Joint ML Estimation of I/Q mismatch, DC Offset, Carrier Frequency, and Channel for Direct-Conversion Receivers

Gate Gil*, Il-hyun Sohn*, Yong H. Lee*
*Dept. of EECS, Korea Advanced Institute of Science and Technology, Taejon, 305-701, Korea
Email: yohlee@ee.kaist.ac.kr

Young Ik Song†, Jin Kyu Park‡
†R & D Team, Samsung Thales Co., Ltd, ‡Agency for Defense Development, Korea

Abstract—**This paper proposes a technique to jointly estimate and compensate for I/Q mismatch, DC offset and carrier frequency offset in frequency selective channel environments for direct-conversion receivers. It estimates the unknown parameters in a joint maximum-likelihood (ML) sense from a training sequence. In addition, a design method of training sequences is proposed to minimize the mean-square errors (MSEs) of the proposed estimates. It is shown through simulation that the proposed technique can provide significant performance improvement over the existing schemes, and the proposed method of designing training sequences is effective as well.**

I. Introduction

The frequency and channel estimation techniques that have been introduced so far assume an ideal receiver in which there are no errors in RF and analog circuit with the exception of frequency offset. This assumption can be justified for receivers based on the traditional superheterodyne principle. However, its justification can be difficult for some other receiver architectures such as those with direct conversion [1], and a low-IF [2], which accompany several deficiencies: for example, both direct conversion and low-IF architectures suffer from I/Q mismatch which is an imbalance between I and Q branch amplitude and phase. For these receivers, it would be recommended to estimate frequency and channel in presence of I/Q mismatch.

In this paper, we develop a data-aided, ML technique that jointly estimate frequency, channel, I/Q mismatch and DC offset, where DC offset is a particular concern in direct conversion receivers. The proposed estimator can be thought of as an extension of the techniques in [3] and [4]. Here, [3] proposes a joint ML scheme for frequency and channel estimation, and [4] proposes a joint least squares (LS) estimate of I/Q mismatch, DC offset and channel. It will be observed that the proposed scheme reduces to the estimate in [3] when there are no I/Q mismatch and DC offset, and that it performs better than the estimate in [4] when frequency offset exists. Like the estimate in [4], the proposed scheme can be applied to both direct-conversion and conventional heterodyne architectures, yet not to those with a low-IF.

This paper is organized as follows. Section II describes the signal model used in the paper. The proposed method is derived in Section III, and its properties are analyzed in Section IV. Section V presents criteria for designing training sequences that minimize the estimation error. Simulation results showing performance characteristics of the proposed technique are presented in Section VI.

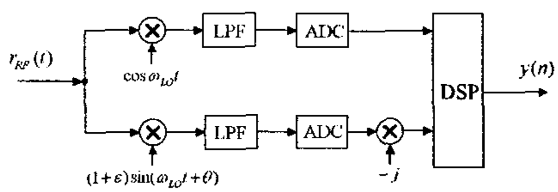

Fig. 1. Direct-conversion receiver model.

II. Signal Model

The receiver model considered in this paper is shown in Fig. 1. Here, $r_{RF}(t)$ is the received signal in front of the mixers; w_{LO} denotes the local oscillator frequency; ε and θ denote the amplitude and phase imbalance, respectively. The received signal samples, taken at a symbol rate, can be expressed as

$$y(n) = \beta_o x(n) + \alpha_o x^*(n) + d_o, \tag{1}$$

where $x(n) = s(n)e^{j2\pi\nu n} + w_o(n)$, ν is the normalized frequency offset, $w_o(n)$ is AWGN, $\beta_o = \frac{1}{2}\left(1 + (1+\varepsilon)e^{-j\theta}\right)$ and $\alpha_o = \frac{1}{2}\left(1 - (1+\varepsilon)e^{j\theta}\right)$, d_o is the DC offset. Assuming a linear modulation (e.g., PSK or QAM), the signal $s(n)$ is expressed as

$$s(n) = \sum_{l=0}^{L-1} h_l a_{n-l} \tag{2}$$

where h_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. Ignoring the first $L-1$ samples, the received data $\{y(0), y(1), \cdots, y(N-1)\}$ that correspond to the training sequence can be written in vector form as

