and MSE values were empirically estimated through 5000 trials.

Fig. 3 shows the mean estimation errors that were obtained via simulation. The results indicate that the LSE was unbiased within its acquisition range ($|\Delta fT| \le 1/2N$). Fig. 4 shows the normalized MSE, $(\sigma_e T)^2$, plotted as a function of E_b/N_0 for $\Delta f = 0$ (the MSEs were almost independent of Δf, as long as $|\Delta fT| \le 1/2N$—see Fig. 8). The MSEs decreased as E_b/N_0 increased. A remarkably good agreement between the analysis and the simulation was observed. Also shown is the nomalized Cramér–Rao lower bound (CRLB) derived in [14]. As N increased, the MSEs approached the CRLB, which was almost achieved when $N = 9$.

V. APPLICATION TO FREQUENCY-SELECTIVE FADING CHANNELS

In this section, the proposed LSE is compared with the estimators in [13] and [14] when they are applied to a Rayleigh fading channel. To facilitate the comparison, many of the simulation parameters were equal to those of [14]. Specifically, the pulse shaping filter, pilot sequence, and model for the transmission medium were identical. A binary-PSK format was assumed and a raised-cosine filter with a rolloff of 0.5 used for pulse shaping. The pilot was 5230F641 in a hexadecimal form (K = 32). It was also assumed that the frequency offset was limited to the following range:[2] $|\Delta fT| \le 0.05$. The channel response $g_k(l)$ in (1) was obtained by extending the time-invariant channel in [14]. Modeling the transmission medium as the typical urban channel of the GSM system [16] with six paths, $g_k(l)$ is given by

$$g_k(l) = \sum_{i=0}^{5}\xi_i(k)h(lT-\tau_i-t_0), \quad -\infty < l < \infty \quad (23)$$

where $\{\xi_i(k)\}$ and $\{\tau_i\}$ are the attenuations and delays of the paths, respectively, and t_0 is the timing phase, which was selected as equal to $\lfloor L/2\rfloor T$—thereby guaranteeing that $\{g_k(l)|0 \le l \le L\}$ encompassed the $L+1$ most significant channel elements. The normalized delays $\{\tau_i/T\}$ were set at $\{0, 0.054, 0.135, 0.432, 0.621, 1.351\}$. For a given i, $\{\xi_i(k)|-\infty < k < \infty\}$ was a zero-mean complex Gaussian random process where the power spectral density (PSD) was bandlimited to a range $\pm f_D$, where f_D was the maximum Doppler shift. For different paths, $\{\xi_i(k)|0 \le i \le 5\}$ were statistically independent and their variances were equal to $\{-3, 0, -2, -6, -8, -10\}$ (in decibels). $\{\xi_i(k)\}$ for the ith path were generated by passing a complex Gaussian white noise through a baseband Doppler filter [18], as shown in Fig. 5. The

[2]The GSM system calls for oscillator stabilities better than 0.1 ppm. If $|\Delta fT| = 0.05$ is allowed, then the oscillator stability for the GSM can be relaxed to 15 ppm.

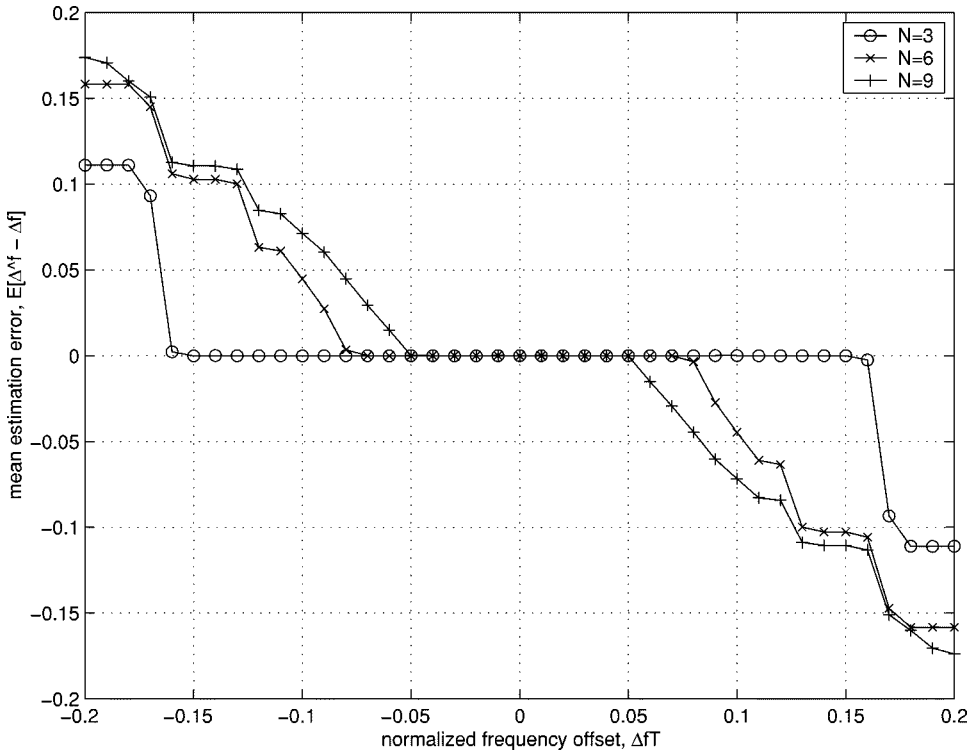

Fig. 3. Mean estimation error (bias) of proposed estimator for fixed ISI channel.

