Least Squares Approach to Data-Aided Frequency Estimation in Frequency-Selective Fading Channels

Eui-Rim Jeong†, Sung-Kwon Jo††, Young-Doo Kim† and Yong H. Lee† *

† Department of Electrical Engineering, Korea Advanced Institute of Science and Technology
373-1 Kusong-dong, Yusong-gu, Taejon, 305-701 KOREA

†† Samsung Electronics Co., LTD.
263, Seohyeon-dong, Bundang-gu, Sungnam-si, Kyungki-do, Korea

Abstract— **A new data-aided frequency estimator for frequency-selective fading channels is introduced. This estimator is developed based on a least squares (LS) error criterion. The proposed method can estimate frequency offsets without the need for channel information. It is shown through simulation that the proposed LS method can be preferred to the existing techniques in mobile communications.**

I. Introduction

Recent communication systems tend to impose stringent requirement on the frequency stability of the transmitter and receiver oscillators. For example, 3rd 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], [3]. Among these, the class of data-aided techniques [4] – [14], which use a training sequence for frequency offset estimation, become popular because they can attain good performance with a short training sequence. The 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, the knowledge of 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 is either assumed to be known [11] or estimated [12]. The estimator in [13] requires statistics of frequency-selective channels. As a consequence, the performance of the estimators in [11]–[13] is degraded when 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 an excellent behavior for a fixed channel; however, its performance can be considerably degraded when the channel is time-varying.

The objective of this paper is to develop a more efficient carrier frequency recovery method for frequency selective fading channels. The proposed approach is based on a least squares (LS) error criterion that does not need any a prior knowledge about channels. It will be shown that the proposed LS estimator can outperform those in [13] and [14] in time-varying channels.

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 Section IV presents simulation results demonstrating the advantage of the proposed estimator over existing methods.

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 a baseband pulse shape; $n(t)$ is AWGN; θ is an initial random phase, and Δf represents a carrier frequency offset. Assuming perfect symbol timing recovery, the output of the receiver filter sampled at $t = kT$ is

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

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

*All correspondence may be addressed to this author.

0-7803-6728-6/01/$10.00 ©2001 IEEE.

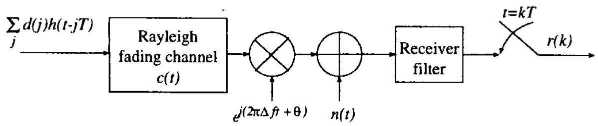

Fig. 1. Baseband system model.

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 $\{d(k)|k=1,...,K\}$ are available, and that $\{g_k(l)\}$ is a fixed ISI channel over the training period, i.e., $g_k(l) = g(l)$ for $k = 1,...,K$. The problem considered in this section is stated as follows: given $\{d(k)|k = 1,...,K\}$ and $\{r(k)|k = 1,...,K\}$, estimate Δf without the knowledge of $\{g(l)|l = 0,...,L\}$. Estimating Δf directly from $\{r(k)\}$ is difficult, because this involves the estimation of the time-varying phase $2\pi\Delta f kT + \theta$ 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

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

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

$$\begin{aligned}\mathbf{d}_m(k) &\triangleq [d(k)d^*(k-m), d(k)d^*(k-m-1), ...,\\ &\quad d(k-1)d^*(k-m), d(k-1)d^*(k-m-1),\\ &\quad ..., d(k-L)d^*(k-m-L)]^T, \qquad (3)\\ \mathbf{g} &\triangleq [|g(0)|^2, g(0)g^*(1), ..., g(1)g^*(0),\\ &\quad |g(1)|^2, g(1)g^*(2), ..., |g(L)|^2]^T. \qquad (4)\end{aligned}$$

Note that $2\pi\Delta f mT$ 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 of k:

$$m + L + 1 \le k \le K. \qquad (5)$$

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

$$\begin{aligned}\boldsymbol{\gamma}_m &= \mathbf{D}_m\mathbf{g}e^{j2\pi\Delta f mT} + \mathbf{n}_m\\ &= \mathbf{D}_m\mathbf{p}_m + \mathbf{n}_m \qquad (6)\end{aligned}$$