$$\mathbf{y} = \beta_o\left(\Gamma(\nu)\mathbf{A}\mathbf{h} + \mathbf{w}_o\right) + \alpha_o\left(\Gamma(\nu)\mathbf{A}\mathbf{h} + \mathbf{w}_o\right)^* + d_o\mathbf{1}, \tag{3}$$

where $\mathbf{y} = [y(0), y(1), \cdots, y(N-1)]^T$; $\Gamma(\nu)$ is a diagonal matrix given by

$$\Gamma(\nu) = diag(1, \cdots, e^{j2\pi\nu(N-1)}); \quad (4)$$

$\mathbf{A}$ 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 $\mathbf{h} = [h_0, h_1, \ldots, h_{L-1}]^T$. Finally, $\mathbf{w}_o = [w_o(0), w_o(1), \cdots, w_o(N-1)]^T$ is a zero-mean Gaussian vector with covariance matrix $\sigma_o^2\mathbf{I}_N$ where $\mathbf{I}_N$ is the N-by-N identity matrix. In (3), 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 parameter β_o, α_o and d_o in (3) are known. A direct calculation using (3) yields

$$\mathbf{y} - d_o\mathbf{1} - \alpha(\mathbf{y} - d_o\mathbf{1})^* = (1-|\alpha|^2)\beta_o(\Gamma(\nu)\mathbf{Ah} + \mathbf{w}_o) \quad (5)$$

where $\alpha = \alpha_o/\beta_o^*$ and $\mathbf{1}$ is an N-dimensional column vector consisting of 1's. The right-hand-side (RHS) of (5) represents a signal which is DC offset and mismatch free. This fact indicates that the DC offset and I/Q mismatch can be easily compensated using the left-hand-side (LHS) of (5), once β_o, α_o and d_o are estimated. Equation (5) 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, (5) can be rewritten as

$$\mathbf{y} - \alpha\mathbf{y}^* = \Gamma(\nu)\mathbf{Ag} + d\mathbf{1} + \mathbf{w}, \quad (6)$$

where $\mathbf{w}$ represents a zero-mean Gaussian vector with covariance matrix $\sigma^2\mathbf{I}_N$, and $\sigma^2 = (1-|\alpha|^2)^2|\beta_0|^2\sigma_0^2$. In the following section, $\mathbf{g}$, d, α and ν are estimated in an ML sense using (6).

III. Joint ML Estimation

For given $\mathbf{g}$, d, α and ν, the vector $\mathbf{y} - \alpha\mathbf{y}^*$ in (6) is Gaussian with 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

$$\begin{array}{c} p(\mathbf{y} - \bar{\alpha}\mathbf{y}^*|\bar{\mathbf{g}}, \bar{d}, \bar{\alpha}, \bar{\nu}) = \frac{1}{(\pi\sigma^2)^N} \times \\ \exp\{-\frac{1}{\sigma^2}\|\mathbf{y} - \bar{\alpha}\mathbf{y}^* - \bar{d}\mathbf{1} - \Gamma(\bar{\nu})\mathbf{A}\bar{\mathbf{g}}\|^2\}, \end{array} \quad (7)$$

where $\bar{\mathbf{g}}$, $\bar{d}$, $\bar{\alpha}$ and $\bar{\nu}$ are trial values of $\mathbf{g}$, d, α and ν, respectively, and $\|\mathbf{x}\|^2 = \mathbf{x}^H\mathbf{x}$. The ML estimates of ($\mathbf{g}$,$d$,$\alpha$, ν) can be obtained by maximizing the following likelihood function over $\bar{\mathbf{g}}$, $\bar{d}$, $\bar{\alpha}$ and $\bar{\nu}$:

$$\Lambda(\bar{\mathbf{g}}, \bar{d}, \bar{\alpha}, \bar{\nu}) = -\|\mathbf{y} - \bar{\alpha}\mathbf{y}^* - \bar{d}\mathbf{1} - \Gamma(\bar{\nu})\mathbf{A}\bar{\mathbf{g}}\|^2. \quad (8)$$

The location of the maximum is obtained through the four step procedure described below [3].