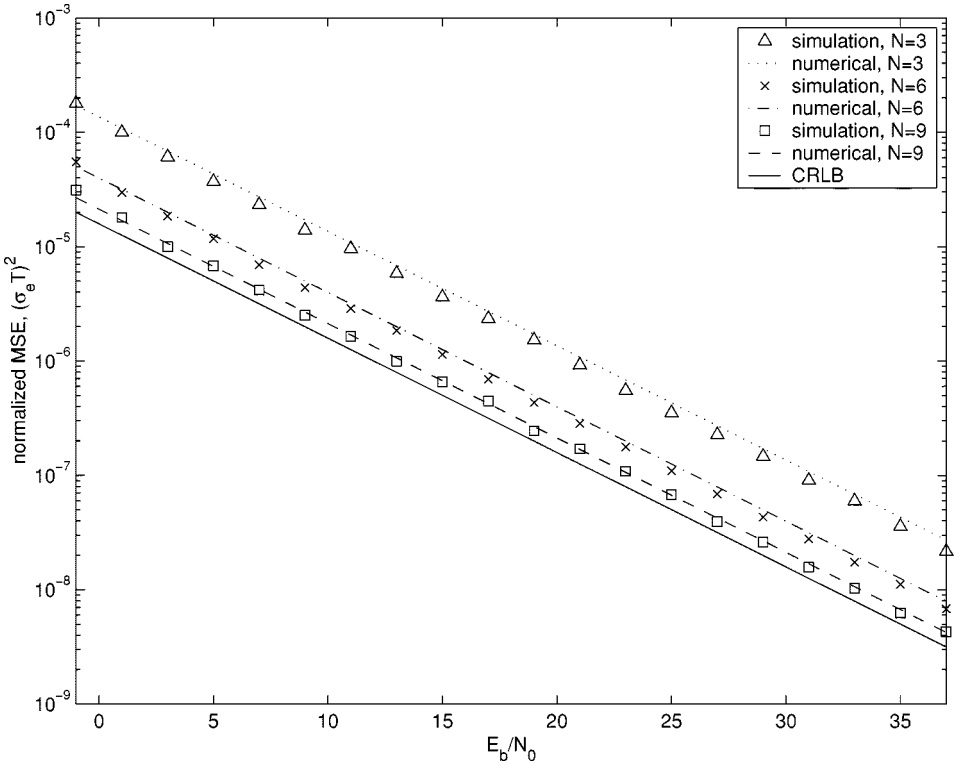

Fig. 4. Normalized MSE of proposed estimator for fixed ISI channel.

Doppler filter was a third-order infinite impulse response (IIR) filter as introduced in [19]. Its system function was

$$H_D(z) = \frac{1 + 3z^{-1} + 3z^{-2} + z^{-3}}{a_0 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}} \tag{24}$$

where $a_0 = \alpha^3 + 0.8\alpha^2 + 1.07\alpha + 0.7$, $a_1 = -3\alpha^3 - 0.8\alpha^2 + 1.07\alpha + 2.1$, $a_2 = 3\alpha^3 - 0.8\alpha^2 - 1.07\alpha + 2.1$, $a_3 = -\alpha^3 + 0.8\alpha^2 - 1.07\alpha + 0.7$ and $\alpha = \pi f_D T$. The magnitude response of this filter, shown in Fig. 6, closely approximated the mobile radio spectrum in [20].

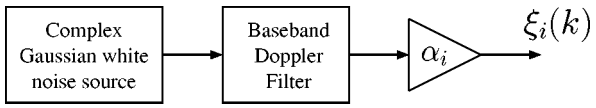

Fig. 5. Generation of $\{\xi_i(k)|1 \le k \le K\}$ for ith path where α_i is positive constant.

After obtaining the time-varying channel $\{g_k(l)\}$, the estimators were implemented as follows.

The Proposed LSE: Equation (16) was evaluated. From (17), $N \le 10$. Since $K = 32$ and $N \le 10$, (18) indicates that $L \le 3$. Therefore, the proposed estimator should assume a channel with a duration of less than or equal to 4.

The Estimator in [13]: This is defined as

$$\Delta\hat{f} = \frac{1}{(K/2+1)\pi T} \arg\left\{\sum_{m=1}^{K/2} \psi(m)\right\} \tag{25}$$

where

$$\psi(m) = \frac{1}{K-m-L} \sum_{k=m+L+1}^{K} r(k)r^*(k-m) \cdot \sum_{i=0}^{L}\sum_{l=0}^{L} \hat{\mathrm{E}}[g_k^*(i)g_{k-m}(l)]d^*(k-i)d(k-m-l) \tag{26}$$

and $\hat{\mathrm{E}}[g_k^*(i)g_{k-m}(l)]$, an estimate of the channel autocorrelation, were obtained by averaging $g_k^*(i)g_{k-m}(l)$ over one million samples, i.e., $\hat{\mathrm{E}}[g_k^*(i)g_{k-m}(l)] = (1/10^6)\sum_{k=m+1}^{10^6} g_k^*(i)g_{k-m}(l)$. The estimator in (25) will be referred to as the *autocorrelation-based estimator* (ABE).

The Estimator in [14]: This estimator is expressed by

$$\Delta\hat{f} = \max_{\Delta\tilde{f}} \left\{-\rho(0) + 2\mathrm{Re}\left[\sum_{m=0}^{\beta K-1} \rho(m)e^{-j2\pi m\Delta\tilde{f}T}\right]\right\} \tag{27}$$

where $\mathrm{Re}\{\cdot\}$ denotes the real part and $\rho(m)$ is a weighted correlation of the received samples, defined as

$$\rho(m) = \begin{cases} \displaystyle\sum_{k=m+1}^{K} B(k-m, k)r(k)r^*(k-m), \\ \qquad m \le K-1 \\ 0, \quad \text{otherwise.} \end{cases} \tag{28}$$

In (28), $B(i, j)$ is the (i, j)th entry of a K-by-K matrix $\mathbf{B}$ which is given by $\mathbf{B} = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$. Here $\mathbf{A}$ is a K-by-L matrix whose (i, j)th entry is defined as $d(i-j)$. Following [14], the parameter β in (27) was fixed at 8. An exhaustive search is needed to obtain $\Delta\hat{f}$ in (27). The estimator in (27) is called the *ML estimator* (MLE).

For the ABE and MLE, L is upper bounded by $K-1$. L was set at 7 for these algorithms, because $\{g_k(l)|0 \le l \le 7\}$ encompassed all significant channel elements, as observed in [14] for the corresponding time-invariant channel.