where $\boldsymbol{\gamma}_m \triangleq [\gamma_m(m+L+1), \gamma_m(m+L+2), ..., \gamma_m(K)]^T$, $\mathbf{n}_m \triangleq [n_m(m+L+1), n_m(m+L+2), ..., n_m(K)]^T$, $\mathbf{D}_m$ is a $(K-m-L)\times(L+1)^2$ dimensional matrix defined as $\mathbf{D}_m \triangleq [\mathbf{d}_m(m+L+1), \mathbf{d}_m(m+L+2), ..., \mathbf{d}_m(K)]^T$ and $\mathbf{p}_m \triangleq \mathbf{g}e^{j2\pi\Delta f mT}$. 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 m=8, $\boldsymbol{\gamma}_m$ has 5 elements and $\mathbf{D}_m$ is a 5×4 matrix. (6) is written as

$$\begin{bmatrix} r(10)r^*(2)\\ r(11)r^*(3)\\ r(12)r^*(4)\\ r(13)r^*(5)\\ r(14)r^*(6)\end{bmatrix} = \begin{bmatrix} d(10)d^*(2) & d(10)d^*(1) & d(9)d^*(2) & d(9)d^*(1)\\ d(11)d^*(3) & d(11)d^*(2) & d(10)d^*(3) & d(10)d^*(2)\\ d(12)d^*(4) & d(12)d^*(3) & d(11)d^*(4) & d(11)d^*(3)\\ d(13)d^*(5) & d(13)d^*(4) & d(12)d^*(5) & d(12)d^*(4)\\ d(14)d^*(6) & d(14)d^*(5) & d(13)d^*(6) & d(13)d^*(5)\end{bmatrix}\cdot$$

$$\begin{bmatrix} |g(0)|^2\\ g(0)g^*(1)\\ g^*(0)g(1)\\ |g(1)|^2\end{bmatrix} e^{j2\pi\Delta f 8T} + \mathbf{n}_8$$

■

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, and 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 error squares:

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

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

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

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 the determined and overdetermined cases; that is, 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:

$$1 \le m \le K - L - (L+1)^2. \qquad (9)$$

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

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 Δf can be estimated by examining the elements of $\hat{\mathbf{p}}_m$. Let $\hat{\mathbf{p}}_m = [\hat{p}_m(1), \hat{p}_m(2), ..., \hat{p}_m((L+1)^2)]^T$. Since $\mathbf{p}_m = \mathbf{g}e^{j2\pi\Delta f mT}$, we may write that $\hat{\mathbf{p}}_m = \hat{\mathbf{g}}e^{j2\pi\Delta \hat{f} mT}$, 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 1st element, from (4), we get

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

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, we can select elements of $\hat{\mathbf{p}}_m$ from which $\Delta\hat{f}$ can be 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 $\sum_{i=0}^{L}\hat{p}_m((L+2)i+1) = \left(\sum_{i=0}^{L}|\hat{g}(i)|^2\right)e^{j2\pi\Delta\hat{f}mT}$,

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

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 LS estimate is an average of the estimates in (14) for all possible values of m. It is written as

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

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

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

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

The parameter N represents the maximum value of m, and must meet the inequalities in (9) and (15). N is determined by considering the tradeoff between the performance and computational cost.

Due to the fact that $\hat{\mathbf{p}}_m$ in (8) only requires $\{d(k)\}$ and $\{r(k)\}$, the estimator in (16) can estimate Δf without the knowledge of channel parameters. It is designed for a fixed ISI channel and should be suitable for slow or moderate fading channels. Its implementation is reasonably simple, because $\left(\mathbf{D}_m^H\mathbf{D}_m\right)^{-1}\mathbf{D}_m^H$ in (8) can be *precalculated* for a given training sequence. The accuracy of this 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 the increase of N causes the reduction of the acquisition range and the increase of K tends to increase transmission overhead.

As mentioned before, this estimator requires a training sequence which guarantees the nonsingularity of $\mathbf{D}_m^H\mathbf{D}_m$. Next we present some training sequences and examine singularities of the corresponding $\mathbf{D}_m^H\mathbf{D}_m$.

Example 2. Table I shows the midamble of the GSM system [17], the preamble of the IS-136 system [18] and the Barker code of length 11. 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 IS-136 preamble, $\mathbf{D}_m^H\mathbf{D}_m$ is nonsingular for all possible values of m. For the GSM midamble and the Barker code, however, $\mathbf{D}_m^H \mathbf{D}_m$ is singular for some m.