First, the parameters $(\bar{d}, \bar{\alpha}, \bar{\nu})$ are fixed in (8) and the likelihood function is maximized over $\bar{\mathbf{g}}$. This yields

$$\hat{\mathbf{g}}(\bar{d}, \bar{\alpha}, \bar{\nu}) = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\Gamma^H(\bar{\nu})\left(\mathbf{y} - \bar{\alpha}\mathbf{y}^* - \bar{d}\mathbf{1}\right), \quad (9)$$

for which $\Lambda(\bar{\mathbf{g}}, \bar{d}, \bar{\alpha}, \bar{\nu})$ achieves a maximum.

Second, substitute (9) into (8), which gives

$$\Lambda(\bar{d}, \bar{\alpha}, \bar{\nu}) = -\|\left(\mathbf{I}_N - \mathbf{C}(\bar{\nu})\right)\left(\mathbf{y} - \bar{\alpha}\mathbf{y}^* - \bar{d}\mathbf{1}\right)\|^2, \quad (10)$$

where $\mathbf{C}(\bar{\nu}) = \Gamma(\bar{\nu})\mathbf{B}\Gamma^H(\bar{\nu})$ and $\mathbf{B} = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$. Maximizing (10) with respect to $\bar{d}$, while fixing $\bar{\alpha}$ and $\bar{\nu}$, yields

$$\hat{d}(\bar{\alpha}, \bar{\nu}) = \mathbf{f}^H(\bar{\nu})\left(\mathbf{y} - \bar{\alpha}\mathbf{y}^*\right), \quad (11)$$

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

Third, following the procedure similar to the 2nd step, (10) reduces to

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

and the estimate of α is given by

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

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

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

$$\Lambda(\bar{\nu}) = -\|\left(\mathbf{I}_N - \mathbf{C}(\bar{\nu})\right)\mathbf{x}(\bar{\nu})\|^2, \quad (14)$$

where $\mathbf{x}(\bar{\nu}) = \mathbf{y} - \hat{\alpha}(\bar{\nu})\mathbf{y}^* - \mathbf{f}^H(\bar{\nu})\left(\mathbf{y} - \hat{\alpha}(\bar{\nu})\mathbf{y}^*\right)\mathbf{1}$.

In summary, the ML estimator reads

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

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

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

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

Notice that (13) and (14) may be rewritten as

$$\hat{\alpha}(\bar{\nu}) = \frac{\mathbf{y}^{*H}\mathbf{y} - \mathbf{y}^{*H}\mathbf{C}(\bar{\nu})\mathbf{y} + \frac{(\mathbf{c}^H(\bar{\nu})\mathbf{y})(\mathbf{c}^H(\bar{\nu})\mathbf{y}^*)^*}{\mathbf{c}^H(\bar{\nu})\mathbf{1}}}{\mathbf{y}^H\mathbf{y} - \mathbf{y}^{*H}\mathbf{C}(\bar{\nu})\mathbf{y}^* + \frac{(\mathbf{c}^H(\bar{\nu})\mathbf{y}^*)(\mathbf{c}^H(\bar{\nu})\mathbf{y}^*)^*}{\mathbf{c}^H(\bar{\nu})\mathbf{1}}} \quad (19)$$

$$\begin{array}{rcl} \Lambda(\bar{\nu}) & = & -\mathbf{y}^H\mathbf{y} + \mathbf{y}^H\mathbf{C}(\bar{\nu})\mathbf{y} + \frac{|\mathbf{c}^H(\bar{\nu})\mathbf{y}|^2}{\mathbf{c}^H(\bar{\nu})\mathbf{1}} \\ & + & |\hat{\alpha}(\bar{\nu})|^2\left(\mathbf{y}^H\mathbf{y} - \mathbf{y}^{*H}\mathbf{C}(\bar{\nu})\mathbf{y}^* - \frac{|\mathbf{c}^H(\bar{\nu})\mathbf{y}^*|^2}{\mathbf{c}^H(\bar{\nu})\mathbf{1}}\right) \end{array} \quad (20)$$