To examine the sensitivity of the proposed LS algorithm to the parameter L, normalized MSEs of the LSE were evaluated for $L \in \{0, 1, 2, 3\}$ (N was fixed at 10). The results, shown in Fig. 7, indicate that the performance of the estimator was insensitive to L when $1 \le L \le 3$. Therefore, L was set at 1 for the LSE considered in the remaining investigations.[3]

Fig. 8 shows the MSEs of the estimators when the normalized frequency offset ΔfT varied from 0 to 0.05. The MSE of the proposed estimator with $N = 10$ was in between those of the MLE and ABE when $f_DT = 10^{-3}$, yet the LSE outperformed the other estimators when $f_DT = 10^{-2}$. Although the MLE exhibited excellent behavior for slowly varying channels, its performance degraded rather rapidly as the Doppler frequency increased. This is confirmed in Fig. 9, which compares the MSEs for various values of f_DT (ΔfT was assumed to be zero). The LSE with $N = 10$ started to perform better than the other estimators when $f_DT = 2\times10^{-3}$, and the performance gap became wider as f_DT increased. These results indicate that the LSE is more robust to channel variations and is a useful alternative to existing techniques in mobile communication applications.

Finally, this section compares the computational complexities of the estimators. Under the assumption that $(\mathbf{D}_m^H\mathbf{D}_m)^{-1}\mathbf{D}_m^H$ in (8) has been precalculated, the number of multiplications and additions required by the LSE was counted and expressed in terms of K, N, and L. These results and the corresponding expressions of the ABE and MLE, as presented in [14], are listed in Table III. The coefficient μ in this table was given by $\mu = 1 - (\log_2\beta + 2(1/\beta - 1))/(\log_2(\beta K))$. Fig. 10 illustrates the number of total operations (multiplications plus additions) as a function of the training period. The curves were computed from the results in Table III with $N = K/2$ and $L \in \{1, 3\}$. The computational complexity of the LSE was in between those of the ABE and MLE.

VI. Application to Transmissions With Antenna Diversity

A. System Model and Frequency Estimation

The system model considered in this section, as shown in Fig. 11, has $\Gamma + 1$ transmitter antennas and one receiver antenna. A space–time encoder is employed to generate each sequence $\{s_i(k)\}$ which is transmitted through the ith antenna. It is assumed that each $\{s_i(k)\}$ experiences flat fading; that is, the channel between the ith transmitter antenna and the receiver antenna is a flat fading channel.[4] The received signal $r(k)$ is expressed as

$$r(k) = e^{j(2\pi\Delta fkT+\theta)} \sum_{i=0}^{\Gamma} s_i(k)g_{i,k} + \eta(k) \tag{29}$$

where $g_{i,k}$ denotes the channel response from the ith transmitter antenna to the receiver antenna. In this model, we assume that the signals from the transmitter antennas arrive at the receiver antenna with the same delays. This assumption is justified by the fact that in the cellular bands, the propagation delays among the transmitter antenna elements, are measured in nanoseconds

[3]The LSE with $L = 1$ works in a mismatched mode where the *assumed* channel length which is equal to 2 is much shorter than the *true* one. It was observed that such a mismatch caused some floor in the curves of normalized MSE, $(\sigma_eT)^2$, versus $\mathrm{E}_b/\mathrm{N}_0$.

[4]The extension of the estimator proposed in this section to a frequency-selective channel is possible, yet cumbersome.

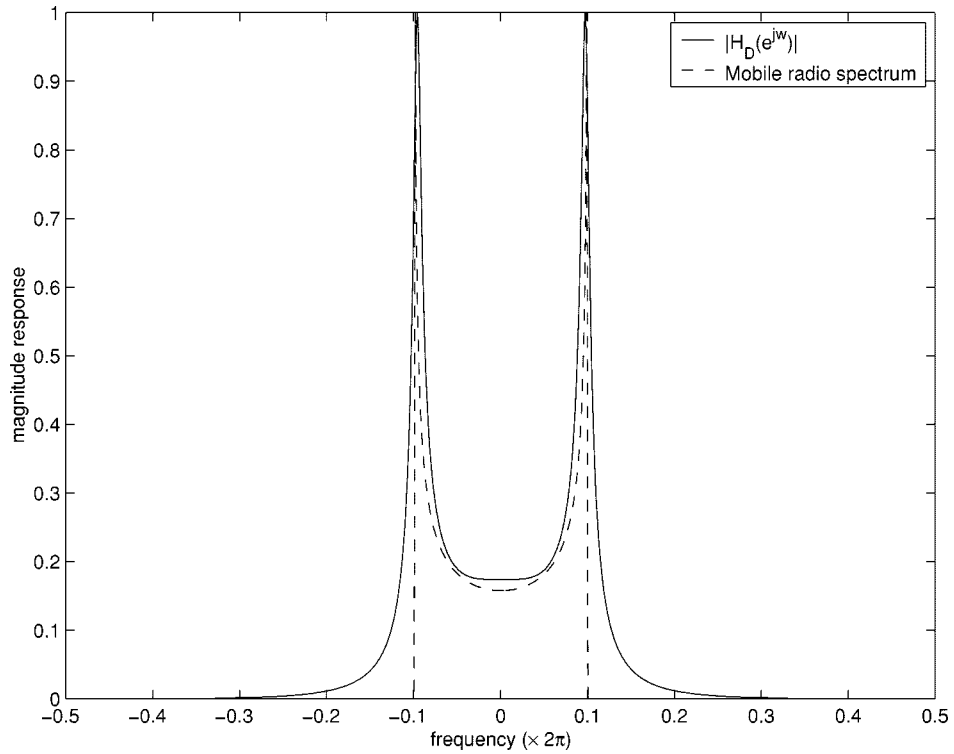

Fig. 6. Magnitude response of Doppler filter in (24) when $f_D T = 0.1$.

