Data-Aided Approach to I/Q Mismatch and DC Offset Compensation in Communication Receivers

Il-Hyun Sohn, *Student Member, IEEE*, Eui-Rim Jeong, and Yong H. Lee, *Senior Member, IEEE*

***Abstract*—A digital signal processing technique for compensating both the I/Q mismatch and the DC offset in communication receivers is derived with an emphasis on direct-conversion architectures. The I/Q mismatch and DC offset are estimated in a least-squares sense using a training sequence. Also, a group of training sequences that minimizes the mean square error of the estimate is determined. The advantages of the proposed technique are demonstrated through computer simulation.**

***Index Terms*—Channel estimation, DC offset, I/Q mismatch.**

I. INTRODUCTION

IN PRACTICAL communication receivers that employ quadrature mixing in the analog front-end, there is always an imbalance, called the I/Q mismatch, between the I and Q branch amplitude and phase. In addition, a DC offset also exists. The I/Q mismatch is a particular concern in direct-conversion [1] and recent low-IF architectures [2], [3]. On the other hand, the DC offset can be easily suppressed in most receiver architectures with the exception of direct-conversion, in which the DC offset is severe. One approach to overcome the I/Q mismatch and DC offset is compensation by digital signal processing (DSP). An off-line technique based on a discrete Fourier transform (DFT) was proposed in [4], while more sophisticated adaptive DSP techniques were introduced in [2], [3], and [5]. Here [2] and [3] propose blind I/Q mismatch compensation techniques that are applicable to most receiver architectures, and [5] proposes a data-aided scheme that can compensate for both the I/Q mismatch and DC offset, although its use is limited to direct-conversion and conventional heterodyne receivers in additive white Gaussian noise (AWGN) channels.

Accordingly, the current paper introduces an alternative DSP technique for DC offset and mismatch compensation. Unlike the schemes in [2], [3], and [5], which are basically adaptive filters, the proposed technique estimates the offset and mismatch values in a least-squares (LS) sense, from a training sequence. It assumes a receiver operating in a frequency-selective channel environment, and jointly estimates the channel parameters. The new technique can be applied to both direct-conversion and conventional heterodyne architectures, yet not to those with a low-IF in [2] and [3].

Manuscript received July 11, 2002. The associate editor coordinating the review of this letter and approving it for publication was Dr. N. Van Stralen.

I. H. Sohn and Y. H. Lee are with the Division of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejon 305-701, Korea (e-mail: yohlee@ee.kaist.ac.kr).

E. R. Jeong is with the Hyundai Syscomm, Inc., Gyunggi-do, Korea.

Digital Object Identifier 10.1109/LCOMM.2002.806451

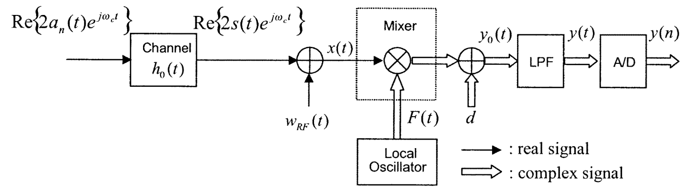


Fig. 1. System employing either direct-conversion or heterodyne architecture.

II. SIGNAL MODEL

The system model considered in the current paper is shown in Fig. 1. Here, $a_n(t)$ is the complex signal to be transmitted; ω_c denotes the carrier frequency; $h_0(t)$ is the impulse response of the frequency selective channel, which is assumed to be quasi-static so that its parameters are constant over a data frame; $w_{RF}(t)$ is zero mean additive white noise in the passband; and d is the DC offset. The passband signal $x(t)$ entering the mixer is represented as $x(t) = s(t)e^{j\omega_c t} + s^*(t)e^{-j\omega_c t} + w_{RF}(t)$, where $s(t)$ is an equivalent lowpass signal at the receiver antenna, given by the convolution of $a_n(t)$ and $h_0(t)$. Note that the receiver in Fig. 1 becomes a direct-conversion receiver when $x(t)$ is an RF signal, and a heterodyne receiver when $x(t)$ is an IF signal. The output of the local oscillator is expressed as $F(t) = \cos(\omega_c t) - j(1+\epsilon)\sin(\omega_c t+\theta) = \beta_0 e^{-j\omega_c t} + \alpha_0 e^{j\omega_c t}$, where ϵ and θ denote the amplitude and phase imbalance, respectively, $\beta_0 = 1/2\{1 + (1+\epsilon)e^{-j\theta}\}$, and $\alpha_0 = 1/2\{1 - (1+\epsilon)e^{j\theta}\}$. Lowpass filtering the signal $y_0(t)$, which is equal to $x(t)F(t) + d$, results in the baseband analog signal $y(t)$. If $y(t)$ is sampled with a symbol rate, this produces $y(n) = \beta_0 s(n) + \alpha_0 s^*(n) + d + \beta_0 w_0(n) + \alpha_0 w_0^*(n)$, where $w_0(n)$ denotes the equivalent band-limited lowpass noise; $s(n) = \sum_{k=0}^{L-1} h_0(k)a_{n-k}$; $h_0(k)$ is the impulse response of the equivalent channel due to an impulse applied k time units earlier; L denotes the channel length; and a_n is a complex information symbol. Suppose that N $(N > L)$ training symbols $\{a_n | n = 0, 1, \ldots, N-1\}$ are transmitted. The received data $\{y(L-1), y(L), \ldots, y(N-1)\}$ that correspond to the training sequence can be written in vector form as