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 $\mathbf{a}$ and $\mathbf{b}$,

$$\begin{array}{rcl} \mathbf{b}^H\mathbf{C}(\bar{\nu})\mathbf{a} & = & -\rho_{ab}(0) + \sum_{m=0}^{N-1}\rho_{ab}(m)e^{-j2\pi m\bar{\nu}} \\ & + & \left(\sum_{m=0}^{N-1}\rho_{ba}(m)e^{-j2\pi m\bar{\nu}}\right)^*, \end{array} \quad (21)$$

where $\rho_{ab}(m) = \sum_{k=m}^{N-1}\mathbf{B}_{k-m,k}a(k)b^*(k-m)$, and $\mathbf{B}_{i,j}$ is the (i,j)-th element of $\mathbf{B}$. Maximization of $\Lambda(\bar{\nu})$ in (20) requires a numerical search over $|\bar{\nu}| \le 0.5$ and fast Fourier transform (FFT) can be applied to evaluating the RHS of (21). The matrices $\mathbf{B}$ and $\mathbf{C}(\bar{\nu})$ in (21) as well as the vectors $\mathbf{f}(\bar{\nu})$ and $\mathbf{c}(\bar{\nu})$ in (11) and (20), respectively, can be precalculated at the initial phase. Following the search procedure in [3], it can be shown that the proposed estimator needs $(8L+40)NK + 12L + 2 + 8(2N^2+5N) + 8NK\eta\log_2(KN)$ real multiplications and $(6L+24)NK + 6L - 4 + 6(2N^2+5N) + 12NK\eta\log_2(KN)$ real additions, where K is a design parameter called the pruning factor and $\eta = 1 - (\log_2 K + 2(1/K-1))/\log_2(KN)$.

IV. Derivation of MSEs

Expressions for the MSEs of the proposed estimator are derived. Throughout this section, $\hat{\nu}$, $\hat{\alpha}$, $\hat{d}$ and $\hat{\mathbf{g}}$ denote the ML estimates in (15)-(18).

An approximate expression for the MSE of $\hat{\nu}$ has been derived in [3] when $\alpha = d = 0$. It is written as

$$E[(\hat{\nu}-\nu)^2] \cong \frac{\sigma^2}{2\mathbf{z}^H(\mathbf{I}_N-\mathbf{B})\mathbf{z}}, \quad (22)$$

where $\mathbf{z} = j2\pi\mathbf{MAg}$ and $\mathbf{M} = diag\{0,1,2,\cdots,N-1\}$. This MSE expression will be adopted because it is reasonably accurate for practical values of α and d, as shown in the following subsection (derivation of the MSE for $\hat{\nu}$ is formidable when α and d are unknown).

Under the assumption that the frequency offset ν is known, the MSEs for $\hat{\alpha}$, $\hat{d}$ and $\hat{\mathbf{g}}$ may be approximated at high signal-to-noise ratio (SNR) as

$$E[|\hat{\alpha}-\alpha|^2] \cong \frac{\sigma^2}{\|(\mathbf{I}_N-\mathbf{C}(\nu))(\mathbf{y}_a^*-\mathbf{f}^H(\nu)\mathbf{y}_a^*\mathbf{1})\|^2}, \quad (23)$$

$$E\left[|\hat{d}-d|^2\right] \cong E[|\hat{\alpha}-\alpha|^2]|\mathbf{f}^H(\nu)\mathbf{y}_a^*|^2 + \frac{\gamma_a\sigma^2}{N-\mathbf{1}^H\mathbf{C}(\nu)\mathbf{1}}, \quad (24)$$

$$\begin{aligned} E[\|\hat{\mathbf{g}}-\mathbf{g}\|^2] \cong\; & \gamma_a\sigma^2 tr\left((\mathbf{A}^H\mathbf{A})^{-1}\right) \\ & + E[|\hat{\alpha}-\alpha|^2]\|\mathbf{A}^\dagger\Gamma^H(\nu)\mathbf{y}_a^*\|^2 \\ & + E\left[|\hat{d}-d|^2\right]\|\mathbf{A}^\dagger\Gamma^H(\nu)\mathbf{1}\|^2 \\ & - 2E[|\hat{\alpha}-\alpha|^2]\xi, \end{aligned} \quad (25)$$

where $\xi = Re\left(\mathbf{f}^H(\nu)\mathbf{y}_a^*(\mathbf{A}^\dagger\Gamma^H(\nu)\mathbf{y}_a^*)^H\mathbf{A}^\dagger\Gamma^H(\nu)\mathbf{1}\right)$, $\gamma_a = 1 + \frac{E[|\hat{\alpha}-\alpha|^2]}{|1-|\alpha|^2|^2}$, and $\mathbf{y}_a = \beta_o\Gamma(\nu)\mathbf{Ah} + \alpha_o(\Gamma(\nu)\mathbf{Ah})^* + d_o\mathbf{1}$.