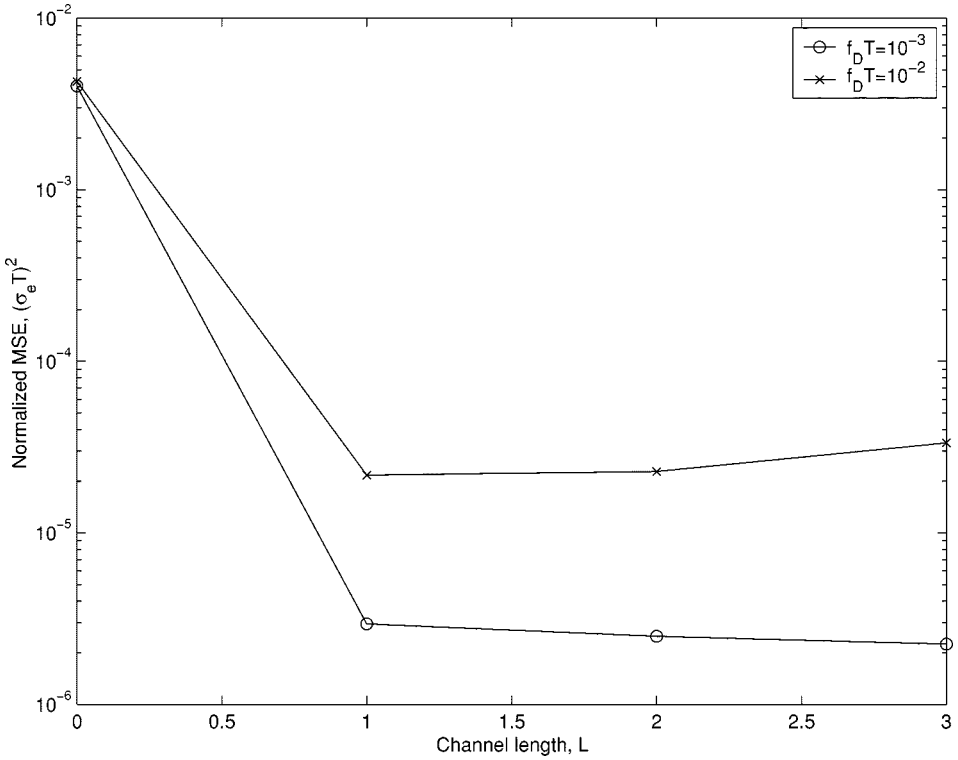

Fig. 7. Sensitivity of the proposed estimator to L ($\mathrm{E}_b/\mathrm{N}_0 = 15$ dB).

[23], while the symbol (or chip in CDMA) durations are measured in microseconds.

When comparing (29) with (1), the right-hand side of (29) is essentially the same as that of (1). The multipath interference observed in (29) is caused by multi-antenna transmissions. Thus, due to the similarity between (1) and (29), the LSE can be directly applied to the estimation of Δf in (29). The estimator for this case is written as

$$\Delta\hat{f} = \frac{1}{N}\sum_{m=1}^{N}\frac{1}{2\pi mT}\arg\left\{\sum_{i=0}^{\Gamma}\hat{p}_m((\Gamma+2)i+1)\right\} \quad (30)$$

where $\hat{\mathbf{p}}_m$ is given by (8), yet $\mathbf{d}_m(k)$ in (3) and $\mathbf{g}$ in (4) should

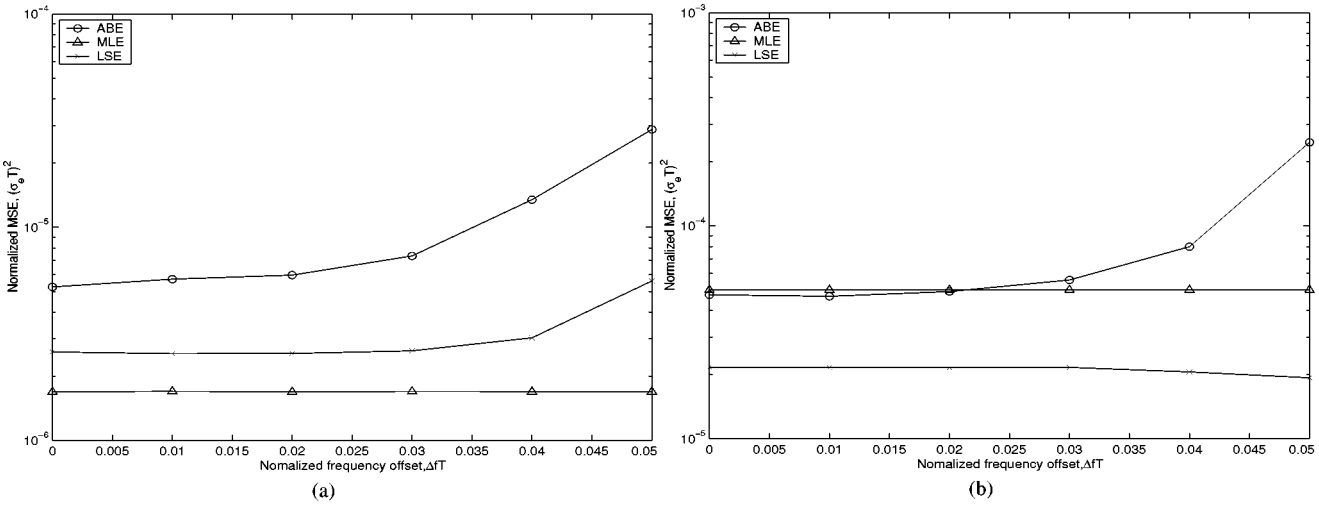

Fig. 8. Performance comparison for frequency-selective fading channel ($E_b/N_0 = 15$ dB). (a) $f_D T = 10^{-3}$. (b) $f_D T = 10^{-2}$.

