Joint ML Estimation of Carrier Frequency, Channel, I/Q Mismatch, and DC Offset in Communication Receivers

Gye-Tae Gil, *Student Member, IEEE*, Il-Hyun Sohn, Jin-Kyu Park, and Yong H. Lee, *Senior Member, IEEE*

***Abstract*—This paper proposes a technique that jointly estimates the I/Q mismatch, dc offset, carrier frequency offset, and channel for receivers in frequency selective channels. The joint estimator is a data-aided technique that requires a training sequence and is derived based on the maximum likelihood (ML) criterion. The characteristics of the joint estimator are analyzed. In particular, it is shown that its accuracy almost achieves the Cramér-Rao lower bound (CRLB). The advantage of the proposed estimator over an existing method is demonstrated through computer simulation.**

***Index Terms*—dc offset, direct conversion, estimation, frequency offset, I/Q mismatch.**

I. Introduction

Frequency offset compensation and channel estimation are important functions of recent radio receivers. By recovering the carrier frequency at a receiver via signal processing, the stringent requirements on the frequency stability of the transmitter and receiver oscillators can be relieved. Estimation of channel parameters is essential for systems with ML-type detection. A popular approach to frequency and channel estimation is to employ data-aided techniques that operate with the help of training sequences. Various *data-aided* frequency and channel estimators can be found in [1]–[8].

Existing frequency and channel estimation techniques assume an ideal receiver in which there are no errors in the radio frequency (RF) and analog circuit, with the exception of a frequency offset. This assumption can be justified for receivers based on the traditional superheterodyne principle. However, its justification can be difficult for certain other receiver architectures, such as those with direct conversion [9], [10], and a low-intermediate frequency (IF) [11], [12], which are accompanied by several deficiencies: for example, both direct conversion and low-IF architectures suffer from an I/Q mismatch that is an imbalance between the I and Q branch amplitudes and phases. Thus, for these receivers, it is recommended that the frequency and channel are estimated in presence of an I/Q mismatch.[1]

In this paper, we develop a data-aided, ML technique that can jointly estimate the frequency, channel, I/Q mismatch and dc offset, where the dc offset is a particular concern in direct conversion receivers. The proposed estimator can be thought of as an extension of the techniques in [4] and [18], where [4] proposes a joint ML scheme for frequency and channel estimation, and [18] proposes a joint least squares (LS) estimate of the I/Q mismatch, dc offset, and channel. It will be observed that the proposed scheme reduces to the estimate in [4] when there is no I/Q mismatch and dc offset, plus it performs better than the estimate in [18] when a frequency offset exists. Like the estimate in [18], the proposed scheme can be applied to both direct-conversion and conventional heterodyne architectures, yet not to those with a low-IF in [11] and [12]. This is because an I/Q mismatch in a low-IF system causes an image signal that comes from an adjacent channel, and the proposed data-aided scheme cannot handle such an image; it can only estimate a self-image.

The remainder of this paper is organized as follows. Section II describes the signal model used in the current study. The proposed method is derived in Section III, and its properties are analyzed in Section IV. In Section V, computer simulation results are presented to demonstrate the performance of the proposed technique in quasistatic channel environments. Section VI contains the conclusions.

II. Signal Model

The signal model considered in the current paper is shown in Fig. 1. Here, $a(t)$ is the complex signal to be transmitted; f_c denotes the carrier frequency; $w_{\mathrm{RF}}(t)$ is zero mean additive white noise in the passband; and d_o is the dc offset. The passband signal entering the mixer can be represented as

$$r_{\mathrm{RF}}(t) = 2\mathrm{Re}\{s(t)e^{j2\pi f_c t}\} + w_{\mathrm{RF}}(t) \tag{1}$$

where $s(t)$ is an equivalent lowpass signal at the receiver antenna, given by the convolution of $a(t)$ and the baseband-equivalent channel impulse response. The receiver shown in Fig. 1 has a direct-conversion architecture, yet it can also be thought

Manuscript received August 5, 2003; revised March 26, 2004 and June 27, 2004. This work was supported in part by the University Information Technology Research Center Program, Government of Korea. This paper was presented in part at the IEEE Vehicular Technology Conference, Jeju, Korea, in April 2003. The review of this paper was coordinated by Dr. K. Molnar.

G.-T. Gil is with the Research Center of Korea Telecom, Daejeon 305-811, Korea (e-mail: gategil@kt.co.kr).

I.-H. Sohn is with the Research and Development Planning Team, SK Telesys, Kyoungkido 441-230, Korea (e-mail: ihsohn@sktelesys.com).

J.-K. Park is with the Agency for Defense Development, Daejeon,Korea (e-mail: jinkpark@hanafos.com).

Y. H. Lee is with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 305-701, Korea (e-mail: yohlee@ee.kaist.ac.kr).

Digital Object Identifier 10.1109/TVT.2004.836919

[1]Signal processing techniques that can compensate for an I/Q mismatch can be found in [13]–[18].

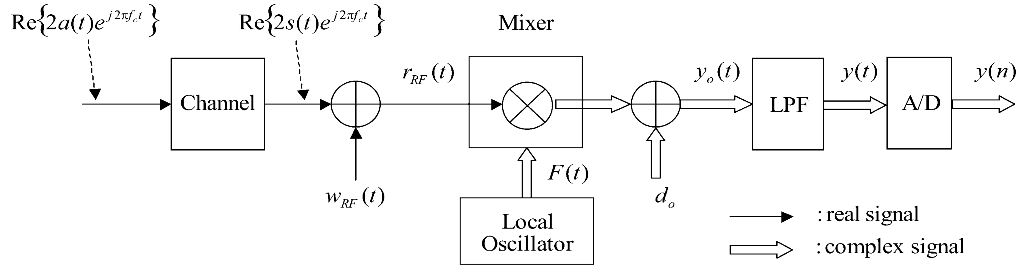

Fig. 1. Signal model.

of as a heterodyne receiver assuming $r_{\mathrm{RF}}(t)$ is an IF signal. The output of the local oscillator is expressed as

$$\begin{aligned} F(t) &= \cos(2\pi f_{\mathrm{LO}}t - \phi) - j(1+\varepsilon)\sin(2\pi f_{\mathrm{LO}}t - \phi + mbi\theta) \\ &= \beta_o e^{j\phi} e^{-j2\pi f_{\mathrm{LO}}t} + \alpha_o e^{-j\phi} e^{j2\pi f_{\mathrm{LO}}t} \end{aligned}$$

where ε and $\boldsymbol{\theta}$ denote the amplitude and phase imbalance, respectively, ϕ represents a random phase offset, $\beta_o = (1/2)(1+(1+\varepsilon)e^{-j\boldsymbol{\theta}})$, and $\alpha_o = (1/2)(1-(1+\varepsilon)e^{j\theta})$. Lowpass filtering the signal $y_o(t)$, which is equal to $r_{\mathrm{RF}}(t)F(t) + d_o$, results in the following baseband signal

$$y(t) = \beta_o x(t) + \alpha_o x^*(t) + d_o$$

where $x(t) = s(t)e^{j(2\pi\Delta f t+\phi)} + w_o(t), \Delta f$ is the frequency offset given by $\Delta f = f_c - f_{\mathrm{LO}}$, and $w_o(t)$ is the low-pass filtered version of $w_{\mathrm{RF}}(t)e^{-j2\pi f_{\mathrm{LO}}t}$. If $y(t)$ is sampled with a symbol rate, $1/T_s$, this produces

$$y(n) = \beta_o x(n) + \alpha_o x^*(n) + d_o$$

where $x(n) = s(n)e^{j(2\pi\nu n+\phi)} + w_o(n), s(n)$ is the sampled version of $s(t)$ in (1), and ν is the normalized frequency offset given by $\Delta f T_s$. Assuming a linear modulation (e.g., PSK or QAM), the signal $s(n)$ is expressed as

$$s(n) = \sum_{l=0}^{L-1} q_l a_{n-l}$$

where q_l is an impulse response of the equivalent channel due to an impulse that is applied l time units earlier; L denotes the channel length; and a_n is a complex information symbol.

Suppose that $N+L-1$ training symbols $\{a_n \,|\, n = -L+1, \ldots, N-1\}$ are transmitted.[2] Ignoring the first $L-1$ samples, the received data $\{y(0), y(1), \ldots, y(N-1)\}$ that correspond to the training sequence can be written in vector form as

$$\mathbf{y} = \beta_o(\Gamma(\nu)\mathbf{Ah} + \mathbf{w}_o) + \alpha_o(\Gamma(\nu)\mathbf{Ah} + \mathbf{w}_o)^* + d_o\mathbf{1} \quad (2)$$

where $\mathbf{y} = [y(0), y(1), \ldots, y(N-1)]^T$ and $\mathbf{1}$ is an N-dimensional vector consisting of 1s; $\Gamma(\nu)$ is a diagonal matrix given by

$$\Gamma(\nu) = \mathrm{diag}(1, \ldots, e^{j2\pi\nu(N-1)});$$

$\mathbf{A}$ is a N-by-L matrix with entries $[\mathbf{A}]_{i,j} = a_{i-j}, 0 \le i \le N-1, 0 \le j \le L-1$, and $\mathbf{h} = e^{j\phi}[q_0, q_1, \ldots, q_{L-1}]^T$. Finally, $\mathbf{w}_o = [w_o(0), w_o(1), \ldots, w_o(N-1)]^T$ is a zero-mean Gaussian vector with the covariance matrix $\sigma_o^2\mathbf{I}_N$ where $\mathbf{I}_N$ is the N-by-N identity matrix. In (2), the terms $\Gamma(\nu)\mathbf{Ah}$ and $(\Gamma(\nu)\mathbf{Ah})^*$ represent the frequency-shifted signal vector and its image vector (self-image), respectively.