■

IV. Application to Frequency-Selective Fading Channels

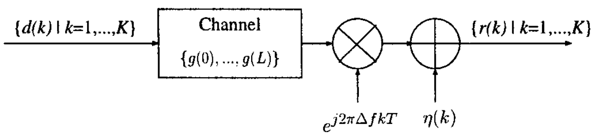

Fig. 2. System model used for simulation.

In this section, the proposed estimator is compared with the estimators in [13] and [14] when they are applied to a Rayleigh fading channel. For the simulation, the system model depicted in Fig. 2 was used. The digital channel $\{g_k(l)\}$ was obtained by sampling (T-spaced) the convolution of an analog channel $c(t)$ and the impulse responses of the transmitter and receiver filters, which were square root raised cosine filters with roll-off factor 0.35 (refer Fig. 1). The analog channel was a two-ray Rayleigh fading channel [16], which was modeled as

$$c(t) = \xi_0(t)\delta(t) + \xi_1(t)\delta(t-\tau) \tag{17}$$

where $\xi_0(t)$ and $\xi_1(t)$ are independent zero mean complex Gaussian processes and τ is the time delay between the two rays. For simplicity, a normalized uniform delay power profile was assumed, i.e., $E[|\xi_0(t)|^2] = E[|\xi_1(t)|^2] = 0.5$. The parameter τ was $T/2$; $L = 1$; the baud rate was 24.3ksymbol/s; and the carrier frequency was 2GHz. The maximum Doppler frequency was: $f_D T = 10^{-3}$ and 10^{-2} which correspond to the vehicle speed 13km/h and 130km/h, respectively. It was assumed that the range of frequency offset is known; $|\Delta f T| \leq 0.05$.

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

The proposed estimator. Equation (16) was evaluated. The parameter N was set to an integer which was less than or equal to 9, because $1/2N \geq 0.05$ from (15).

The estimator in [13]. This is defined as

$$\Delta\hat{f} = \frac{1}{(K/2+1)\pi T} \arg\left\{\sum_{m=1}^{K/2} \psi(m)\right\} \tag{18}$$

where

$$\psi(m) = \frac{1}{K-m-L} \sum_{k=m+L+1}^{K} r(k) r^*(k-m) \cdot$$

$$\sum_{i=0}^{L}\sum_{l=0}^{L} \hat{\mathrm{E}}\left[g_k^*(i) g_{k-m}(l)\right] d^*(k-i) d(k-m-l) \tag{19}$$

and $\hat{\mathrm{E}}[g_k^*(i) g_{k-m}(l)]$, an estimate of the channel autocorrelation, was obtained by averaging $g_k^*(i) g_{k-m}(l)$ over one million samples, i.e., $\hat{\mathrm{E}}[g_k^*(i) g_{k-m}(l)] = \frac{1}{10^6}\sum_{k=m+1}^{10^6} g_k^*(i) g_{k-m}(l)$.

The estimator in [14]. This estimator is expressed by

$$\Delta\hat{f} = \max_{\Delta\tilde{f}} \left\{-\rho(0) + 2\mathrm{Re}\left[\sum_{m=0}^{\beta K-1} \rho(m) e^{-j2\pi m \Delta\tilde{f} T}\right]\right\} \tag{20}$$

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

$$\rho(m) = \begin{cases} \sum_{k=m+1}^{K} B(k-m,k) r(k) r^*(k-m), & m \leq K-1 \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

In (21), $B(i,j)$ 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 (20) was fixed at 8. An exhaustive search is needed to obtain $\Delta\hat{f}$ in (20).

Fig. 3 shows the MSEs of the estimators when the normalized frequency offset $\Delta f T$ varies from 0 to 0.05. As expected, the performance of the LS estimator was enhanced by increasing N. The MSE of the proposed estimator with $N = 9$ was in between those of the estimators in [13] and [14] when $f_D T = 10^{-3}$ and $|\Delta f T| \leq 0.045$, but the LS estimator outperformed the others when $f_D T = 10^{-2}$. Although the estimator in [14] exhibited excellent behavior for slowly-varying channels, its performance was degraded rather rapidly as the Doppler frequency increased. This is confirmed in Fig. 4 which compares the MSEs for various values of $f_D T$ ($\Delta f T$ was assumed to be zero). The proposed LS estimator with $N = 9$ started to perform better than the others when $f_D T = 2 \times 10^{-3}$, and the performance gap became wider as $f_D T$ increased. These results indicate that the LS estimator is more robust