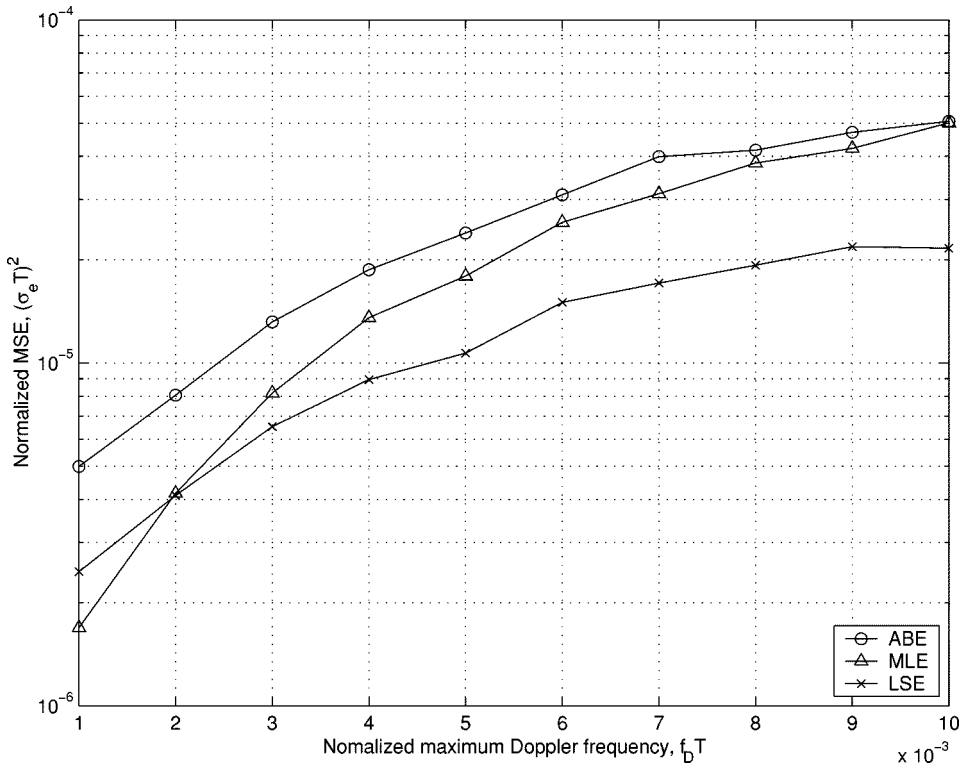

Fig. 9. Performance comparison when the Doppler frequency varies from 10^{-3} to 10^{-2} ($\Delta f T = 0$ and $E_b/N_0 = 15$ dB).

be modified as

$$\mathbf{d}_m(k) \triangleq [s_0(k)s_0^*(k-m),\, s_0(k)s_1^*(k-m),\, \ldots,\, s_1(k)s_0^*(k-m),\, s_1(k)s_1^*(k-m),\, \ldots,\, s_\Gamma(k)s_\Gamma^*(k-m)]^T \tag{31}$$

$$\mathbf{g} \triangleq [|g_0|^2,\, g_0 g_1^*,\, \ldots,\, g_1 g_0^*,\, |g_1|^2,\, g_1 g_2^*,\, \ldots,\, |g_\Gamma|^2]. \tag{32}$$

B. Application to 3GPP Systems

Fig. 12 illustrates the transmitter and receiver for the common pilot channel (CPICH) of the 3GPP system employing transmitter diversity. There are two transmission antennas ($\Gamma = 1$). The CPICH symbol generator produces the pilot sequences, $s_0(k)$ and $s_1(k)$, shown in Fig. 13. These sequences are multiplied with the same channelization and scrambling codes. Thereafter, they are transmitted through multipath fading channels. The received signal is input to a RAKE receiver with $W + 1$ fingers, which separately detect the $W + 1$ strongest multipath components. At each finger, the input signal with a chip rate of $1/T_c$ is descrambled and then correlated with the locally generated channelization code. The output of the wth finger, $r_w(k)$, is obtained by decimating the resulting signal. Owing to the autocorrelation characteristic of the codes, the multipath interference at each finger can be ignored, and $r_w(k)$

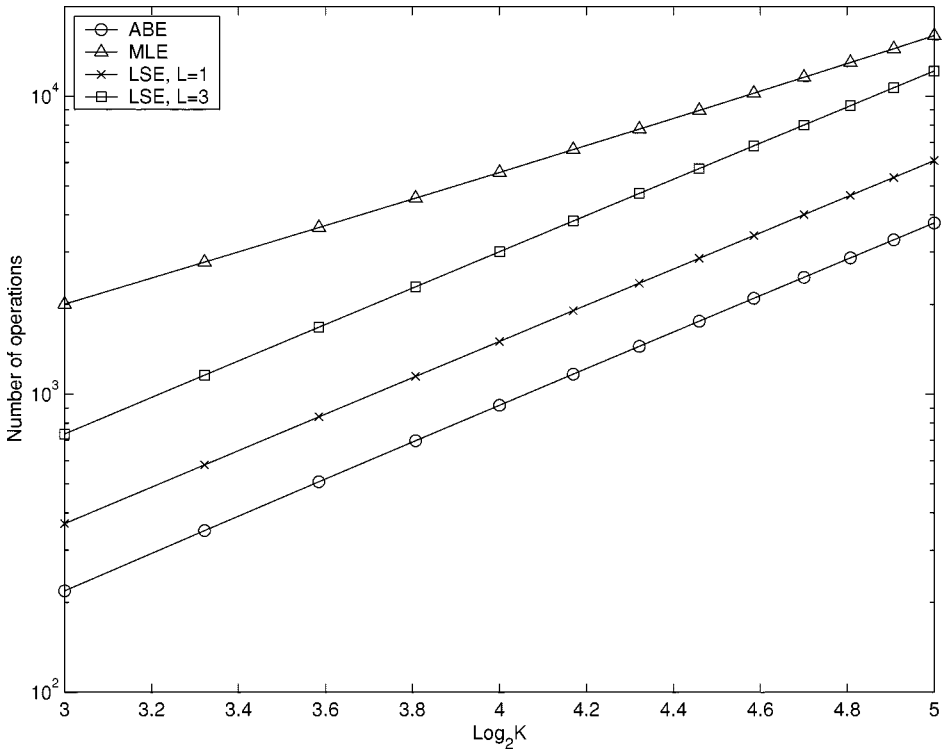

Fig. 10. Complexity comparison.