Assume, for the time being, that the parameters β_o, α_o, and d_o in (2) are known. A direct calculation using (2) yields

$$\mathbf{y} - d_o\mathbf{1} - \alpha(\mathbf{y} - d_o\mathbf{1})^* = (1-|\alpha|^2)\beta_o(\Gamma(\nu)\mathbf{Ah} + \mathbf{w}_o) \quad (3)$$

where $\alpha = \alpha_o/\beta_o^*$. The right-hand side (RHS) of (3) represents a signal that is dc-offset and mismatch-free, which indicates that the dc offset and I/Q mismatch can be easily compensated using the left-hand side (LHS) of (3), once β_o, α_o, and d_o are estimated. Equation (3) is also useful for estimating β_o, α_o, d_o, and ν. To simplify the notation, let $d = d_o - \alpha d_o^*$, $\mathbf{g} = (1-|\alpha|^2)\beta_o\mathbf{h}$ and $\mathbf{w} = (1-|\alpha|^2)\beta_o\mathbf{w}_o$. Then, (3) can be rewritten as

$$\mathbf{y} - \alpha\mathbf{y}^* = \Gamma(\nu)\mathbf{Ag} + d\mathbf{1} + \mathbf{w} \quad (4)$$

where $\mathbf{w}$ represents a zero-mean Gaussian vector with the covariance matrix $\sigma^2\mathbf{I}_N$, and $\sigma^2 = (1-|\alpha|^2)^2|\beta_0|^2\sigma_0^2$. Fig. 2 illustrates the relation among the vectors in (4) when $\mathbf{w} = \mathbf{0}$. The desired term $\Gamma(\nu)\mathbf{Ag}$ lies on the subspace spanned by the columns of $\Gamma(\nu)\mathbf{A}$. The projection matrix associated with this subspace, which is denoted by $S_{\mathbf{A}}$, is given by

$$\mathbf{C}(\nu) = \Gamma(\nu)\mathbf{B}\Gamma^H(\nu) \quad (5)$$

where

$$\mathbf{B} = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H.$$

The subspace that is orthogonal to $S_{\mathbf{A}}$ is denoted by $S_{\mathbf{A}}^{\perp}$. It can be shown that the projection matrix of $S_{\mathbf{A}}^{\perp}$ is given by $\mathbf{I}_N - \mathbf{C}(\nu)$. In the following section, $\mathbf{g}, d, \alpha$, and ν are estimated in an ML sense using (4). The subspaces and their projection matrices will be useful for gaining some insight into the estimators.

III. Joint ML Estimation

For given $\mathbf{g}, d, \alpha$, and ν, the vector $\mathbf{y} - \alpha\mathbf{y}^*$ in (4) is Gaussian with the mean $\Gamma(\nu)\mathbf{Ag} + d\mathbf{1}$, and covariance matrix $\sigma^2\mathbf{I}_N$. Therefore, the conditional probability density function (pdf) for $\mathbf{y} - \bar{\alpha}\mathbf{y}^*$ is written as

$$\begin{aligned} p(\mathbf{y} - \bar{\alpha}\mathbf{y}^* | \bar{\mathbf{g}}, \bar{d}, \bar{\alpha}, \bar{\nu}) &= \frac{1}{(\pi\sigma^2)^N} \\ &\times \exp\left\{-\frac{1}{\sigma^2}\|\mathbf{y} - \bar{\alpha}\mathbf{y}^* - \bar{d}\mathbf{1} - \Gamma(\bar{\nu})\mathbf{A}\bar{\mathbf{g}}\|^2\right\} \quad (6) \end{aligned}$$

where $\bar{\mathbf{g}}, \bar{d}, \bar{\alpha}$, and $\bar{\nu}$ are trial values of $\mathbf{g}, d, \alpha$, and ν, respectively, and $\|\mathbf{x}\|^2 = \mathbf{x}^H\mathbf{x}$. The ML estimates of $(\mathbf{g}, d, \alpha, \nu)$ can

[2]The $L-1$ symbols at the start of the training sequence are *precursors*, which are given by $a_{-L+1} = a_{N-L+1}, a_{-L+2} = a_{N-L+2}, \ldots, a_{-1} = a_{N-1}$.

be obtained by maximizing the following likelihood function over $\bar{\mathbf{g}}, \bar{d}, \bar{\alpha}$, and $\bar{\nu}$:

$$\Lambda(\bar{\mathbf{g}}, \bar{d}, \bar{\alpha}, \bar{\nu}) = -\|\mathbf{y} - \bar{\alpha}\mathbf{y}^* - \bar{d}\mathbf{1} - \Gamma(\bar{\nu})\mathbf{A}\bar{\mathbf{g}}\|^2. \tag{7}$$

The location of the maximum is obtained through the four-step procedure described below [4]. First, the parameters $(\bar{d}, \bar{\alpha}, \bar{\nu})$ are fixed in (7) and the likelihood function is maximized over $\bar{\mathbf{g}}$. This yields

$$\hat{\mathbf{g}}(\bar{d}, \bar{\alpha}, \bar{\nu}) = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\Gamma^H(\bar{\nu})(\mathbf{y} - \bar{\alpha}\mathbf{y}^* - \bar{d}\mathbf{1}), \tag{8}$$

for which $\Lambda(\bar{\mathbf{g}}, \bar{d}, \bar{\alpha}, \bar{\nu})$ achieves the maximum. Second, substitute (8) into (7), which gives

$$\Lambda(\bar{d}, \bar{\alpha}, \bar{\nu}) = -\|\left(\mathbf{I}_N - \mathbf{C}(\bar{\nu})\right)\left(\mathbf{y} - \bar{\alpha}\mathbf{y}^* - \bar{d}\mathbf{1}\right)\|^2 \tag{9}$$

where $\mathbf{C}(\bar{\nu})$ is the projection matrix in (5). Maximizing (9) with respect to $\bar{d}$, while fixing $\bar{\alpha}$ and $\bar{\nu}$, yields

$$\hat{d}(\bar{\alpha}, \bar{\nu}) = \mathbf{f}^H(\bar{\nu})(\mathbf{y} - \bar{\alpha}\mathbf{y}^*) \tag{10}$$

where

$$\mathbf{f}^H(\bar{\nu}) = \frac{\mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\bar{\nu}))}{\|(\mathbf{I}_N - \mathbf{C}(\bar{\nu}))\mathbf{1}\|^2}. \tag{11}$$

Third, following the procedure similar to the 2nd step, (9) becomes

$$\Lambda(\bar{\alpha}, \bar{\nu}) = -\|(\mathbf{I}_N - \mathbf{C}(\bar{\nu}))(\mathbf{y} - \bar{\alpha}\mathbf{y}^* - \mathbf{f}^H(\bar{\nu})\left(\mathbf{y} - \bar{\alpha}\mathbf{y}^*\right)\mathbf{1})\|^2 \tag{12}$$

and the estimate of α is given by

$$\hat{\alpha}(\bar{\nu}) = \mathbf{p}^H(\bar{\nu})(\mathbf{y} - \mathbf{f}^H(\bar{\nu})\mathbf{y}\mathbf{1}) \tag{13}$$

where

$$\mathbf{p}(\bar{\nu}) = \frac{(\mathbf{I}_N - \mathbf{C}(\bar{\nu}))(\mathbf{y}^* - \mathbf{f}^H(\bar{\nu})\mathbf{y}^*\mathbf{1})}{\|(\mathbf{I}_N - \mathbf{C}(\bar{\nu}))(\mathbf{y}^* - \mathbf{f}^H(\bar{\nu})\mathbf{y}^*\mathbf{1})\|^2}.$$

Finally, using (13) in (12), it is found that ν can be estimated by maximizing

$$\begin{aligned}\Lambda(\bar{\nu}) = -\|(\mathbf{I}_N - \mathbf{C}(\bar{\nu}))(\mathbf{y} - \hat{\alpha}(\bar{\nu})\mathbf{y}^* \\ -\mathbf{f}^H(\bar{\nu})(\mathbf{y} - \hat{\alpha}(\bar{\nu})\mathbf{y}^*)\mathbf{1})\|^2.\end{aligned} \tag{14}$$

In summary, the ML estimator reads

$$\hat{\nu}_{\mathrm{ML}} = \arg\max_{\bar{\nu}}\{\Lambda(\bar{\nu})\} \tag{15}$$

$$\hat{\alpha}_{\mathrm{ML}} = \hat{\alpha}(\bar{\nu})|_{\bar{\nu}=\hat{\nu}_{\mathrm{ML}}} \tag{16}$$

$$\hat{d}_{\mathrm{ML}} = \hat{d}(\bar{\alpha}, \bar{\nu})|_{\bar{\alpha}=\hat{\alpha}_{\mathrm{ML}}, \bar{\nu}=\hat{\nu}_{\mathrm{ML}}} \tag{17}$$

$$\hat{\mathbf{g}}_{\mathrm{ML}} = \hat{\mathbf{g}}(\bar{d}, \bar{\alpha}, \bar{\nu})|_{\bar{d}=\hat{d}_{\mathrm{ML}}, \bar{\alpha}=\hat{\alpha}_{\mathrm{ML}}, \bar{\nu}=\hat{\nu}_{\mathrm{ML}}}. \tag{18}$$

The characteristics of the proposed ML estimator in (15)–(18) can be stated as follows.

1) The proposed estimator reduces to the ML estimator in [4] when $\alpha = d = 0$.
2) In (8), $\mathbf{y} - \alpha\mathbf{y}^* - d\mathbf{1}$ represents the signal vector after compensating for the I/Q mismatch and dc offset. Therefore, (8) simply suggests I/Q-mismatch and dc-offset compensation prior to channel estimation. This estimator reduces to the ML channel estimator in [6] when $\nu = \alpha = d = 0$.

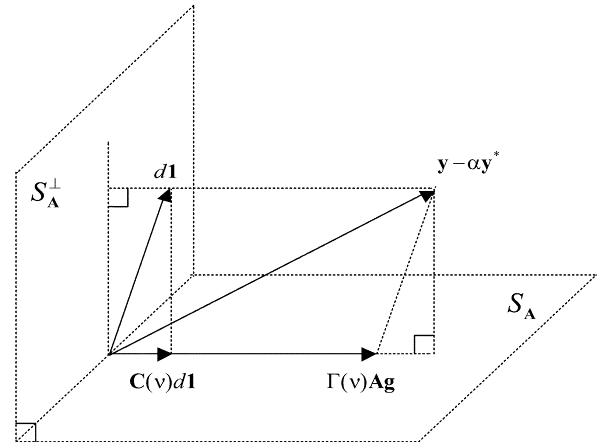

Fig. 2. Geometrical illustration of the terms in (4) when $\mathbf{w} = \mathbf{0}$. Here, $S_{\mathbf{A}}$ is the subspace spanned by the columns of $\Gamma(\nu)\mathbf{A}$, and $S_{\mathbf{A}}^{\perp}$ denotes the subspace orthogonal to $S_{\mathbf{A}}$.