V. Training Sequence Design

Conditions for minimizing the MSEs of $\hat{\nu}$, $\hat{\alpha}$, $\hat{d}$ and $\hat{\mathbf{g}}$ are derived. Then, a procedure for selecting an optimal TS is described.

A. Conditions for minimum MSEs

The minimization of the MSE of $\hat{\nu}$ has been considered in [5] and [6]. The result can be stated as follows.

Property 1: The MSE of $\hat{\nu}$ in (22) can be minimized if

$$tr(\mathbf{Q})det(\mathbf{Q}) \quad (26)$$

is maximized, where $\mathbf{Q} = C_{\mathbf{gg}}^{\frac{1}{2}}\mathbf{A}^H\mathbf{M}(\mathbf{I}_N-\mathbf{B})\mathbf{MA}C_{\mathbf{gg}}^{\frac{1}{2}}$, $C_{\mathbf{gg}} = E[\mathbf{gg}^H]$, and $C_{\mathbf{gg}}$ is assumed to be diagonal.

In [6], training sequences maximizing (26) are obtained when $\mathbf{C_{gg}} = \mathbf{I}$, and they are shown to be useful for reducing MSEs of $\hat{\nu}$ in practical channel environments.

Property 2: Simultaneous minimization of the MSEs of $\hat{\alpha}$, $\hat{d}$ and $\hat{\mathbf{g}}$ in (23), (24) and (25) requires the following conditions :

$$\mathbf{A}^H\mathbf{A} = N\mathbf{I}_L \quad (27)$$

$$(\Gamma(\nu)\mathbf{A})^H(\Gamma(\nu)\mathbf{A})^* = \mathbf{0} \quad (28)$$

$$(\Gamma(\nu)\mathbf{A})^H\mathbf{1} = \mathbf{0} \quad (29)$$

Proof: see [7] for the proof. ∎

It is interesting to note that the conditions in (27), (28) and (29) tend to decouple the estimation of α, d and $\mathbf{g}$. Actually, if we multiply both sides of (6) with $(\Gamma(\nu)\mathbf{A})^T$, then

$$\begin{aligned} (\Gamma(\nu)\mathbf{A})^T\mathbf{y} &= \alpha(\Gamma(\nu)\mathbf{A})^T\mathbf{y}^* + d(\Gamma(\nu)\mathbf{A})^T\mathbf{1} \\ &+ (\Gamma(\nu)\mathbf{A})^T\Gamma(\nu)\mathbf{Ag} + (\Gamma(\nu)\mathbf{A})^T\mathbf{w} \\ &= \alpha(\Gamma(\nu)\mathbf{A})^T\mathbf{y}^* + (\Gamma(\nu)\mathbf{A})^T\mathbf{w}, \end{aligned} \quad (30)$$

where the second equality comes from (28) and (29). From (30), α can be estimated irrespective of d and $\mathbf{g}$. Similarly, from (29), multiplying both sides of (6) with $\mathbf{1}^H$ yields

$$\begin{aligned} \mathbf{1}^H\mathbf{y} - \alpha\mathbf{1}^H\mathbf{y}^* &= Nd + \mathbf{1}^H\Gamma(\nu)\mathbf{Ag} + \mathbf{1}^H\mathbf{w} \\ &= Nd + \mathbf{1}^H\mathbf{w}, \end{aligned} \quad (31)$$

and d can be estimated irrespective of $\mathbf{g}$. Finally, from (29), multiplying both sides of (6) with $(\Gamma(\nu)\mathbf{A})^H$ yields

$$\begin{aligned} (\Gamma(\nu)\mathbf{A})^H(\mathbf{y}-\alpha\mathbf{y}^*) &= d(\Gamma(\nu)\mathbf{A})^H\mathbf{1} + \mathbf{A}^H\mathbf{Ag} \\ &+ (\Gamma(\nu)\mathbf{A})^H\mathbf{w} \\ &= \mathbf{A}^H\mathbf{Ag} + (\Gamma(\nu)\mathbf{A})^H\mathbf{w}, \end{aligned} \quad (32)$$

and estimation of $\mathbf{g}$ is not influenced by d.