TABLE III
COMPUTATIONAL LOAD

algorithm	real products	real additions
ABE	$3N(2K-N-1)+1$	$2N(2K-N-1)-2$
MLE	$2K[2K+2+\beta\mu\log_2(\beta K)]$	$K[3K+1+3\beta\mu\log_2(\beta K)]$
LSE	$1+N+4(L+1)\sum_{m=1}^{N}(K-m)$	$N-1+2LN+4(L+1)\sum_{m=1}^{N}(K-m)$

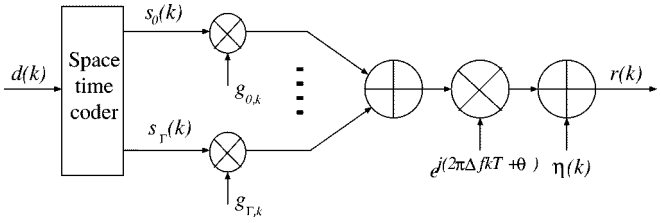

Fig. 11. Baseband system model with transmitter diversity.

is expressed as

$$r_w(k) = e^{j(2\pi\Delta f kT+\theta)} \sum_{i=0}^{1} s_i(k) g_{i,k}(w) + \eta(k) \qquad (33)$$

where $g_{i,k}(w)$ is the channel response associated with the wth finger and the ith transmitter antenna. Since (33) can be viewed as a special case of (29), it is possible to apply the estimator in (30) to the estimation of Δf in (33).

The performance of the proposed estimator was examined using computer simulation. It was easily observed that the ABE and MLE could be applied to the signals in (29) and (33); therefore, they were also applied to the 3GPP system and compared with the proposed LSE. In addition, to examine the effect of transmitter diversity on frequency estimation, the estimator in [11], as proposed for flat fading channels, was applied to the case of single antenna transmission ($\Gamma = 0$) and compared with the estimators associated with two transmitter antenna elements ($\Gamma = 1$). The simulation was performed following the 3GPP specifications: the CPICH symbol rate was 15 ksymbol/s and the carrier frequency was 2 GHz. The observation interval K was 20. Fig. 14 shows the simulation results. First, the performances of the LSE, ABE, and MLE were compared. As in the case of Fig. 8, the MSE of the proposed LSE was in between those of the MLE and ABE when the vehicle speed was 3 km/h ($f_DT = 3.7 \times 10^{-4}$), yet the LSE outperformed the other estimators when the vehicle speed was 120 km/h ($f_DT = 1.5\times10^{-2}$). The MLE performed the best for $f_DT = 3.7 \times 10^{-4}$, yet exhibited the worst performance for $f_DT = 1.5\times10^{-2}$. When comparing the estimator in [11] with the other estimators, the diversity gain was significant when the channel fade rate was low. However, in a fast fading environment, the former performed better than the MLE and was almost comparable to the ABE. In this case, only the LSE achieved some diversity gain. These results indicate that the proposed method is preferable to the other methods in 3GPP mobile communication applications.

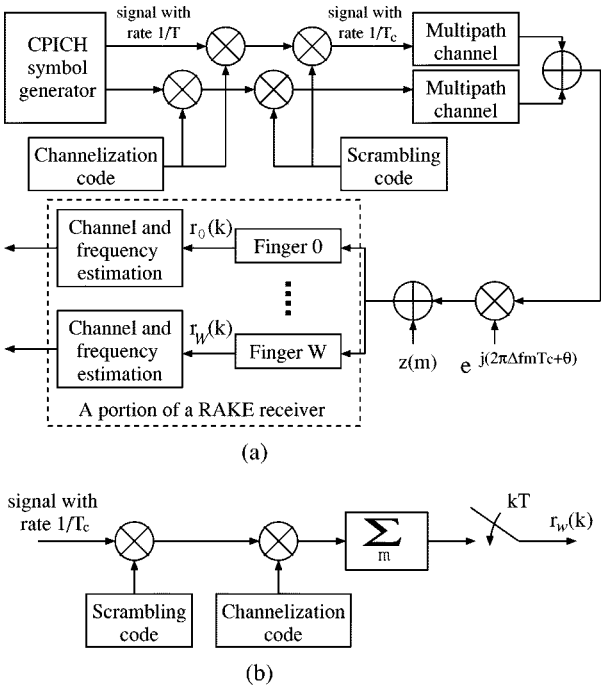

Fig. 12. (a) System model for CPICH of 3GPP where T_c is chip period. (b) One finger of RAKE receiver in (a).

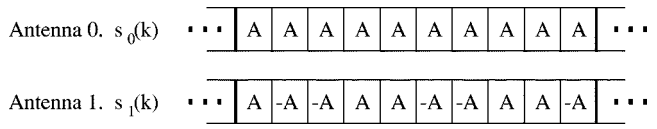

Fig. 13. Symbol pattern for common pilot channel of 3GPP ($A = 1 + j$).

In the 3GPP system, the observation interval K can be increased as desired. Therefore, to reduce the MSE of the proposed method, it is recommended to increase K instead of increasing N. As a result, a desirable value of MSE can be achieved without sacrificing the acquisition range.

VII. Conclusion

An LS frequency offset estimator for frequency-selective fading channels was derived. The proposed estimator does not require any channel information and is approximately unbiased. Its MSE tends to approach the CRLB. Simulation results showed that the proposed estimator is preferable to existing techniques in mobile communications. It was also demonstrated that the proposed LSE is useful for estimating the frequency offsets of the 3GPP system employing transmitter diversity. Expanding the LSE to consider Doppler-induced dispersion remains as a future research topic.

Appendix

A. Derivation of (19)

Using (6), $\hat{\mathbf{p}}_m$ in (8) can be rewritten as

$$\hat{\mathbf{p}}_m = \mathbf{p}_m + \mathbf{D}_m^{+}\mathbf{n}_m \tag{A1}$$

where $\mathbf{D}_m^{+} = (\mathbf{D}_m^H\mathbf{D}_m)^{-1}\mathbf{D}_m^H$.