3) Again referring to Fig. 2, (10) evaluates $(\mathbf{I}_N - \mathbf{C}(\nu))(\mathbf{y} - \bar{\alpha}\mathbf{y}^*)$, which is the projection of $\mathbf{y} - \bar{\alpha}\mathbf{y}^*$ onto $S_{\mathbf{A}}^{\perp}$. Note that $(\mathbf{I}_N - \mathbf{C}(\nu))(\mathbf{y} - \bar{\alpha}\mathbf{y}^*) = (\mathbf{I}_N - \mathbf{C}(\nu))d\mathbf{1}$. Therefore, under the noise-free condition $(\mathbf{w} = \mathbf{0})$,

$$\begin{aligned}\hat{d} &= (\mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\nu))^H(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{1})^{-1}\mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\nu))d\mathbf{1} \\ &= d\end{aligned}$$

where the second equality comes from $(\mathbf{I}_N - \mathbf{C}(\nu))^H(\mathbf{I}_N - \mathbf{C}(\nu)) = (\mathbf{I}_N - \mathbf{C}(\nu))$ (since $\mathbf{I}_N - \mathbf{C}(\nu)$ is the projection matrix of $S_{\mathbf{A}}^{\perp}$, it is idempotent and Hermitian symmetric).
4) The estimate of α in (13) minimizes the difference between $(\mathbf{I}_N - \mathbf{C}(\bar{\nu}))(\mathbf{y} - \bar{\alpha}\mathbf{y}^*)$ and $(\mathbf{I}_N - \mathbf{C}(\bar{\nu}))\hat{d}(\bar{\alpha}, \bar{\nu})\mathbf{1}$ [see (12)]. This minimization can be justified as follows. Assume the noise-free case $(\mathbf{w} = \mathbf{0})$. Let $\bar{\nu} = \nu$. Then

$$\Lambda(\bar{\alpha}, \nu) = -\|(\mathbf{I}_N - \mathbf{C}(\nu))(\mathbf{y} - \bar{\alpha}\mathbf{y}^*) - (\mathbf{I}_N - \mathbf{C}(\nu))\hat{d}(\bar{\alpha}, \nu)\mathbf{1}\|^2.$$

The cost $\Lambda(\bar{\alpha}, \nu)$ achieves its maximum $(\Lambda(\bar{\alpha}, \nu) = 0)$ when $\bar{\alpha} = \alpha$, because under the noise-free environment $\hat{d}(\alpha, \nu) = d$ and $(\mathbf{I}_N - \mathbf{C}(\nu))(\mathbf{y} - \alpha\mathbf{y}^*) = (\mathbf{I}_N - \mathbf{C}(\nu))d\mathbf{1}$.
5) Notice that (13) and (14) can be rewritten as

$$\begin{aligned}\hat{\alpha}(\bar{\nu}) &= \frac{\mathbf{y}^{*H}\mathbf{y} - \mathbf{y}^{*H}\mathbf{C}(\bar{\nu})\mathbf{y} + \frac{(\mathbf{c}^H(\bar{\nu})\mathbf{y})(\mathbf{c}^H(\bar{\nu})\mathbf{y}^*)^*}{\mathbf{c}^H(\bar{\nu})\mathbf{1}}}{\mathbf{y}^H\mathbf{y} - \mathbf{y}^{*H}\mathbf{C}(\bar{\nu})\mathbf{y}^* + \frac{(\mathbf{c}^H(\bar{\nu})\mathbf{y}^*)(\mathbf{c}^H(\bar{\nu})\mathbf{y}^*)^*}{\mathbf{c}^H(\bar{\nu})\mathbf{1}}} \\ \Lambda(\bar{\nu}) &= -\mathbf{y}^H\mathbf{y} + \mathbf{y}^H\mathbf{C}(\bar{\nu})\mathbf{y} + \frac{|\mathbf{c}^H(\bar{\nu})\mathbf{y}|^2}{\mathbf{c}^H(\bar{\nu})\mathbf{1}} \\ &\quad + |\hat{\alpha}(\bar{\nu})|^2\left(\mathbf{y}^H\mathbf{y} - \mathbf{y}^{*H}\mathbf{C}(\bar{\nu})\mathbf{y}^* - \frac{|\mathbf{c}^H(\bar{\nu})\mathbf{y}^*|^2}{\mathbf{c}^H(\bar{\nu})\mathbf{1}}\right)\end{aligned} \tag{19}$$

where $\mathbf{c}^H(\bar{\nu}) = \mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\bar{\nu}))$. The quadratic terms $\mathbf{y}^H\mathbf{C}(\bar{\nu})\mathbf{y}, \mathbf{y}^{*H}\mathbf{C}(\bar{\nu})\mathbf{y}$ and $\mathbf{y}^{*H}\mathbf{C}(\bar{\nu})\mathbf{y}^*$ in these equations can be evaluated using the following relationship. For any two N-dimensional column vectors

TABLE I
COMPUTATIONAL COMPLEXITIES

Algorithm	Real products	Real additions
Proposed method	$16N^2+48N$ $+12L+12$ +(8L+40)NK $+8NK\eta log_2(KN)$	$12N^2+38N$ $+6L+2$ +(6L+24)NK $+12NK\eta log_2(KN)$
Frequency estimator in [4]	$4N^2+4N$ $+2NK\eta log_2(KN)$	$3N^2+N$ $+3NK\eta log_2(KN)$
LS method in [18]	8NL+8L+4	NL+2N+2L-1

$$\mathbf{u}=[u_0,u_1,\ldots,u_{N-1}]^T \text{ and } \mathbf{v}=[v_0,v_1,\ldots,v_{N-1}]^T,$$

$$\mathbf{v}^H\mathbf{C}(\bar{\nu})\mathbf{u}=-\rho_{uv}(0)+\sum_{m=0}^{N-1}\rho_{uv}(m)e^{-j2\pi m\bar{\nu}} + \left(\sum_{m=0}^{N-1}\rho_{vu}(m)e^{-j2\pi m\bar{\nu}}\right)^* \quad (20)$$

where $\rho_{uv}(m)=\sum_{k=m}^{N-1}\mathbf{B}_{k-m,k}u_k v_{k-m}^*$ and $\mathbf{B}_{i,j}$ is the (i,j)-th element of $\mathbf{B}$. The RHS of (20) can be evaluated by applying the fast Fourier transform (FFT).

6) The maximization of $\Lambda(\bar{\nu})$ in (19) requires a numerical search over $|\bar{\nu}|\le 0.5$. The matrices $\mathbf{B}$ and $\mathbf{C}(\bar{\nu})$ in (20) as well as the vectors $\mathbf{f}(\bar{\nu})$ and $\mathbf{c}(\bar{\nu})$ in (10) and (19), respectively, can be precalculated in the initial phase. Following the search procedure in [4], the number of multiplications and additions required by the proposed estimator was obtained and summarized in Table I, where K is a design parameter called the pruning factor and $\eta=1-(\log_2 K+2(1/K-1))/\log_2(KN)$. For comparison, the computational complexities of the estimators in [4] and [18], which can be used in a cascaded fashion for the joint estimation of ν,α,d, and $\mathbf{g}$, are also shown. The proposed method always needs more computation than the cascade of [4] and [18]: the computational complexity of the former is about four times as large as that of the latter.

7) It is necessary that $N\ge L+1$, because $\mathbf{A}^H\mathbf{A}$ becomes singular when $N<L$, and $\mathbf{C}(\bar{\nu})$ reduces to $\mathbf{I}_N$ when $N=L$.

8) When $L=1$, the channel has no inter-symbol interference and the matrix $\mathbf{A}$ reduces to the vector $\mathbf{a}=[a_0,a_1,\ldots,a_{N-1}]^T$. Then, (14) becomes

$$\Lambda(\bar{\nu})=-\|\mathbf{x}(\bar{\nu})\|^2+\frac{1}{\mathbf{a}^H\mathbf{a}}|\mathbf{a}^H\Gamma^H(\bar{\nu})\mathbf{x}(\bar{\nu})|^2$$

where $\mathbf{x}(\bar{\nu})=\mathbf{y}-\hat{\alpha}(\bar{\nu})\mathbf{y}^*-\hat{d}(\hat{\alpha}(\bar{\nu}),\bar{\nu})\mathbf{1}$. Note from (4) that $\mathbf{x}(\bar{\nu})$ represents the signal after I/Q mismatch and dc offset compensation. The ML frequency estimate is given by

$$\hat{\nu}=\arg\max_{\bar{\nu}}\left\{-\sum_{n=0}^{N-1}|x_n(\bar{\nu})|^2 + \frac{1}{\sum_{n=0}^{N-1}|a_n|^2}\left|\sum_{n=0}^{N-1}x_n(\bar{\nu})a_n^*e^{-j2\pi\bar{\nu}n}\right|^2\right\}$$

where $x_n(\bar{\nu})$ denotes the nth element of $\mathbf{x}(\bar{\nu})$. This equation then reduces to the ML estimate in [4] for $L=1$ when $\mathbf{x}(\bar{\nu})=\mathbf{y}$.

IV. Performance Analysis

The mean and mean squared error (mse) of the ML estimators are examined under the assumption of a high signal-to-noise ratio (SNR). When $\alpha=d=0$, it is shown in [4] that $E[\hat{\nu}]\cong\nu$ and

$$E[(\hat{\nu}-\nu)^2]\cong\frac{\sigma^2}{2\mathbf{z}^H(\mathbf{I}_N-\mathbf{B})\mathbf{z}} \quad (21)$$

where $\hat{\nu}$ is the ML frequency estimate, $\mathbf{z}=j2\pi\mathbf{\Phi}\mathbf{A}\mathbf{g}$, and $\mathbf{\Phi}=\mathrm{diag}\{0,1,2,\ldots,N-1\}$. This mse expression will be adopted because it is reasonably accurate for practical values of α and d, and the derivation of the bias and mse for $\hat{\nu}$ is formidable when α and d are unknown.

The means and mses for $\hat{\alpha}$, $\hat{d}$, and $\hat{\mathbf{g}}$ are derived under the assumption that the frequency offset ν is known. Following from (4), the relation between the parameters $(\alpha,d,\mathbf{g})$ and their estimates $(\hat{\alpha},\hat{d},\hat{\mathbf{g}})$ can be expressed as

$$\hat{\alpha}=\alpha+\mathbf{p}^H(\nu)(\mathbf{w}-\mathbf{f}^H(\nu)\mathbf{w}\mathbf{1}) \quad (22)$$

$$\hat{d}=d-(\hat{\alpha}-\alpha)\mathbf{f}^H(\nu)\mathbf{y}^*+\mathbf{f}^H(\nu)\mathbf{w} \quad (23)$$

$$\hat{\mathbf{g}}=\mathbf{g}+\mathbf{A}^\dagger\Gamma^H(\nu)\mathbf{w}-(\hat{\alpha}-\alpha)\mathbf{A}^\dagger\Gamma^H(\nu)\mathbf{y}^* - (\hat{d}-d)\mathbf{A}^\dagger\Gamma^H(\nu)\mathbf{1} \quad (24)$$

where $\hat{\alpha}=\hat{\alpha}(\bar{\nu})|_{\bar{\nu}=\nu}, \hat{d}=\hat{d}(\bar{\alpha},\bar{\nu})|_{\bar{\alpha}=\hat{\alpha},\bar{\nu}=\nu}$, $\hat{\mathbf{g}}=\hat{\mathbf{g}}(\bar{d},\bar{\alpha},\bar{\nu})|_{\bar{d}=\hat{d},\bar{\alpha}=\hat{\alpha},\bar{\nu}=\nu}$; $\mathbf{f}(\nu)$ and $\mathbf{p}(\nu)$ are defined in (10) and (13), respectively, and $\mathbf{A}^\dagger=(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ (see Appendix A for derivation). To proceed, it is worthwhile to define $\mathbf{y}_a$ which is the deterministic part of $\mathbf{y}$ in (2)

$$\mathbf{y}_a=\beta_o\Gamma(\nu)\mathbf{A}\mathbf{h}+\alpha_o(\Gamma(\nu)\mathbf{A}\mathbf{h})^*+d_o\mathbf{1}. \quad (25)$$