B. Training sequence selection method

It can be simply shown that no finite sequence can satisfy the conditions in (28) and (29) in a contiguous range of ν. So, for the selection of an optimal training sequence, a qualification measure is introduced. From (30), (31), and (32), it can be shown that the interference power for estimation of each parameter for a particular ν can be minimized by minimizing the qualification measure defined as

$$\begin{aligned} J(\nu) &= \|C_{\mathbf{gg}}^{\frac{1}{2}}(\Gamma(\nu)\mathbf{A})^H(\Gamma(\nu)\mathbf{A})^*\|_F^2 \\ &+ \|C_{\mathbf{gg}}^{\frac{1}{2}}(\Gamma(\nu)\mathbf{A})^H\mathbf{1}\|^2. \end{aligned} \quad (33)$$

where $C_{\mathbf{gg}} = \mathbf{I}_L$ for standard white Gaussian channel. Because the condition of $\mathbf{A}^H\mathbf{A} = N\mathbf{I}$ greatly simplifies the estimators of ν, α, d and $\mathbf{g}$, and it also contributes to lower the MSEs, the qualification measure was defined to select an optimal sequence among the sequences satisfying this condition. One more thing reflected in (33) is that the impact of $\|C_{\mathbf{gg}}^{\frac{1}{2}}(\Gamma(\nu)\mathbf{A})^H(\Gamma(\nu)\mathbf{A})^*\|_F^2$ and $\|C_{\mathbf{gg}}^{\frac{1}{2}}(\Gamma(\nu)\mathbf{A})^H\mathbf{1}\|^2$ on the SNR can be shown to be approximately equal.

Since the qualification measure assumes negligible frequency estimation error, and ν is a random variable, an optimal sequence with the least value of $E[J(\nu)]$ should be chosen from those that have sufficiently large value of $tr(\mathbf{Q})det(\mathbf{Q})$ in (26).

VI. Simulation Results

In this section, assuming that the starting point of the training symbols is perfectly known, simulation results are provided to verify the accuracy of the analytical results, the effectiveness of the designing method of training sequences, and the performance improvement of the proposed method compared to the estimators in [3] and [4].

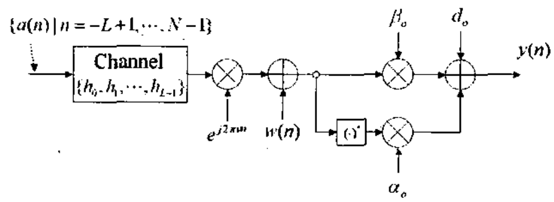

Fig. 2. System model used for simulation.

A. Simulation model

Fig. 2 shows the system model used for simulation. The parameters applied in the simulation was as follows: $\epsilon = 0.1$, $\theta = 10^\circ$, $d_o = 0.1 \times \frac{1}{\sqrt{2}}(1+i)\sqrt{E_s}$, where E_s is the average power of the training symbol defined by $E_s = \frac{1}{N} E\left[\sum_{n=0}^{N-1} |s(n)|^2\right]$. The frequency offset was assumed to be uniformly distributed in [−0.033 0.033]. For the simulation in random channels, the typical urban (TU) channel model of the European GSM system with six path [3, p.1584] ($L = 8$) was assumed. Because more than 99.95% of the channel power is concentrated in the 4 taps, bit error rate (BER) simulation was done with $L = 4$. This value of L is long enough to compare the BER performances of estimators. On the other hand, the MSE performance for GSM TU channel was evaluated through simulation with $L = 8$.

As in [3], the frequency offset was estimated using a two-step procedure: a coarse search with the pruning factor $K = 8$ followed by a fine search that employs a parabolic interpolation. In the simulation, the MSE values were empirically estimated through 10,000 trials by randomly generating $\nu \in [-0.033, -0.033]$.

The BER was measured by sending 19 random data symbols(that is, 38 bits) right after the 19 symbol training sequence. The parameter values estimated from the training sequence were used in detection process of the transmitted 19 data symbols. Given the estimate of each parameter, the distortion of the signal due to I/Q mismatch, DC offset and frequency offset can be simply compensated by a canceller given by $\mathbf{y}_c = \Gamma^H(\hat{\nu})\left(\mathbf{y} - \hat{\alpha}\mathbf{y}^* - \hat{d}\mathbf{1}\right)$. With the compensated signal and the channel estimate $\hat{\mathbf{g}}$, maximum-likelihood sequence detector (MLSD) was used as an example implementation of the detector.