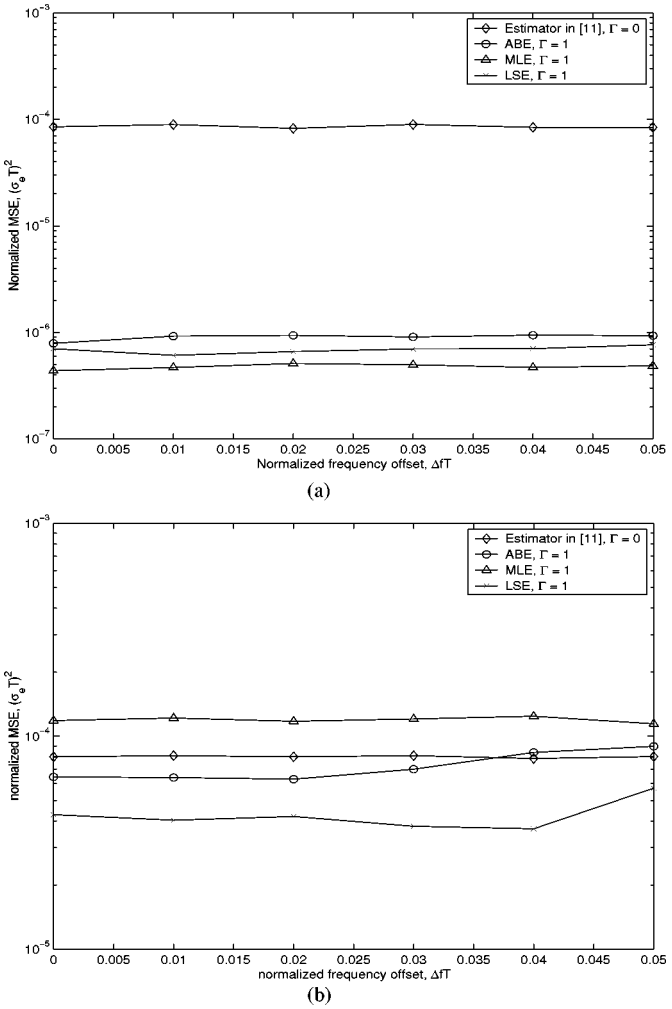

Fig. 14. Performance comparison for transmissions with antenna diversity. (a) Vehicle speed = 3 km/h ($f_D T = 3.7\times10^{-4}$) and $E_b/N_0 = 10$ dB. (b) Vehicle speed = 120 km/h ($f_D T = 1.5 \times 10^{-2}$) and $E_b/N_0 = 10$ dB.

From (2), the kth element of $\mathbf{n}_m$, $n_m(k+m+L)$ is approximated as

$$\begin{aligned} n_m(k+m+L) &\approx \eta^*(k+L)e^{j(2\pi\Delta f(k+m+L)T+\theta)}\sum_{l=0}^{L} d(k+m+L-l)g(l) \\ &\quad +\eta(k+m+L)e^{-j(2\pi\Delta f(k+L)T+\theta)}\sum_{l=0}^{L} d^*(k+L-l)g^*(l) \end{aligned} \tag{A2}$$

where $\eta(k+m+L)\eta^*(k+L)$ is neglected under the assumption of high SNR. Since $\eta(k)$ is white Gaussian noise with zero-mean, $\mathrm{E}[n_m(k+m+L)] \approx 0$ and $\mathrm{E}[\mathbf{n}_m] \approx \mathbf{0}$. Using (A1) in (16), the value inside the brace in (16) becomes

$$\begin{aligned} &\sum_{i=0}^{L}\hat{p}_m((L+2)i+1) \\ &\qquad = \sum_{i=0}^{L}|g(i)|^2 e^{j2\pi\Delta fmT} + \sum_{i=0}^{L}\alpha_m((L+2)i+1) \end{aligned} \tag{A3}$$

where $\alpha_m(i)$ is the ith element of $\mathbf{D}_m^+\mathbf{n}_m$. Equation (A3) can be approximated by

$$\sum_{i=0}^{L}\hat{p}_m((L+2)i+1) \approx \exp\left(j\left(2\pi\Delta f mT + \mathrm{Im}\left[e^{-j2\pi\Delta f mT}\sum_{i=0}^{L}\alpha_m((L+2)i+1)\right]\right)\right) \tag{A4}$$

where, without loss of generality, it is assumed that $\sum_{i=0}^{L}|g(i)|^2 = 1$ and $\mathrm{Im}[\cdot]$ represents the imaginary part [4]. Using (A4) in (16) we get (19).

B. Derivation of MSE

By squaring $\Delta\hat{f}-\Delta f$ in (19) and taking the expectation, we get

$$\begin{aligned}
&\mathrm{E}[(\Delta\hat{f}-\Delta f)^2] \\
&\approx \frac{1}{N^2}\mathrm{E}\left[\left\{\sum_{m=1}^{N}\frac{1}{2\pi mT}\cdot\mathrm{Im}\left[e^{-j2\pi\Delta f mT}\sum_{i=0}^{L}\alpha_m((L+2)i+1)\right]\right\}\right. \\
&\qquad\left.\cdot\left\{\sum_{m'=1}^{N}\frac{1}{2\pi m'T}\cdot\mathrm{Im}\left[e^{-j2\pi\Delta f mT}\sum_{i'=0}^{L}\alpha_{m'}((L+2)i'+1)\right]\right\}\right] \\
&= \frac{1}{N^2}\mathrm{E}\left[\left\{\sum_{m=1}^{N}\frac{1}{2\pi mT}\cdot\mathrm{Im}\left[e^{-j2\pi\Delta f mT}\sum_{k=1}^{K-m-L}\sum_{i=0}^{L}\cdot D_m^+((L+2)i+1,\,k)n_m(k)\right]\right\}\right. \\
&\qquad\left.\cdot\left\{\sum_{m'=1}^{N}\frac{1}{2\pi m'T}\cdot\mathrm{Im}\left[e^{-j2\pi\Delta f m'T}\sum_{k'=1}^{K-m'-L}\sum_{i'=0}^{L}\cdot D_m^+((L+2)i'+1,\,k')n_{m'}(k')\right]\right\}\right].
\end{aligned} \tag{A5}$$