Using $\mathbf{y}_a$, (2) is rewritten as

$$\begin{aligned}\mathbf{y}&=\mathbf{y}_a+\beta_o\mathbf{w}_o+\alpha_o\mathbf{w}_o^*\\&=\mathbf{y}_a+\frac{1}{1-|\alpha|^2}\mathbf{w}+\frac{\alpha}{1-|\alpha|^2}\mathbf{w}^*\end{aligned} \quad (26)$$

where the second equality comes from $\mathbf{w}=(1-|\alpha|^2)\beta_o\mathbf{w}_o$ and $\alpha=\alpha_o/\beta_o^*$. The unbiasedness of the estimators in (22)–(24) can be seen if $\mathbf{y}$ is replaced with $\mathbf{y}_a$ in (22)–(24). Therefore, the estimators are approximately unbiased when $\mathrm{SNR}\gg 1$. Specifically, the bias of $\hat{\alpha}$ may be approximated by

$$\begin{aligned}E[\hat{\alpha}-\alpha|\nu]&=E[\mathbf{p}^H(\nu)(\mathbf{w}-\mathbf{f}^H(\nu)\mathbf{w}\mathbf{1})]\\&\cong\mathbf{p}_a^H(\nu)E[\mathbf{w}-\mathbf{f}^H(\nu)\mathbf{w}\mathbf{1}]\end{aligned}$$

where $\mathbf{p}_a(\nu)$ is an approximation of $\mathbf{p}(\nu)$, which is obtained by replacing $\mathbf{y}$ with $\mathbf{y}_a$ in $\mathbf{p}(\nu)$, i.e.,

$$\mathbf{p}_a(\nu)=\frac{(\mathbf{I}_N-\mathbf{C}(\nu))(\mathbf{y}_a^*-\mathbf{f}^H(\nu)\mathbf{y}_a^*\mathbf{1})}{\|(\mathbf{I}_N-\mathbf{C}(\nu))(\mathbf{y}_a^*-\mathbf{f}^H(\nu)\mathbf{y}_a^*\mathbf{1})\|^2}. \quad (27)$$

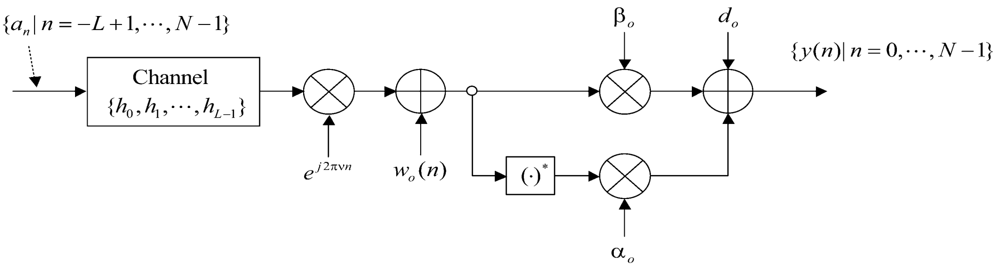

Fig. 3. System model used for simulation.

Because $E[\mathbf{w}] = \mathbf{0}$ and $\mathbf{f}^H(\nu)$ is deterministic

$$E[\hat{\alpha} - \alpha | \nu] \cong 0.$$

In a similar manner, it can be shown from (23) and (24) that

$$E[\hat{d} - d | \nu] \cong 0$$

and

$$E[\hat{\mathbf{g}} - \mathbf{g} | \nu] \cong \mathbf{0}.$$

The mse of $\hat{\alpha}$ can be approximated as

$$\begin{aligned} &E[|\hat{\alpha} - \alpha|^2 | \nu] \\ &\quad = E[|\mathbf{p}^H(\nu)(\mathbf{w} - \mathbf{f}^H(\nu)\mathbf{w}\mathbf{1})|^2] \\ &\quad \cong \mathbf{p}_a^H(\nu) E[(\mathbf{w} - \mathbf{f}^H(\nu)\mathbf{w}\mathbf{1})(\mathbf{w} - \mathbf{f}^H(\nu)\mathbf{w}\mathbf{1})^H]\mathbf{p}_a(\nu). \end{aligned} \tag{28}$$

After the calculation outlined in Appendix B, (28) is rewritten as

$$E[|\hat{\alpha} - \alpha|^2 | \nu] \cong \frac{\sigma^2}{\|(\mathbf{I}_N - \mathbf{C}(\nu))(\mathbf{y}_a^* - \mathbf{f}^H(\nu)\mathbf{y}_a^*\mathbf{1})\|^2}. \tag{29}$$

The mse for $\hat{d}$ is derived as follows: from (23) and (26), the conditional mse given $\hat{\alpha}$ and ν can be expressed as

$$\begin{aligned} &E[|\hat{d} - d|^2 | \hat{\alpha}, \nu] \\ &\quad = |\hat{\alpha} - \alpha|^2 |\mathbf{f}^H(\nu)\mathbf{y}_a^*|^2 + \gamma(\hat{\alpha}) E[|\mathbf{f}^H(\nu)\mathbf{w}|^2] \\ &\quad = |\hat{\alpha} - \alpha|^2 |\mathbf{f}^H(\nu)\mathbf{y}_a^*|^2 + \gamma(\hat{\alpha}) \frac{\sigma^2}{N - \mathbf{1}^H\mathbf{C}(\nu)\mathbf{1}} \end{aligned} \tag{30}$$

where $\gamma(\hat{\alpha}) = (|1 - \hat{\alpha}\alpha^*|^2 + |\hat{\alpha} - \alpha|^2)/|1 - |\alpha|^2|^2$. The second equality in (30) is true because $E[|\mathbf{f}^H(\nu)\mathbf{w}|^2] = \sigma^2 \mathbf{f}^H(\nu)\mathbf{f}(\nu)$ and

$$\begin{aligned} \mathbf{f}^H(\nu)\mathbf{f}(\nu) &= \frac{1}{\|(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{1}\|^2} \\ &= \frac{1}{N - \mathbf{1}^H\mathbf{C}(\nu)\mathbf{1}} \end{aligned} \tag{31}$$

where (31) comes from $\mathbf{C}(\nu)\mathbf{C}(\nu) = \mathbf{C}(\nu)$. Taking the expectation over $\hat{\alpha}$ and approximating $E[\gamma(\hat{\alpha})]$ by $E[\gamma(\hat{\alpha})] \cong 1 + E[|\hat{\alpha} - \alpha|^2]/|1 - |\alpha|^2|^2 \triangleq \gamma_a$ produces

$$\begin{aligned} E[|\hat{d} - d|^2 | \nu] \cong E[|\hat{\alpha} - \alpha|^2 | \nu] |\mathbf{f}^H(\nu)\mathbf{y}_a^*|^2 \\ + \frac{\gamma_a \sigma^2}{N - \mathbf{1}^H\mathbf{C}(\nu)\mathbf{1}}. \end{aligned} \tag{32}$$

The mse for $\hat{\mathbf{g}}$ can be derived in a similar manner. Using (26) in (24), the conditional mse given ν, $\hat{\alpha}$, and $\hat{d}$ is written as

$$\begin{aligned} &E[\|\hat{\mathbf{g}} - \mathbf{g}\|^2 | \hat{d}, \hat{\alpha}, \nu] \\ &\quad = \gamma(\hat{\alpha})\sigma^2 \mathrm{tr}((\mathbf{A}^H\mathbf{A})^{-1}) + |\hat{\alpha} - \alpha|^2 \|\mathbf{A}^\dagger \Gamma^H(\nu)\mathbf{y}_a^*\|^2 \\ &\qquad + |\hat{d} - d|^2 \|\mathbf{A}^\dagger \Gamma^H(\nu)\mathbf{1}\|^2 \\ &\qquad + 2\mathrm{Re}((\hat{\alpha} - \alpha)^*(\hat{d} - d)(\mathbf{A}^\dagger \Gamma^H(\nu)\mathbf{y}_a^*)^H \mathbf{A}^\dagger \Gamma^H(\nu)\mathbf{1}). \end{aligned} \tag{33}$$

From (23), it can be shown that

$$E[(\hat{\alpha} - \alpha)^*(\hat{d} - d) | \nu] \cong -E[|\hat{\alpha} - \alpha|^2 | \nu] \mathbf{f}^H(\nu)\mathbf{y}_a^*.$$

Hence, the expectation of (33) over $\hat{\alpha}$ and $\hat{d}$ becomes

$$\begin{aligned} &E[\|\hat{\mathbf{g}} - \mathbf{g}\|^2 | \nu] \\ &\quad \cong \gamma_a \sigma^2 \mathrm{tr}((\mathbf{A}^H\mathbf{A})^{-1}) + E[|\hat{\alpha} - \alpha|^2 | \nu] \|\mathbf{A}^\dagger \Gamma^H(\nu)\mathbf{y}_a^*\|^2 \\ &\qquad + E[|\hat{d} - d|^2 | \nu] \|\mathbf{A}^\dagger \Gamma^H(\nu)\mathbf{1}\|^2 - 2E[|\hat{\alpha} - \alpha|^2 | \nu] \\ &\qquad \times \mathrm{Re}(\mathbf{f}^H(\nu)\mathbf{y}_a^*(\mathbf{A}^\dagger \Gamma^H(\nu)\mathbf{y}_a^*)^H \mathbf{A}^\dagger \Gamma^H(\nu)\mathbf{1}). \end{aligned} \tag{34}$$

The mse expressions in (29), (32), and (34) are verified by simulation in the following section.

V. Simulation Results

Computer simulations were conducted to examine the performance of the proposed estimator. In the first set of simulations, the analytical results of the previous section were confirmed and extended. Then, performances of the proposed estimator and an existing scheme were compared.