$$\mathbf{y} = \beta_0 \mathbf{A}\mathbf{h}_0 + \alpha_0 \mathbf{A}^* \mathbf{h}_0^* + d\mathbf{1} + \beta_0 \mathbf{w}_0 + \alpha_0 \mathbf{w}_0^* \quad (1)$$

where $\mathbf{y} = [y(L-1), y(L), \ldots, y(N-1)]^T$; $\mathbf{h}_0 = [h_0(0), h_0(1), \ldots, h_0(L-1)]^T$; $\mathbf{1} = [1, 1, \ldots, 1]^T$ is an $(N-L+1)$-dimensional column vector whose elements are all 1; $\mathbf{w}_0 = [w_0(L-1), w_0(L), \ldots, w_0(N-1)]^T$; and $\mathbf{A}$ is an $(N-L+1)$-by-L data matrix with entries $[\mathbf{A}]_{i,j} = a_{i-j+L-1}$, $0 \le i \le N-L$, $0 \le j \le L-1$. To simplify the notation, let $\mathbf{h} = \beta_0 \mathbf{h}_0$ and $\mathbf{w} = \beta_0 \mathbf{w}_0$. Then, $\mathbf{y}$ in (1) can be rewritten as

$$\mathbf{y} = \mathbf{A}\mathbf{h} + \alpha \mathbf{A}^* \mathbf{h}^* + d\mathbf{1} + \mathbf{w} + \alpha \mathbf{w}^* \quad (2)$$

where $\alpha = \alpha_0/\beta_0^*$. In (2), $\alpha\mathbf{A}^*\mathbf{h}^*$ denotes the image signal caused by the I/Q mismatch, and $d\mathbf{1}$ represents the DC offset. The constant α is called the mismatch parameter.

III. Proposed Compensation Technique

A. Derivation

Suppose, for the time being, the parameters α and d in (2) are known. A direct calculation using (2) yields

$$(\mathbf{y} - d\mathbf{1}) - \alpha(\mathbf{y} - d\mathbf{1})^* = (1- \mid \alpha \mid^2)(\mathbf{Ah} + \mathbf{w}). \quad (3)$$

The right-hand side (RHS) of (3) represents the desired signal vector multiplied with a known constant, because $\mathbf{Ah} + \mathbf{w}$ 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 (3). The proposed compensation technique consists of two steps: estimation of α and d followed by the correction procedure in (3). Next, an LS estimator is developed for α and d.

In (2), three parameters α, d, and $\mathbf{h}$ need to be estimated. Due to the image vector $\alpha\mathbf{A}^*\mathbf{h}^*$ and offset vector $d\mathbf{1}$, the RHS of (2) is quite different from the standard linear model. In the proposed estimation, the RHS of (2) is linearized first. If the vector $\mathbf{g}$ is defined as

$$\mathbf{g} = [\mathbf{h}^T, \alpha\mathbf{h}^H, d]^T \quad (4)$$

then (2) becomes

$$\mathbf{y} = \mathbf{Bg} + \mathbf{z} \quad (5)$$

where $\mathbf{B} = [\mathbf{A}, \mathbf{A}^*, \mathbf{1}]$ and $\mathbf{z} = \mathbf{w} + \alpha\mathbf{w}^*$. Note that the RHS of (5) has the standard linear form. The LS estimate of $\mathbf{g}$, which minimizes the cost $(\mathbf{y} - \mathbf{B}\hat{\mathbf{g}})^H(\mathbf{y} - \mathbf{B}\hat{\mathbf{g}})$, is given by

$$\hat{\mathbf{g}} = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H\mathbf{y}. \quad (6)$$