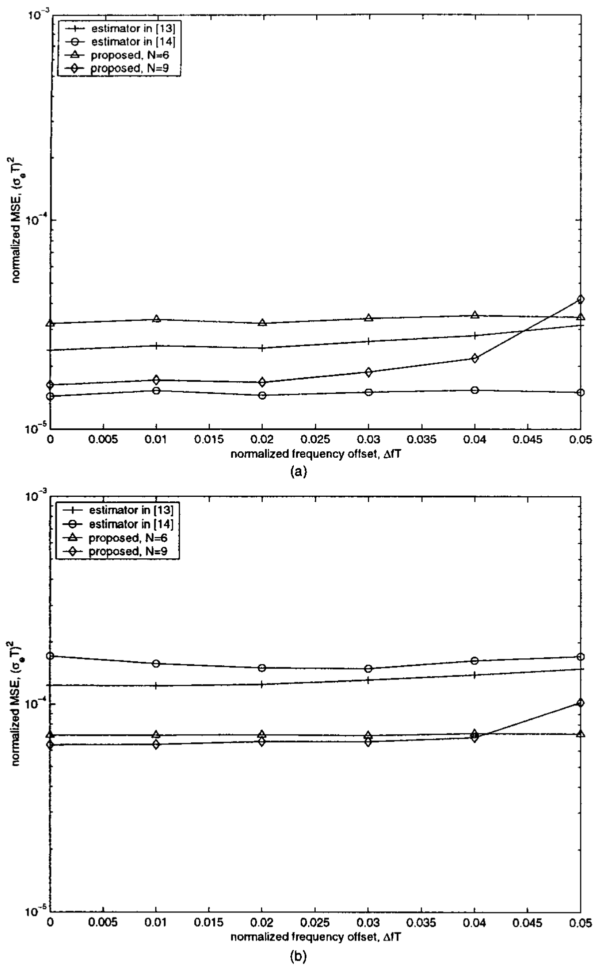

Fig. 3. Performance comparison for a frequency-selective fading channel (E_b/N_0=15dB). (a) $f_D T = 10^{-3}$ (13km/h) (b) $f_D T = 10^{-2}$ (130km/h).

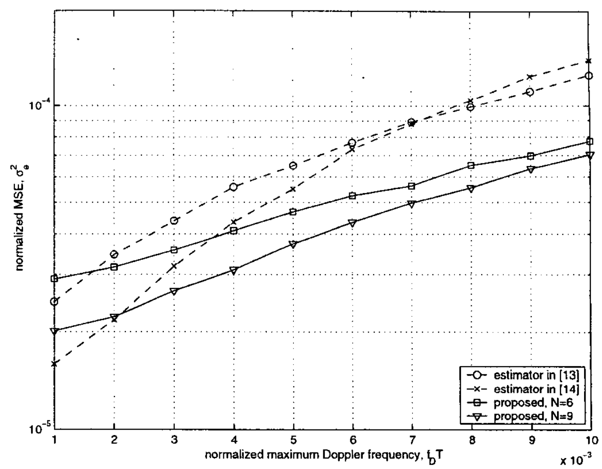

Fig. 4. Performance comparison when the Doppler frequency varies from 10^{-3} to 10^{-2} ($\Delta fT = 0$ and E_b/N_0=15dB).

to channel variations and is a useful alternative to the existing techniques in mobile communication applications.

References

[1] 3G TS 25.201 (V3.0.2), *Physical layer FDD*, Mar. 2000.

[2] U. Mengali and A. N. D'Andrea, *Synchronization techniques for digital receivers*, Plenum Press, New York, 1997.

[3] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital communication receivers*, John Wiley & Sons, New York, 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," *GLOBECOM*, vol. 1, pp. 365–369, Dec 1991.

[8] M. P. Fitz, "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, "Futher 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*, Addison-Wesley Publishing Company, Massachusetts, 1991.

[16] K. Feher, *Wireless digital communications*, Prentice Hall, New Jersey, 1995.

[17] GSM Recommendation 05.05, *Radio Transmission and Reception*, V.3.11.0, ETSI/PT, 1990.

[18] Electronic Industries Association, *EIA/TIA IS-136: dual-mode mobile station-base station compatibility standard*, TIA, Washington D.C., 1994.