Fig. 3 shows the system model used for the simulation. The training sequence was the midamble of the GSM system, given by $\{a(n) | n = 0, \ldots, N-1\} = \{1, -j, 1, j, 1, -j, -1, -j, -1, j, -1, -j, -1, j, -1, -j\} (N = 16)$ [4]. Two types of channel models were considered: a fixed channel given by $\{h_0, h_1, h_2\} = \{1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}\} (L = 3)$, and the typical urban channel model of the GSM system with six paths, which was employed in [4]. The GSM channel response was represented as

$$h_k = \sum_{i=0}^{5} \xi_i g_T(kT_s - \tau_i - t_0) \tag{35}$$

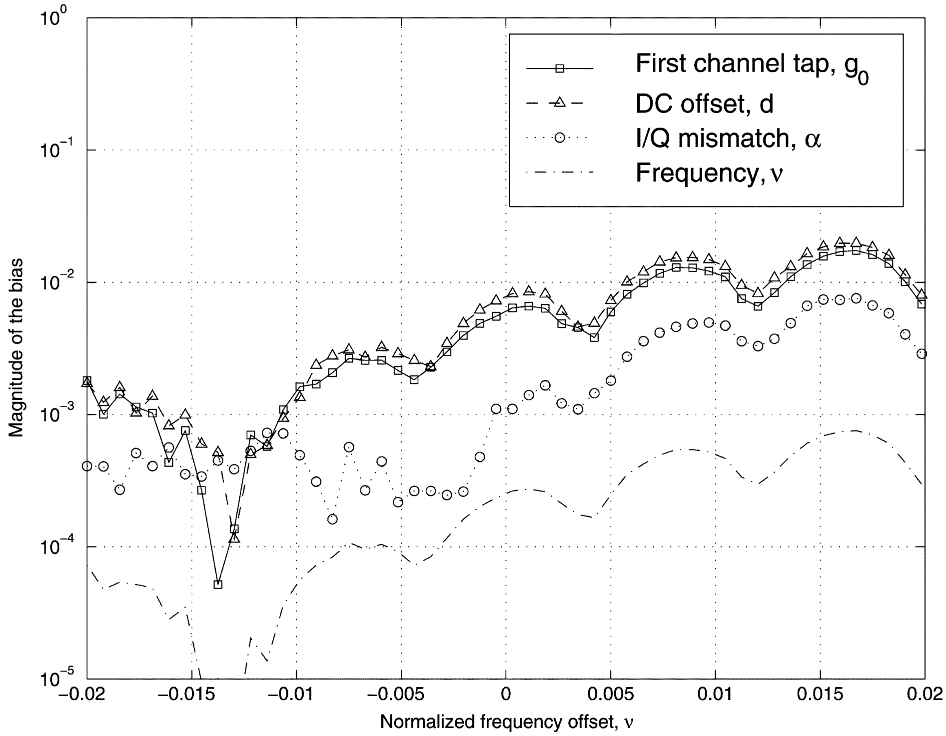

Fig. 4. Bias versus ν when $\mathrm{E_b/N_o} = 20$ dB and $|d_o| = 0.1$.

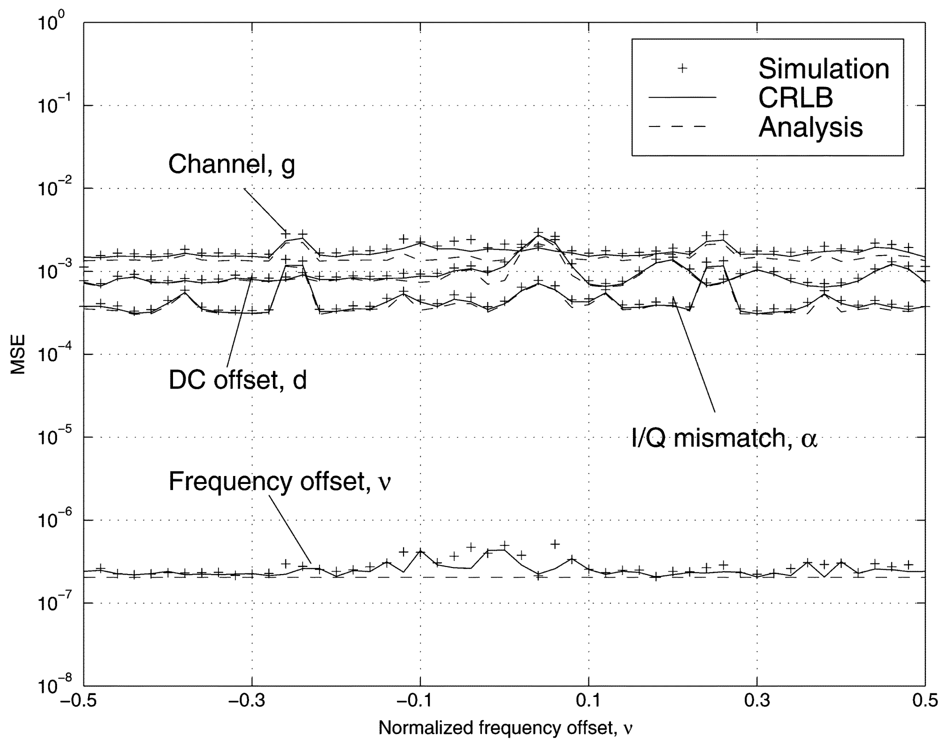

Fig. 5. mses versus ν when $\mathrm{E_b/N_o} = 20$ dB and $|d_o| = 0.1$.

where $g_T(t)$ is the impulse response of the raised-cosine filter with a rolloff of 0.5; $\{\xi_i\}$ and $\{\tau_i\}$ are the attenuations and path delays, respectively; and t_0 is the timing phase. All the parameters for generating h_k in (35) were equal to those of [4]. For this GSM channel, $L = 8$. The simulation was performed for various values of $\nu \in [-0.5, 0.5]$, which was the maximum estimation range of the frequency estimator.[3] The I/Q mismatch param-

[3]In practice, the frequency estimation range becomes considerably narrower than its maximum value due to various design issues such as adjacent channel effects.

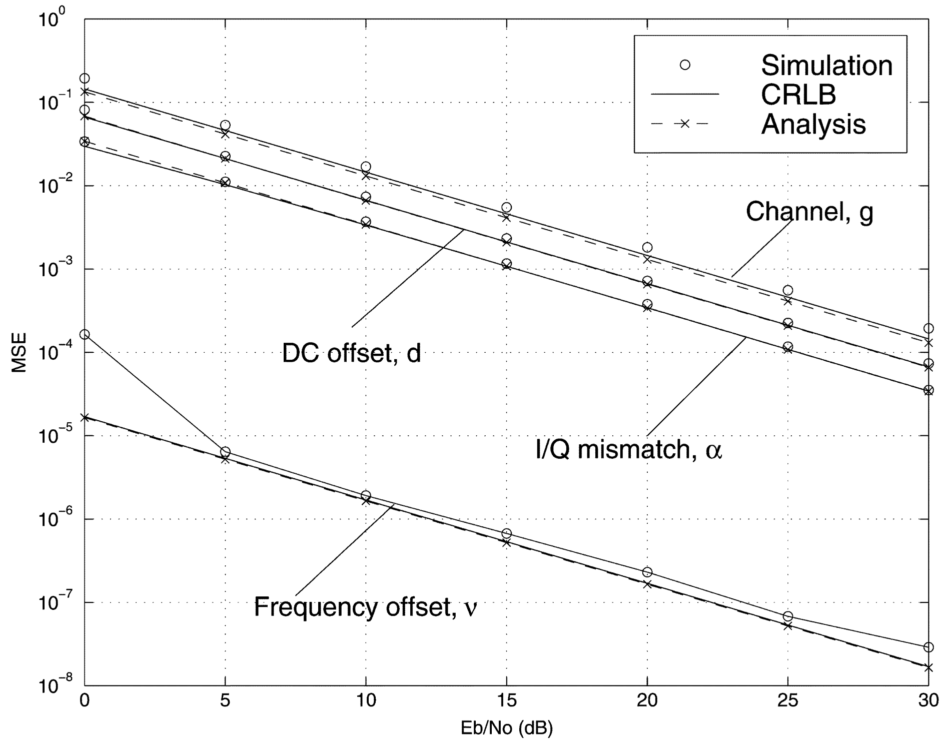

Fig. 6. mses versus E_b/N_o when $\nu = 0.1$ and $|d_o| = 0.1$.

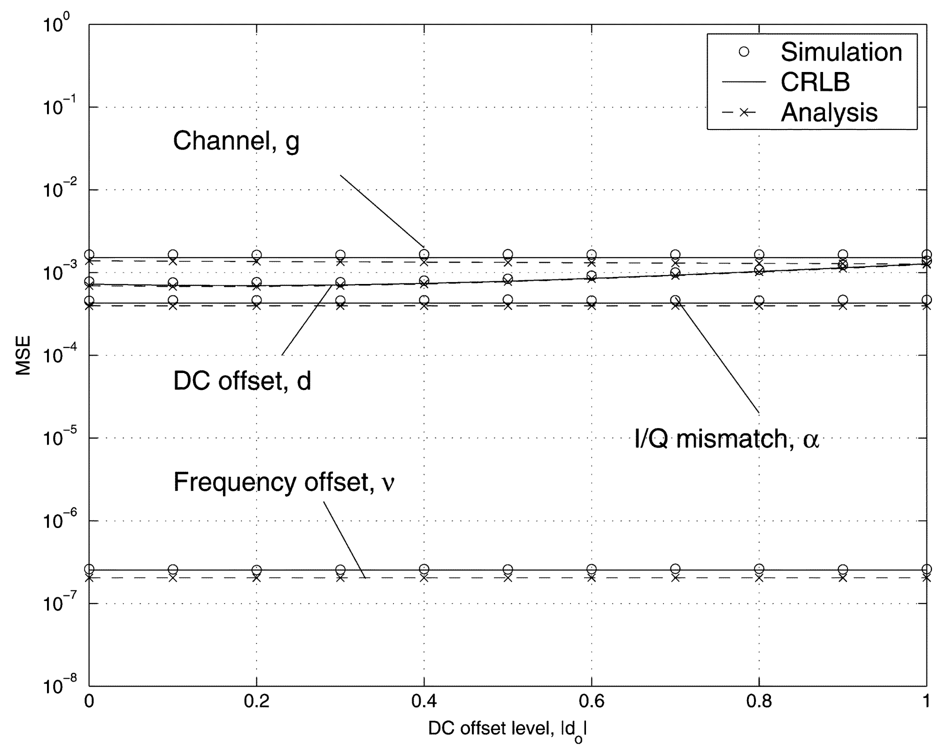

Fig. 7. mses versus $|d_o|$ when $E_b/N_o = 20$ dB and $\nu = 0.1$.

eter α was set at $-0.048 - j0.0873$, corresponding to $\epsilon = 0.1$ and $\boldsymbol{\theta} = 10^o$. The dc offset was given by $d_o = |d_o| \times (1+j)/\sqrt{2}$, where $|d_o| \in [0, 1]$. The pruning factor K for evaluating $\hat{\nu}$ was 8 ($K = 8$). The mean and mse values were estimated through 10 000 trials. The Cramér-Rao lower bounds (CRLB) of the estimates (see Appendix C for derivation) were also evaluated and compared with the mses.