After obtaining $\hat{\mathbf{g}}$, $\mathbf{h}$ and d are estimated by comparing the RHS of (4) with $\hat{\mathbf{g}}$. Specifically, if $\hat{\mathbf{g}} = [\hat{g}(0), \hat{g}(1), \ldots, \hat{g}(2L)]^T$, then $\hat{\mathbf{h}} = [\hat{g}(0), \hat{g}(1), \ldots, \hat{g}(L-1)]^T$ and $\hat{d} = \hat{g}(2L)$. An estimate of the mismatch parameter α is obtained by fitting $\alpha\hat{\mathbf{h}}^*$ to $\hat{\mathbf{g}}_M$, where $\hat{\mathbf{g}}_M = [\hat{g}(L), \hat{g}(L+1), \ldots, \hat{g}(2L-1)]^T$. The estimate $\hat{\alpha}$ is given by

$$\hat{\alpha} = \arg\min_{\alpha} \left\{ (\hat{\mathbf{g}}_M - \alpha\hat{\mathbf{h}}^*)^H(\hat{\mathbf{g}}_M - \alpha\hat{\mathbf{h}}^*) \right\} \quad (7)$$

$$= \frac{(\hat{\mathbf{h}}^T\hat{\mathbf{g}}_M)}{(\hat{\mathbf{h}}^T\hat{\mathbf{h}}^*)}. \quad (8)$$

B. Properties of Estimate

It is assumed that $\mathbf{w} = \beta_0\mathbf{w}_0$ in (2) is a random vector with a zero-mean and diagonal covariance matrix $E[\mathbf{ww}^H] = \sigma_w^2\mathbf{I}$. This assumption is justified because $\mathbf{w}_0$ and β_0 are independent. Under this assumption, the estimate $\hat{\mathbf{g}}$ is unbiased [6]. This fact indicates that $\hat{\mathbf{h}}$ and $\hat{d}$ are also unbiased. For $\hat{\alpha}$ in (8), it is difficult to check the unbiasedness, therefore, this will be examined via simulation in Section IV.

The mean square error (MSE) of $\hat{\mathbf{g}}$ is: $\sigma_e^2 = E\{(\hat{\mathbf{g}} - \mathbf{g})^H(\hat{\mathbf{g}} - \mathbf{g})\} = E\{\text{tr}((\hat{\mathbf{g}} - \mathbf{g})(\hat{\mathbf{g}} - \mathbf{g})^H)\} = \text{tr}\{(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H E\{\mathbf{zz}^H\}\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\}$, where $\text{tr}\{\cdot\}$ denotes the trace of a matrix. Following the assumption, it is straightforward to show that $E\{\mathbf{zz}^H\} = \sigma_z^2\mathbf{I}$, where $\sigma_z^2 = (1 + |\alpha|^2)\sigma_w^2$. Then, the MSE of $\hat{\mathbf{g}}$ becomes $\sigma_e^2 = \sigma_z^2\text{tr}\{(\mathbf{B}^H\mathbf{B})^{-1}\}$. This MSE represents the sum of the MSEs associated with $\hat{\mathbf{h}}$, $\hat{\alpha}\hat{\mathbf{h}}^*$, and $\hat{d}$.

TABLE I
Training Sequences

Item	Preamble 1	Preamble 2
Training sequence	$\{-1, -1, 1, 1, 1, 1, 1, -1, j, -j, j, -1, -j, -1, -1, 1, -1, -1\}$	$\{-j, 1, 1, j, -j, 1, -1, j, -j, -1, -1, j, j, 1, -1, -j, -j, 1\}$
Properties	$\mathbf{A}^H\mathbf{A} = 16\mathbf{I}$, $\mathbf{A}^T\mathbf{A} \neq \mathbf{0}$, $\mathbf{A}^T\mathbf{1} = \mathbf{0}$	$\mathbf{A}^H\mathbf{A} = 16\mathbf{I}$, $\mathbf{A}^T\mathbf{A} = \mathbf{0}$, $\mathbf{A}^T\mathbf{1} = \mathbf{0}$

C. Training Sequence Design

This subsection identifies the class of training sequences that minimizes σ_e^2.