B. Verification of the analytical results

The analytical results of the MSEs are compared with the simulation results assuming the channel parameters are fixed at $\{h_0, h_1, h_2\} = \{2/3, 2/3, 1/3\}$ ($L = 3$). The training sequence $\{a(n)\}$ was one of the midambles of the GSM system[8] of length 16 ($N = 16$), which was represented as in Table I[2].

Fig. 3 shows the MSEs of $\hat{\nu}$, $\hat{\alpha}$, $\hat{d}$ and $\hat{\mathbf{g}}$ at $\nu = 0.1$. This figure compares the empirical MSEs to the analytical MSEs of (23), (24) and (25). It is seen that the empirical MSEs are reasonably close to the analytical results and the Cramér-Rao lower bounds (CRLB) of the estimates (see [7]) as well.

[2] "GSM sequence" indicates TSC#2 for normal burst in [8] which is differentially encoded and then GMSK modulated.

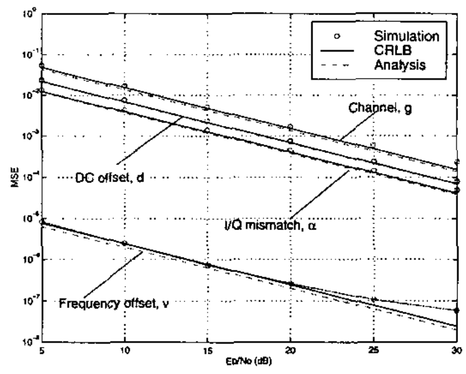

Fig. 3. MSEs versus Eb/No at $\nu = 0.1$.

C. Performance comparison

The performance of the proposed estimator was compared with the MLE#1 in [3] and LS estimator in [4] by using GSM midamble of $N = 16$ in GSM TU channel. Fig. 4 shows that the MSEs of $\hat{\nu}$ and $\hat{\mathbf{g}}$ of MLE#1 exhibit error floor. The crossover between MLE#1 and the proposed estimator occurs at 10 dB with $N = 16$ for both $\hat{\nu}$ and $\hat{d}$. LS method fails to estimate the channel correctly. Fig. 5 shows the MSE of $\hat{\alpha}$ and $\hat{d}$. While LS method fails to estimate correctly, the proposed ML shows MSE monotonically decreasing as Eb/No increases. Fig. 6 shows the BER versus Eb/No. In the presence of I/Q mismatch and DC offset, the proposed ML estimator outperforms the others at Eb/No greater than 6 dB. The Eb/No gain of the proposed estimator with respect to MLE#1 is about 3 dB at 10^{-3} of BER level. The performance loss compared to ideal receiver at 10^{-3} of BER level is lowered to around 1 dB.

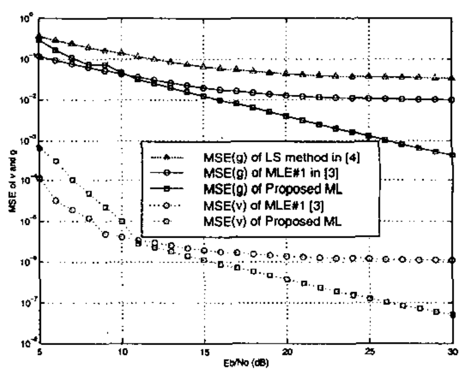

Fig. 4. Performance of frequency and channel estimates.

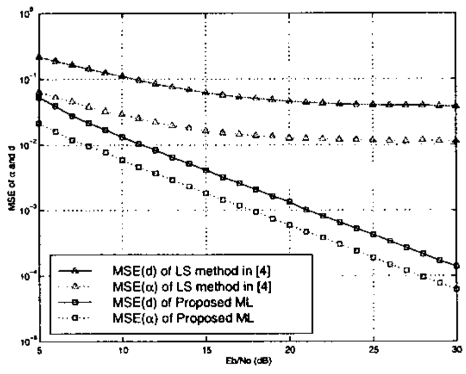

Fig. 5. Performance of I/Q mismatch and DC offset estimates.