A. Comparison Between Analytical and Simulation Results

The mean and mse expressions of the previous section were checked using the fixed channel with $L = 3$. The mean values obtained through the simulation indicated that the ML estimates were affected by some bias. Fig. 4 illustrates this bias when $E_b/N_o = 20$ dB and $|d_o| = 0.1$ (for simplicity, only the bias of

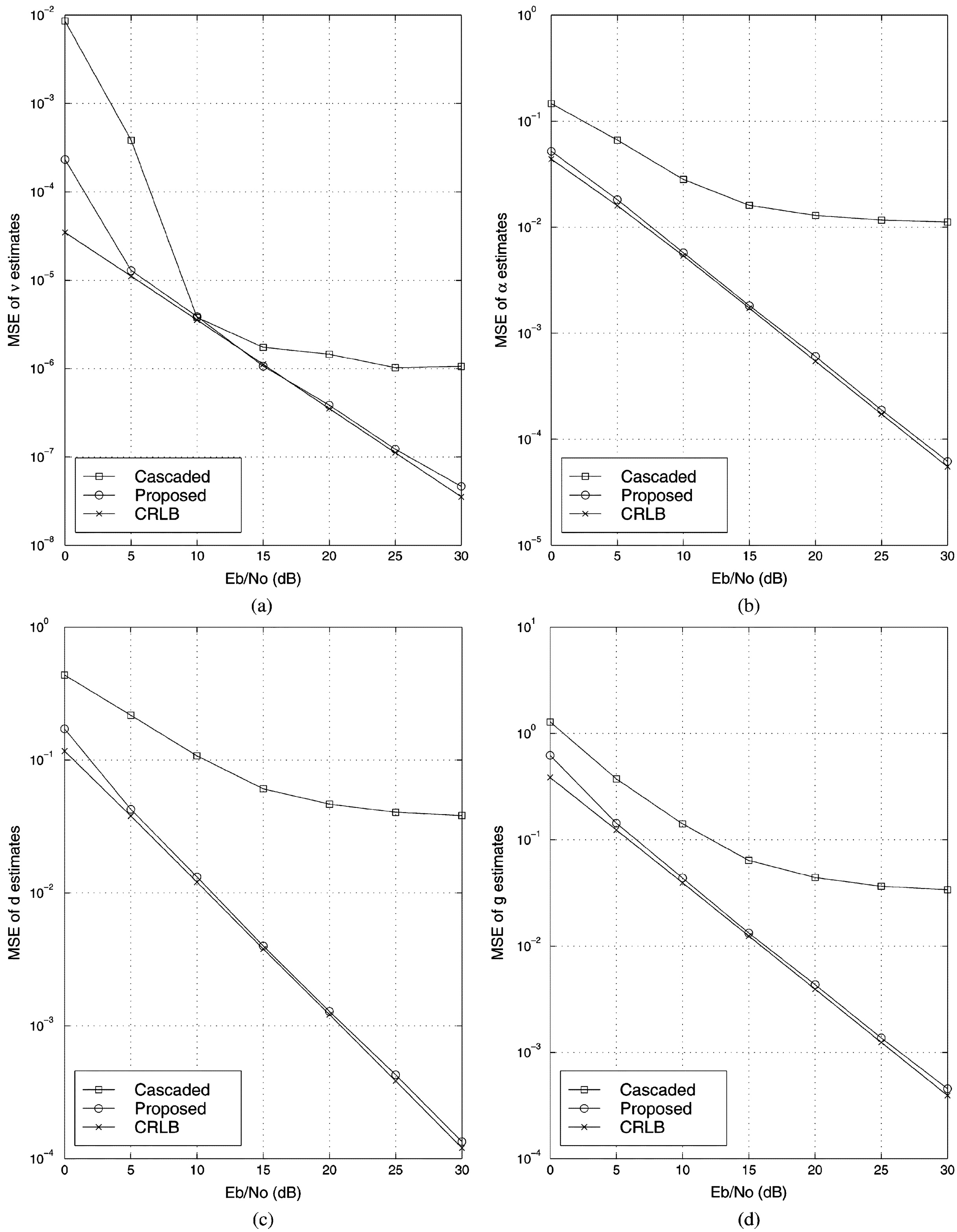

Fig. 8. mses of the estimators and the corresponding CRLBs for GSM channel, when ν was uniformly distributed in $[-0.033, 0.033]$. (a) Frequency offset. (b) I/Q mismatch. (c) dc offset. (d) Channel.

the first channel tap g_0 is shown). As observed in [4], a microbias appeared for the ML frequency estimate $\hat{\nu}$. The bias of $\hat{\nu}$ in turn caused some increase in the bias of the other estimates, which then fluctuated depending on the former. It was observed that the bias of α, d, and g_0 reduced to about 10^{-4} when $\hat{\nu} = \nu$ (perfect frequency estimation).

Figs. 5–7 compare the mses from the simulation with those from (29), (32), and (34), plus the corresponding CRLBs. A remarkably good agreement was observed between the analysis and simulation results. The estimation performance was robust to frequency and dc offsets (see Figs. 5 and 7). Furthermore, the mses were almost identical to the CRLBs. (In some cases, the

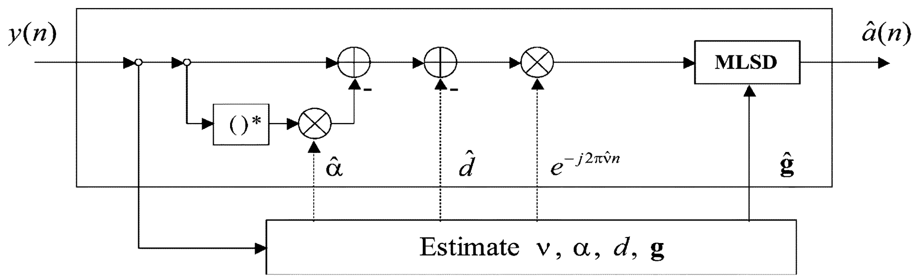

Fig. 9. Digital signal processing for recovering $\{a_n\}$.

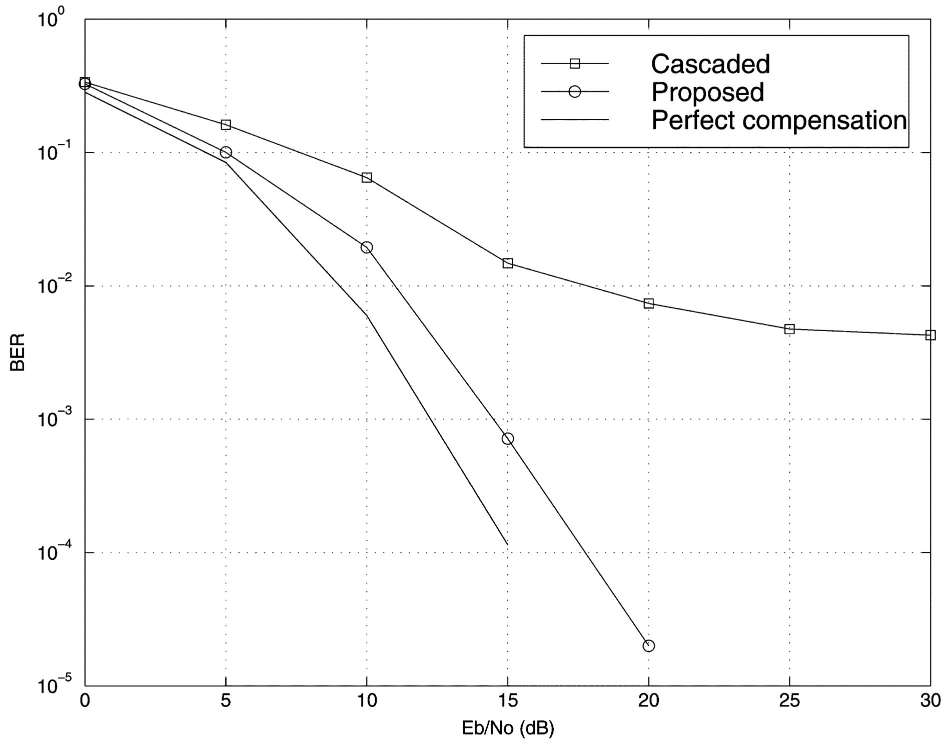

Fig. 10. Comparison of BER performances for the GSM channel.

analytical mses were slightly smaller than the CRLBs, which occurred because the former were obtained under the assumption that $\hat{\nu} = \nu$.)

B. Performance Comparison

The performance of the proposed ML estimate was examined for the GSM channel with $L = 8$. For each simulation run, $\{\xi_i\}$ in (35) and the frequency offset ν were generated using Gaussian and uniform random number generators, respectively. It was assumed that $\nu \in [-0.033, 0.033]$ and $|d_o| = 0.1$. For comparison, a cascade of the estimators in [4] and [18] was considered. This technique, which will be referred to as the cascaded method, estimates and compensates for the frequency offset ν using the frequency estimator in [4], then estimates α, d, and $\mathbf{g}$ using the LS method in [18].

Fig. 8 shows the mses of the estimates and the corresponding CRLBs (the CRLBs were obtained by averaging the CRLBs from each simulation run). The proposed scheme outperformed the cascaded method. The mses of the former were close to the CRLBs, while those of the latter exhibited error floors. This happened because I/Q imbalance and dc offset deteriorate the frequency estimation performance of the method in [4]. It was also observed that the LS method in [18] behaved like the proposed ML method when $\nu = 0$. However, it exhibited a severe performance degradation even for small values of ν. This is the reason why the frequency offset was estimated prior to the estimation of the other parameters.

The performances of the estimators for the GSM channel were also examined in terms of the bit error rates (BERs). The receiver structure for the BER simulation is shown in Fig. 9. The I/Q mismatch, dc, and frequency offsets were compensated according to (4), then information symbols, which were quadrature phase shift keying (QPSK) symbols, were detected by applying the maximum likelihood sequence detector (MLSD). The transmitted data were grouped into frames consisting of 19 training symbols followed by 19 information symbols.[4] For simplicity in implementation, the MLSD assumed a channel of length 4 and only used the 4 taps of the GSM channel, which contain 99.95% of the channel power. The BERs are shown in Fig. 10. As expected, the proposed

[4]The duration of the information symbols was limited due to the lack of a phase tracking scheme that can compensate for the residual frequency offset.

scheme outperformed the cascaded method, which exhibited an error floor. When the $\mathrm{BER} = 10^{-3}$, the proposed scheme was only 2 dB worse than the case of perfect compensation.

VI. Conclusion

A data-aided method for jointly estimating the I/Q mismatch, dc offset, frequency offset, and channel for receivers in a frequency-selective quasistatic channel was proposed. This estimator is particularly useful for direct-conversion receivers. The performance of the estimator was investigated analytically and by simulation, which demonstrated that the mses of the estimates were close to the CRLBs, and that the proposed scheme could outperform an existing technique. Further work in this area will include the following topics.