Property 1: Suppose that the training symbols are PSK. Then the MSE σ_e^2 is minimized if $\mathbf{A}$ in (1) satisfies the following:

$$\mathbf{A}^H\mathbf{A} = c\mathbf{I} \quad (9)$$

$$\mathbf{A}^T\mathbf{A} = \mathbf{0} \quad (10)$$

$$\mathbf{A}^T\mathbf{1} = \mathbf{0} \quad (11)$$

where c is a scalar constant.

Proof: Following the approach in [7, p. 788], it is possible to show that σ_e^2 is minimized if $\mathbf{B}$ satisfies $\mathbf{B}^H\mathbf{B} = c\mathbf{I}$. Using $\mathbf{B} = [\mathbf{A}, \mathbf{A}^*, \mathbf{1}]$ in this equation, (9) – (11) are obtained. ■

Since the conditions in (9) and (10) are necessary for minimizing the MSEs of $\hat{\mathbf{h}}$ and $\hat{\alpha}\hat{\mathbf{h}}^*$, these conditions are also useful for estimating $\mathbf{h}$ and α (in [7], (9) has been derived for the channel estimation). The condition in (11) minimizes the MSE of $\hat{d}$.

IV. Simulation Results

The parameters used in the simulation were as follows: channel length $L = 3$; quasi-static channel $\mathbf{h}$ was composed of independent identically distributed complex Gaussian random variables with a zero mean and variance $1/L$; magnitude mismatch $\epsilon = 0.1$; phase mismatch $\theta = 10^\circ$; and DC offset $d = 0.1 + j0.1$. In this case, $\alpha = -0.048 - j0.0873$. Two kinds of training sequences of length $N = 18$ were considered, as shown in Table I and referred to as Preambles 1 and 2. Preamble 2 was a desired training sequence that satisfied all the conditions in Property 1, whereas Preamble 1 met the conditions in (9) and (11), yet violated (10). The results presented in this section are the average of 100 000 trials.

Fig. 2 shows the mean estimation error of $\hat{\alpha}$. The estimate $\hat{\alpha}$ exhibited some bias at a low signal-to-noise-ratio (SNR), however, the bias was small compared with α. Note that Preamble 2 yielded a smaller bias than Preamble 1. For a high SNR, $\hat{\alpha}$ was almost unbiased.

Fig. 3 shows the MSEs of $\hat{\mathbf{h}}$, $\hat{\alpha}$, and $\hat{d}$. In the case of $\hat{\mathbf{h}}$ and $\hat{\alpha}$, Preamble 2 yielded smaller MSEs than Preamble 1, because the

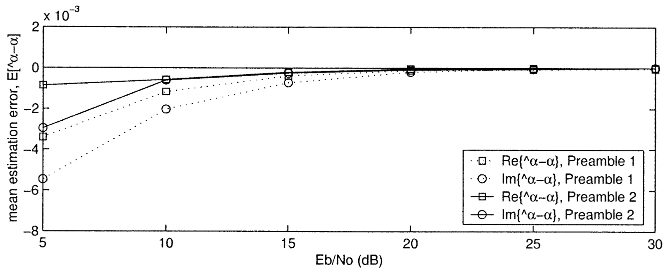
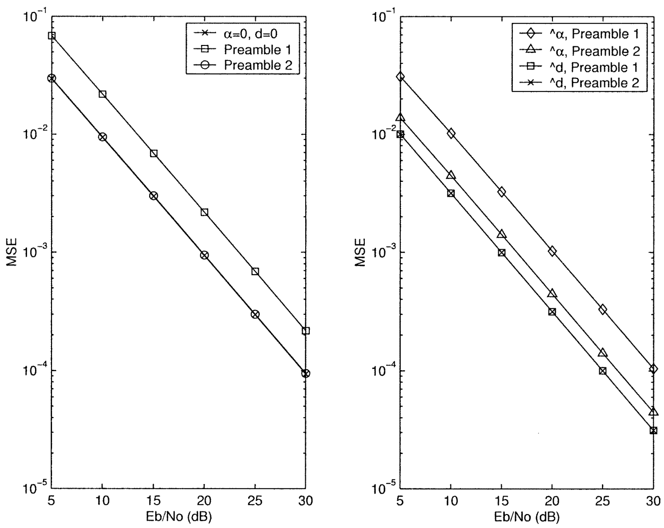
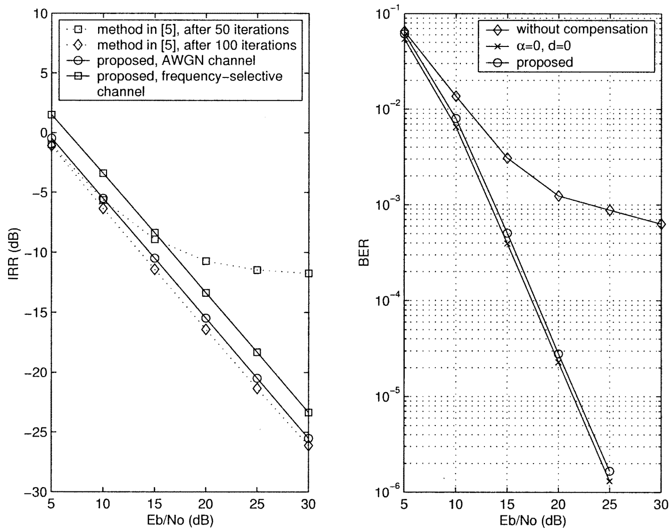