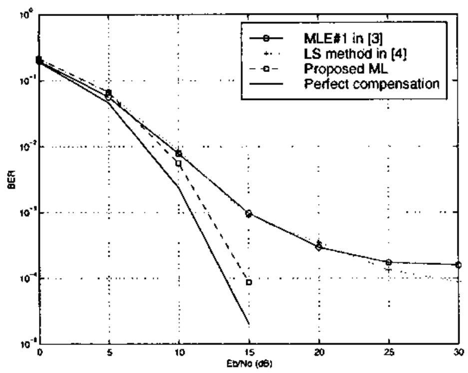

Fig. 6. Comparison of BER performance with GSM sequence.

D. Effectiveness of training sequence selection method

Here, the performance of an optimal sequence is compared with the GSM midamble in GSM TU channel. Assuming no a priori knowledge on the value of C_{gg} in (33), we design an optimal sequence for standard white Gaussian channel (i.e., $\mathbf{C_{gg}} = \mathbf{I}_L$). Table I contains an optimal sequence for $N =$

TABLE I

SOME TRAINING SEQUENCES FOR $N = 16$, $L = 4$

sequence	$\{a(n), n = 0, \cdots, N-1\}$
GSM	1, -j, 1, j, 1, -j, -1, -j, -1, j, -1, -j, -1, j, -1, -j
optimal	1, j, -1, j, -1, -j, 1, -j, -1, j, -1, j, 1, j, -1, j

$16, L = 4$ found. BER of each Training sequence is plotted in Fig. 7. The performance gain of the optimal sequence over GSM sequence was about 0.7 dB at BER of 10^{-3}. The optimal sequence lowers Eb/No loss due to I/Q mismatch and DC offset to 0.3 dB from about 1 dB with GSM sequence, at 10^{-3} level of BER.

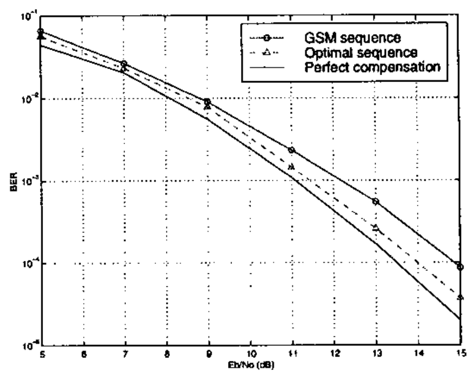

Fig. 7. BER performances of two sequences.

VII. CONCLUSION

A new baseband signal processing method for jointly estimating frequency offset, channel, I/Q mismatch and DC offset in a frequency-selective quasi-static channel was proposed. Because the proposed method estimates the frequency, channel, I/Q mismatch and DC offset jointly, it should be suitable for fast acquisition in the initial stage of communication. Simulation results indicate that the MSE of the proposed estimates approaches the CRLB, and their MSEs can be reduced by employing an optimal sequence. The performance loss due to I/Q mismatch and DC offset at 10^{-3} of BER could be lowered to around 1 dB by use of the proposed estimation method. In addition, the optimal sequence designed by the proposed design rule improved the performance of the proposed estimator by about 0.7 dB.

REFERENCES

[1] A. A. Abidi, *"Direct-conversion radio transceivers for digital communications,"* *IEEE Journal of solid-state circuits*, vol. 30, pp. 1399-1410, Dec. 1995.

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

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

[4] 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. Letter*, 2002.

[5] Young-Doo Kim, Jae Kun Lim, Chang-Ho Suh, Eui-Rim Jeong, and Yong H. Lee, "Carrier frequency estimation for transmissions with antenna diversity," *IEEE 55th VTC Spring*, vol. 3, pp. 1569-1573, 2002.

[6] Young-Doo Kim, Jae Kun Lim, Gate Gil, Chang-Ho Suh and Yong H. Lee, "Traing sequence design for MIMO antenna system," in preparation.

[7] Gate Gil, Y. H. Lee, "Joint ML Estimation fo frequency offset, I/Q mismatch and DC iffset in communication receivers," in preparation.

[8] Digital cellular telecommunications system (Phase 2+), "Multiplexing and multiple access on the radio path," ETSI TS 100 908 (3GPP TS 05.02) V8.9.0 Release 1999(2001-04).