1) Development of an I/Q mismatch compensation technique for low-IF receivers.
2) Extension of the proposed method to the cases with frequency-dependent I/Q mismatch.
3) Study of improved cascaded approaches.

Appendix A
Proof of (22)–(24)

Using (9), the term $\mathbf{y} - \mathbf{f}^H\mathbf{y}\mathbf{1}$ in (13) can be written as

$$\mathbf{y} - \mathbf{f}^H(\nu)\mathbf{y}\mathbf{1} = \alpha(\mathbf{y}^* - \mathbf{f}^H(\nu)\mathbf{y}^*\mathbf{1}) + \Gamma(\nu)\mathbf{A}\mathbf{g} + \mathbf{w} - \mathbf{f}^H(\nu)\mathbf{w}\mathbf{1} \quad (36)$$

where the equality comes from $\mathbf{f}^H(\nu)\Gamma(\nu)\mathbf{A}\mathbf{g} = 0$ (see Fig. 2). Substituting (36) into (13) yields

$$\hat{\alpha}(\nu) = \alpha + \mathbf{p}^H(\nu)(\Gamma(\nu)\mathbf{A}\mathbf{g} + \mathbf{w} - \mathbf{f}^H\mathbf{w}\mathbf{1}) \quad (37)$$

because $(\mathbf{I}_N - \mathbf{C}(\nu))(\mathbf{I}_N - \mathbf{C}(\nu)) = \mathbf{I}_N - \mathbf{C}(\nu)$ and thus $\mathbf{p}^H(\nu)(\mathbf{y}^* - \mathbf{f}^H(\nu)\mathbf{y}^*\mathbf{1}) = 1$. Since $(\mathbf{I}_N - \mathbf{C}(\nu))\Gamma(\nu)\mathbf{A} = \mathbf{0}$, then (22) follows from (37).

To derive (23), (4) is rewritten as

$$\mathbf{y} - \hat{\alpha}\mathbf{y}^* = d\mathbf{1} + \Gamma(\nu)\mathbf{A}\mathbf{g} - (\hat{\alpha} - \alpha)\mathbf{y}^* + \mathbf{w}. \quad (38)$$

Using (38) in (10), $\hat{d}$ is written as

$$\hat{d} = \mathbf{f}^H(\nu)(d\mathbf{1} + \Gamma(\nu)\mathbf{A}\mathbf{g} - (\hat{\alpha} - \alpha)\mathbf{y}^* + \mathbf{w}). \quad (39)$$

(23) follows from (39), because $\mathbf{f}^H(\nu)\mathbf{1} = 1$ and $\mathbf{f}^H(\nu)\Gamma(\nu)\mathbf{A}\mathbf{g} = 0$.

From (8), $\hat{\mathbf{g}}$ is written as

$$\hat{\mathbf{g}} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\Gamma^H(\nu)(\mathbf{y} - \hat{\alpha}\mathbf{y}^* - \hat{d}\mathbf{1}) \quad (40)$$

and from (4)

$$\mathbf{y} - \hat{\alpha}\mathbf{y}^* - \hat{d}\mathbf{1} = \Gamma(\nu)\mathbf{A}\mathbf{g} - (\hat{\alpha} - \alpha)\mathbf{y}^* - (\hat{d} - d)\mathbf{1} + \mathbf{w}. \quad (41)$$

Using (41) in (40) yields (24). This is true because $\Gamma^H(\nu)\Gamma(\nu) = \mathbf{I}_N$ and $(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\Gamma^H(\nu)\Gamma(\nu)\mathbf{A} = \mathbf{I}_L$.

Appendix B
Proof of (29)

Because $E[\mathbf{w}\mathbf{w}^H] = \sigma^2\mathbf{I}_N$

$$E[(\mathbf{f}^H(\nu)\mathbf{w}\mathbf{1})\mathbf{w}^H|\nu] = \sigma^2\mathbf{1}\mathbf{f}^H(\nu) \quad (42)$$

and

$$E[(\mathbf{f}^H(\nu)\mathbf{w}\mathbf{1})(\mathbf{f}^H(\nu)\mathbf{w}\mathbf{1})^H|\nu] = \sigma^2\mathbf{f}^H(\nu)\mathbf{f}(\nu)\mathbf{1}\mathbf{1}^H. \quad (43)$$

Using (42) and (43) in (28) yields

$$E[|\hat{\alpha} - \alpha|^2|\nu] \cong \sigma^2\mathbf{p}_a^H(\nu)(\mathbf{I}_N - \mathbf{1}\mathbf{f}^H(\nu) - \mathbf{f}(\nu)\mathbf{1}^H + \mathbf{f}^H(\nu)\mathbf{f}(\nu)\mathbf{1}\mathbf{1}^H)\mathbf{p}_a(\nu) \quad (44)$$

where $\mathbf{f}^H(\nu)$ and $\mathbf{p}_a(\nu)$ is defined by (11) and (27), respectively. Note that

$$\mathbf{p}_a^H(\nu)\mathbf{1} = \frac{(\mathbf{y}_a^* - \mathbf{f}^H(\nu)\mathbf{y}_a^*\mathbf{1})^H(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{1}}{\|(\mathbf{I}_N - \mathbf{C}(\nu))(\mathbf{y}_a^* - \mathbf{f}^H(\nu)\mathbf{y}_a^*\mathbf{1})\|^2} \quad (45)$$

and

$$(\mathbf{f}^H(\nu)\mathbf{y}_a^*\mathbf{1})^H(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{1} = (\mathbf{y}_a^*)^H(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{1}. \quad (46)$$

Here, (46) holds because

$$(\mathbf{f}^H(\nu)\mathbf{y}_a^*)^* = \frac{(\mathbf{y}_a^*)^H(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{1}}{\mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{1}}.$$

Using (46) in (45) produces $\mathbf{p}_a^H(\nu)\mathbf{1} = 0$. Hence, (44) reduces to

$$E[|\hat{\alpha} - \alpha|^2|\nu] \cong \sigma^2\mathbf{p}_a^H(\nu)\mathbf{p}_a(\nu). \quad (47)$$

The RHS of (47) is simplified as (29), because

$$\mathbf{p}_a^H(\nu)\mathbf{p}_a(\nu) = \frac{1}{\|(\mathbf{I}_N - \mathbf{C}(\nu))(\mathbf{y}_a^* - \mathbf{f}^H(\nu)\mathbf{y}_a^*\mathbf{1})\|^2}.$$

Appendix C
Derivation of CRLBs

The CRLBs of ν, α, d, and $\mathbf{g}$ can be derived following the procedure in [19]. To begin, define a vector

$$\boldsymbol{\theta} = (\nu \quad \alpha \quad \alpha^* \quad d \quad d^* \quad \mathbf{g}^T \quad \mathbf{g}^H)^T.$$

Let $\hat{\boldsymbol{\theta}}$ be an unbiased estimate of $\boldsymbol{\theta} \in \mathbb{C}^{(2L+5)\times 1}$. It can be proved that

$$\mathrm{cov}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}) - \mathbf{M}^{-1} \quad (48)$$

is positive semidefinite where $\mathrm{cov}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}) = E[\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}^H], \mathbf{M} = E[((\partial ln f(\boldsymbol{\theta}))/(\partial\boldsymbol{\theta}^T))^H(\partial ln f(\boldsymbol{\theta}))/(\partial\boldsymbol{\theta}^T)], f(\boldsymbol{\theta})$ denotes the pdf in (6) and $(\partial ln f(\boldsymbol{\theta}))/(\partial\boldsymbol{\theta}^T)$ is a $(2L+5)$-dimensional row vector written by the first equation shown at the bottom of the next page. From (48)

$$\mathrm{var}(\hat{\nu}) \geq [\mathbf{M}^{-1}]_{11} \quad (49)$$

$$\mathrm{var}(\hat{\alpha}) \geq [\mathbf{M}^{-1}]_{22} \tag{50}$$
$$\mathrm{var}(\hat{d}) \geq [\mathbf{M}^{-1}]_{44} \tag{51}$$
$$E[\|\hat{\mathbf{g}} - \mathbf{g}\|^2] \geq \sum_{i=6}^{6+L-1} [\mathbf{M}^{-1}]_{ii} \tag{52}$$

and the RHS of (49)–(52) represent the CRLBs. After some calculation, it can be shown that $(\partial lnf(\boldsymbol{\theta}))/(\partial\nu) = (1/\sigma_n^2)(\mathbf{z}^H\Gamma^H(\nu)\mathbf{w}_o + \mathbf{w}_o^H\Gamma(\nu)\mathbf{z})$, $(\partial lnf(\boldsymbol{\theta}))/(\partial\alpha) = (1/\sigma_n^2)\mathbf{w}_o^H\mathbf{y}^*, (\partial lnf(\boldsymbol{\theta}))/(\partial\alpha^*) = ((\partial lnf(\boldsymbol{\theta}))/(\partial\alpha))^*$, $(\partial lnf(\boldsymbol{\theta}))/(\partial d) = (1/\sigma_n^2)\mathbf{w}_o^H\mathbf{1}, (\partial lnf(\boldsymbol{\theta}))/(\partial d^*) = ((\partial lnf(\boldsymbol{\theta}))/(\partial d))^*$, $(\partial lnf(\boldsymbol{\theta}))/(\partial\mathbf{g}^T) = (1/\sigma_n^2)\mathbf{w}_o^H\Gamma(\nu)\mathbf{A}$, and $(\partial lnf(\boldsymbol{\theta}))/(\partial\mathbf{g}^{*T}) = ((\partial lnf(\boldsymbol{\theta}))/(\partial\mathbf{g}^T))^*$, where $\mathbf{z}$ is defined in (21). Therefore, $\mathbf{M}$ can be expressed as

$$\mathbf{M} = \frac{1}{\sigma^2}\begin{pmatrix} \mathbf{D} & \mathbf{C}^H \\ \mathbf{C} & \mathbf{R} \end{pmatrix}$$

where $\mathbf{D}, \mathbf{C}$, and $\mathbf{R}$ are 3×3, $(2L+2) \times 3$, and $(2L+2) \times (2L+2)$ matrices, respectively, given by the second equation at the bottom of the page. Here

$$s = \mathbf{y}_a^H\mathbf{y}_a + \sigma^2\frac{2 + |\alpha|^2 N(N+1)}{(1-|\alpha|^2)^2}$$
$$c = \sigma^2\frac{N(N+1)\alpha^2}{(1-|\alpha|^2)^2}$$

and $\mathbf{y}_a$ is defined in (25). Using the identity in the inverses of the block matrices [20, p. 728] and after lengthy computations

$$[\mathbf{M}^{-1}]_{11} = \sigma^2\left(u_{11} - \frac{2\mathrm{Re}(u_{22}|u_{12}|^2 - cu_{12}^2)}{|u_{22}|^2 - |c|^2}\right)^{-1}$$
$$[\mathbf{M}^{-1}]_{22} = \sigma^2\left(u_{22} - \frac{|u_{12}|^2}{u_{11}} - \frac{|cu_{11} - (u_{12}^*)^2|^2}{u_{11}(u_{11}u_{22} - |u_{12}|^2)}\right)^{-1}$$