Fig. 2. Mean estimation error of $\hat{\alpha}$.



(a) (b)

Fig. 3. (a) MSE of channel estimate $\hat{\mathbf{h}}$. (b) MSE of mismatch parameter estimate $\hat{\alpha}$ and DC offset estimate $\hat{d}$.



(a) (b)

Fig. 4. (a) Comparison of IRRs. (b) Comparison of BERs.

latter violated (10). For $\hat{d}$, the two preambles exhibited identical performances, because both of them satisfied (11). Fig. 3(a) also shows the MSEs of $\hat{\mathbf{h}}$ when $\alpha = 0$ and $d = 0$. The MSEs were identical to those associated with Preamble 2, and the performance of the proposed estimator equipped with Preamble 2 was not degraded by the mismatch and DC offset.

The behavior of the proposed compensator was examined by evaluating the image rejection ratio (IRR) [8], defined by $10\log_{10}(\gamma_{out}/\gamma_{in})$, where γ_{out} and γ_{in} denote the image-to-signal power ratio at the output and input of the compensator, respectively. For comparison, the adaptive technique in [5] was also considered, under the assumption of an AWGN channel: this scheme employs a least-mean-square (LMS) algorithm for adaptation and resembles a 3-tap equalizer. The step-size for the LMS adaptation was 0.15. Fig. 4(a) compares the IRRs. The results produced by the proposed compensator, which estimated the parameters from Preamble 2 consisting of 18 training symbols, were almost comparable to those with the adaptive technique using 100 iterations (i.e., 100 training symbols). This occurred because the LMS algorithm had a relatively slow rate of convergence. Fig. 4(a) also shows the IRRs for the proposed method when the channel was frequency-selective, as in the case of Fig. 3. The IRRs remained almost the same, irrespective of the channel. This indicates that the proposed compensator was able to behave well in any channel environment.

Finally, the degradation of the bit error rate (BER) was examined for a QPSK system employing maximum-likelihood (ML) sequence detection. The frequency-selective channel introduced at the beginning of this section was considered, along with Preamble 2. Fig. 4(b) compares BERs of the receivers with/without mismatch and DC offset compensation (the ML channel estimate in [7] was employed for the "without-compensation" case). When the compensation was not performed, the receiver exhibited an error-floor about a BER of 6×10^{-4}. This error-floor disappeared after the proposed compensation, and the receiver performance became close to the case with $\alpha = 0$ and $d = 0$.

V. Conclusion

A new I/Q mismatch and DC offset compensation technique, based on LS estimation, was proposed, and a group of training sequences that minimizes the MSE of the estimator determined. Simulation results indicated that the proposed method performed well in a frequency-selective channel environment.

References

[1] S. Mirabbasi and K. Martin, "Classical and modern receiver architectures," *IEEE Commun. Mag.*, vol. 38, no. 11, pp. 132–139, Nov. 2000.
[2] L. Yu and W. M. Snelgrove, "A novel adaptive mismatch cancellation system for quadrature IF radio receivers," *IEEE Trans. Circuits Syst. II*, vol. 46, pp. 789–801, June 1999.
[3] 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.
[4] F. E. Churchill, G. W. Ogar, and B. J. Thompson, "The correction of I and Q errors in a coherent processor," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, pp. 131–137, Jan. 1981.
[5] J. K. Cavers and M. W. Liao, "Adaptive compensation for imbalance and offset losses in direct conversion tranceivers," *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 581–588, Nov. 1993.
[6] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1996.
[7] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers*. New York: Wiley, 1998.
[8] B. Razavi, *RF Microelectronics*. Englewood Cliffs, NJ: Prentice Hall, 1998.