If we only keep those terms that correspond to $m = m'$ and $k = k'$ and ignore the rest, then

$$\mathrm{E}[(\Delta\hat{f}-\Delta f)^2] \approx \frac{1}{N^2}\sum_{m=1}^{N}\frac{1}{4\pi^2m^2T^2}\sum_{k=1}^{K-m-L}\cdot\left|\sum_{i=0}^{L}D_m^+((L+2)i+1,\,k)\right|^2\mathrm{E}[|n_m(k)|^2]. \tag{A6}$$

Using (A2) in (A6), the expression in (22) is obtained.

REFERENCES

[1] 3G TS 25.211 (V3.5.0). (2000, Dec.) Physical channels and mapping of transport channels onto physical channels (FDD). [Online]. Available: http://www.3gpp.org/ftp/Specs/2000-12/R1999/25_series/25 211-350.zip

[2] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York: Plenum, 1997.

[3] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers*. New York: Wiley, 1998.

[4] S. A. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 832–835, Nov. 1985.

[5] S. Kay, "A fast and accurate single frequency estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1987–1990, Dec. 1989.

[6] M. P. Fitz and W. C. Lindsey, "Decision-directed burst-mode carrier synchronization techniques," *IEEE Trans. Commun.*, vol. 40, pp. 1644–1653, Oct. 1992.

[7] M. P. Fitz, "Planar filtered techniques for burst mode carrier synchronization," in *Proc. IEEE Global Telecommun. Conf.*, Orlando, FL, 1991, pp. 365–369.

[8] ——, "Further results in the fast estimation of a single frequency," *IEEE Trans. Commun.*, vol. 42, pp. 862–864, Feb./Mar./Apr. 1994.

[9] M. Luise and R. Reggiannini, "Carrier frequency recovery in all-digital modems for burst-mode transmissions," *IEEE Trans. Commun.*, vol. 43, pp. 1169–1178, Feb./Mar./Apr. 1995.

[10] U. Mengali and M. Morelli, "Data-aided frequency estimation for burst digital transmission," *IEEE Trans. Commun.*, vol. 45, pp. 23–25, Jan. 1997.

[11] W. Y. Kuo and M. P. Fitz, "Frequency offset compensation of pilot symbol assisted modulation in frequency flat fading," *IEEE Trans. Commun.*, vol. 45, pp. 1412–1416, Nov. 1997.

[12] M. Morelli, U. Mengali, and G. M. Vitetta, "Further results in carrier frequency estimation for transmissions over flat fading channels," *IEEE Commun. Letters*, vol. 2, pp. 327–330, Dec. 1998.

[13] M. G. Hebley and D. P. Taylor, "The effect of diversity on a burst-mode carrier-frequency estimator in the frequency-selective multipath channel," *IEEE Trans. Commun.*, vol. 46, pp. 553–560, Apr. 1998.

[14] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmissions over selective channels," *IEEE Trans. Commun.*, vol. 48, pp. 1580–1589, Sept. 2000.

[15] L. L. Scharf, *Statistical Signal Processing*. Natick, MA: Addison-Wesley, 1991.

[16] GSM Recommendation 05.05, "Radio Transmission and Reception,", V.8.7.1, ETSI/PT, 1999.

[17] Electronic Industries Association, *EIA/TIA IS-136: Dual-Mode Mobile Station-Base Station Compatibility Standard*. Washington D.C.: TIA, 1994.

[18] W. P. Chou and P. J. McLane, "16-state nonlinear equalizer for IS-54 digital cellular channels," *IEEE Trans. Veh. Technol.*, pp. 12–25, Feb. 1996.

[19] G. S. Liu and C. H. Wei, "Timing recovery techniques for digital cellular radio with $\pi/4$-DQPSK modulation," in *Proc. IEEE Int. Conf. Communications*, Chicago, IL, 1992, pp. 319–323.

[20] W. C. Jakes, *Microwave Mobile Communication*. New York: Wiley, 1974.

[21] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of FDD transmission systems using transmitter antenna diversity," in *Proc. IEEE Vehicular Technology Conf.*, Secaucus, NJ, May 1993, pp. 508–511.

[22] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[23] T. L. Marzetta, B. Hochwald, and C. B. Papadias, "A transmitter diversity scheme for wideband CDMA systems based on space–time spreading," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 48–60, Jan. 2001.

Eui-Rim Jeong received the B.S., M.S., and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Taejon, Korea, in 1995, 1997, and 2001, respectively.

Since 2001, he has been with Hyundai Syscomm Co. in Korea. His research interests include synchronization, modem structure of wireless digital communication, and equalization.

Sung-Kwon Jo received the B.S. degree in electrical engineering from Yonsei University, Seoul, Korea, in 1991, and the M.S. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Taejon, Korea, in 1993 and 2000, respectively.

Since 1998, he has been with Samsung Electronics Co., Ltd. in Korea, where he engages in research and development of mobile communication system. His research interests include digital signal processing, adaptive filter theory, and digital communication theory with a special emphasis on wireless communications.

Yong H. Lee (SM'99) was born in Seoul, Korea, on July 12, 1955. He received the B.S. and M.S. degrees in electrical engineering from Seoul National University, Seoul, Korea, in 1978 and 1980, respectively, and the Ph.D. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1984.

From 1984 to 1988, he was an Assistant Professor in the Department of Electrical and Computer Engineering, State University of New York at Buffalo. Since 1989, he has been with the Department of Electrical Engineering at KAIST, where he is currently a Professor. His research activities include communication signal processing, which includes synchronization, interference cancellation, and efficient filter design.