where

$$u_{11} = 2\mathbf{z}^H(\mathbf{I}_N - \mathbf{B})\mathbf{z} - a^{-2}|\mathbf{z}^H\Gamma^H(\nu)(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{1}|^2$$
$$u_{12} = \mathbf{z}^H\Gamma^H(\nu)(\mathbf{I}_N - \mathbf{C}(\nu))[\mathbf{I}_N - a^{-1}\mathbf{1}\mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\nu))]\mathbf{y}_a^*$$
$$u_{22} = s - \mathbf{y}_a^T\mathbf{C}(\nu)\mathbf{y}_a^* - a^{-1}|\mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{y}_a^*|^2$$

and $a = \|(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{1}\|^2$. In addition

$$[\mathbf{M}^{-1}]_{44} = \sigma^2 a^{-1} + \sigma^2 a^{-2}[v_{11}|q_{11}|^2 + v_{22}|q_{12}|^2 + 2\mathrm{Re}(v_{12}q_{11}q_{12}^*)]$$
$$\sum_{i=6}^{6+L-1}[\mathbf{M}^{-1}]_{ii} = \sigma^2(\mathrm{tr}((\mathbf{A}^H\mathbf{A})^{-1}) + a^{-1}\|(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\Gamma^H(\nu)\mathbf{1}\|^2) + \sigma^2\left[v_{11}\|\mathbf{q}_{31}\|^2 + v_{22}\|\mathbf{q}_{32}\|^2 + 2\mathrm{Re}(v_{12}\mathbf{q}_{32}^H\mathbf{q}_{31})\right]$$

where

$$v_{11} = \left(u_{11} - \frac{2\mathrm{Re}(u_{22}|u_{12}|^2 - c|u_{12}|^2)}{|u_{22}|^2 - |c|^2}\right)^{-1}$$
$$v_{12} = \left(u_{12}^* - \frac{u_{11}\left(|c|^2 - u_{22}^2\right) + u_{22}|u_{12}|^2 - cu_{12}^2}{c^*u_{12}^* - u_{12}u_{22}^*}\right)^{-1},$$
$$v_{22} = \left(u_{22} - \frac{u_{11}|c|^2 - c^*|u_{12}|^2 - cu_{12}^2 + |u_{12}|^2u_{22}}{u_{11}u_{22} - |u_{12}|^2}\right)^{-1}$$
$$q_{11} = \mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\nu))\Gamma(\nu)\mathbf{z}$$
$$q_{12} = \mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\nu))\mathbf{y}_a^*$$
$$\mathbf{q}_{31} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\Gamma^H(\nu)\left[\mathbf{I}_N - \frac{1}{a}\mathbf{1}\mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\nu))\right]\Gamma(\nu)\mathbf{z}$$

and

$$\mathbf{q}_{32} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\Gamma^H(\nu)\left[\mathbf{I}_N - \frac{1}{a}\mathbf{1}\mathbf{1}^H(\mathbf{I}_N - \mathbf{C}(\nu))\right]\mathbf{y}_a^*.$$

$$\frac{\partial lnf(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^T} = \begin{pmatrix} \frac{\partial lnf(\boldsymbol{\theta})}{\partial\nu} & \frac{\partial lnf(\boldsymbol{\theta})}{\partial\alpha} & \frac{\partial lnf(\boldsymbol{\theta})}{\partial\alpha^*} & \frac{\partial lnf(\boldsymbol{\theta})}{\partial d} & \frac{\partial lnf(\boldsymbol{\theta})}{\partial d^*} & \frac{\partial lnf(\boldsymbol{\theta})}{\partial\mathbf{g}^T} & \frac{\partial lnf(\boldsymbol{\theta})}{\partial\mathbf{g}^{*T}} \end{pmatrix}.$$

$$\mathbf{D} = \begin{pmatrix} 2\mathbf{z}^H\mathbf{z} & \mathbf{z}^H\Gamma^H(\nu)\mathbf{y}_a^* & \left(\mathbf{z}^H\Gamma^H(\nu)\mathbf{y}_a^*\right)^* \\ \left(\mathbf{z}^H\Gamma^H(\nu)\mathbf{y}_a^*\right)^* & s & c \\ \mathbf{z}^H\Gamma^H(\nu)\mathbf{y}_a^* & c^* & s \end{pmatrix}$$
$$\mathbf{C} = \begin{pmatrix} (\mathbf{z}^H\Gamma^H(\nu)\mathbf{1})^* & \mathbf{y}_a^H\mathbf{1} & 0 \\ \mathbf{z}^H\Gamma^H(\nu)\mathbf{1} & 0 & \mathbf{y}_a^T\mathbf{1} \\ (\mathbf{z}^H\mathbf{A})^H & \left(\mathbf{y}_a^T\Gamma(\nu)\mathbf{A}\right)^H & 0 \\ (\mathbf{z}^H\mathbf{A})^T & 0 & (\mathbf{y}_a^T\Gamma(\nu)\mathbf{A})^T \end{pmatrix}$$

and

$$\mathbf{R} = \begin{pmatrix} N & 0 & \mathbf{1}^H\Gamma(\nu)\mathbf{A} & \mathbf{0} \\ 0 & N & 0 & \left(\mathbf{1}^H\Gamma(\nu)\mathbf{A}\right)^* \\ \left(\mathbf{1}^H\Gamma(\nu)\mathbf{A}\right)^H & 0 & \mathbf{A}^H\mathbf{A} & \mathbf{0} \\ 0 & \left(\mathbf{1}^H\Gamma(\nu)\mathbf{A}\right)^T & \mathbf{0} & (\mathbf{A}^H\mathbf{A})^* \end{pmatrix}.$$

REFERENCES

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

[2] H. Meyer, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. New York: Wiley, 1998.

[3] 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.

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

[5] E. R. Jeong, S. K. Jo, and Y. H. Lee, "Least squares frequency estimation in frequency-selective channels and its application to transmissions with antenna diversity," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 2369–2380, Dec. 2001.

[6] S. N. Crozier, D. D. Falconer, and S. A. Mahmoud, "Least sum of squared errors (LSSE) channel estimation," *Proc. Inst. Elect. Eng.*, pt. F, vol. 138, pp. 371–378, Aug. 1991.

[7] A. N. D'Andrea, A. Diglio, and U. Mengali, "Symbol-aided channel estimation with nonselective rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 44, pp. 41–49, Feb. 1995.

[8] C. Tellambura, Y. J. Guo, and S. K. Barton, "Channel estimation using aperiodic binary sequences," *IEEE Commun. Lett.*, vol. 2, pp. 140–142, May 1998.

[9] A. A. Abidi, "Direct-conversion radio transceivers for digital communications," *IEEE J. Solid-State Circuits*, vol. 30, pp. 1399–1410, Dec. 1995.

[10] B. Razavi, "Design considerations for direct-conversion receivers," *IEEE Trans. Circuits Syst.*, vol. 44, pp. 428–435, Jun. 1997.

[11] C. C. Chen and C. C. Huang, "On the architecture and performance of a hybrid image rejection receiver," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1029–1040, Jun. 2001.

[12] J. Crols and M. Steyaert, "Low-IF topologies for high-performance analog front ends of fully integrated receivers," *IEEE Trans. Circuits Syst. II*, vol. 45, pp. 269–282, Mar. 1998.

[13] F. E. Churchill, G. W. Ogar, and B. J. Thomson, "The correction of I and Q errors in a coherent processor," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 17, pp. 131–137, Jan. 1981.

[14] J. K. Cavers and M. W. Liao, "Adaptive compensation for imbalance and offset losses in direct conversion tranceivers," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 581–588, Nov. 1993.

[15] L. Yu and W. M. Sneglove, "A novel adaptive mismatch cancellation system for quadrature IF radio receivers," *IEEE Trans. Circuits Syst. II*, vol. 46, pp. 789–801, Jun. 1999.

[16] K. Pun, J. E. Franca, and C. Azeredo-Leme, "Wideband digital correction of I and Q mismatch in quadrature radio receivers," in *Proc. IEEE ISCAS*, Geneva, Switzerland, May 2000, pp. 661–664.

[17] M. Valkama, M. Renfors, and V. Koivunen, "Advanced methods for I/Q imbalance compensation in communication receivers," *IEEE Trans. Signal Processing*, vol. 49, pp. 2335–2344, Oct. 2001.

[18] I. H. Sohn, E. R. Jeong, and Y. H. Lee, "Data-aided approach to I/Q mismatch and DC offset compensation in communication receivers," *IEEE Commun. Lett.*, vol. 6, pp. 547–549, Dec. 2002.

[19] A. van den Bos, "A Cramér-Rao bound for complex parameters," *IEEE Trans. Signal Processing*, vol. 42, p. 2859, Oct. 1994.

[20] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 2000.

Gye-Tae Gil (S'00) was born in Daejeon, Korea, on July 24, 1966. He received the B.S. degree in electrical engineering from Hanyang University, Seoul, Korea, in 1989, and the M.S. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Taejon, in 1992 and 2004, respectively.

Since 1991, he has been with the research center of Korea Telecom, where he is currently a senior member of research staff. His research interests are in the area of communication signal processing, which includes synchronization, interference cancellation, and adaptive filter design.

Il-Hyun Sohn received the B.S. degree in electrical engineering from Seoul National University, Seoul, Korea, in 1989, and the M.S. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Taejon, in 1991 and 2003, respectively.

From April 1991 to December 1995, he was at the Korea Telecom, Seoul, where he was involved in the projects of speech signal processing. Since 2003, he has been with the SK Telesys, Seoul, where he is mainly involved in the projects of wireless communication system. His research interests include digital signal processing and digital communication systems.

Jin-Kyu Park received the B.S. degree in electronic engineering from Kyung-Book National university in 1979, and the M.S degrees in electrical and electronic engineering from the Korea Advanced Institute and Science Technology (KAIST), Taejon, in 1988.

He joined the Agency for Defense Development (ADD) of the Republic of Korea, where he worked as a researcher in several kinds of military development projects since 1979. From 1979 to 1985, he was involved in the development of Surface to Surface Missile System, and from 1986 to 2000, he was working on the development of the Surface to Air Missile Systems. Since 2002, he has been a Project Manager for the development of radar system in ADD.

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 with the Department of Electrical and Computer Engineering, State University of New York, Buffalo. Since 1989, he has been with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, where he currently is a Professor. His research activities are in the area of communication signal processing, which includes synchronization, estimation, detection, interference cancellation, and efficient filter design for CDMA, TDMA, and orthogonal frequency-division multiplexing systems. He is also interested in designing and implementing transceivers.