Joint Calibration of Transmitter and Receiver Impairments in Direct-Conversion Radio Architecture

Chen-Jui Hsu and Wern-Ho Sheen, *Member, IEEE*

***Abstract*—Direct-conversion radio architecture is a low-cost, low-power and small-size design that has been widely employed in today's wireless devices. This architecture, however, induces radio impairments such as I-Q imbalance and dc offset that may incur severe degradation in communication performance if left uncompensated. In this paper, a new method is proposed to calibrate simultaneously a transceiver's own transmitter and receiver radio impairments with no dedicated analog circuit in the feedback loop. Based on a unified time-domain approach, the proposed method is able to calibrate jointly the frequency-independent I-Q imbalance, frequency-dependent I-Q imbalance and dc offset and is applicable to any type of communication systems (single-carrier, multiple-carrier, etc.). The existing methods in the literature either need a dedicated analog circuit in the feedback loop and/or are applicable only to a particular type of systems with some radio impairments present. The issue of training sequence design is also investigated to optimize the calibration performance, and analytical and simulation results show that the performance loss due to radio impairments can be recovered by the proposed method.**

***Index Terms*—Direct-conversion transceiver, self-calibration, I-Q imbalance, DC offset.**

I. Introduction

DIRECT-CONVERSION radio architecture is a low-cost, low power and small-size design that has gained popularity in today's wireless devices [1][2]. The radio impairments of this architecture such as I-Q imbalance and dc offset, however, result in a severe degradation in communication performance if left uncompensated [3]-[9]. This is particularly true in the next-generation high data-rate systems where a wide bandwidth and a high-order modulation are deemed to be employed. Removal and/or compensation of the radio impairments in the direct-conversion radio architecture have been an area of extensive research. Generally speaking, two types of techniques have been proposed [3]-[19]: one is calibration and the other is estimation/compensation. Calibration is a technique used to remove the effects of a transceiver's own radio impairments [9]-[19], whereas the estimation/compensation technique is to counteract the cascaded transmitter and receiver radio impairments at the receiving side [3]-[9]. Both types of techniques find their applications in real systems [3]-[19]. In this work, we focus on the digital calibration technique.

Many interesting works have been devoted to the calibration of the radio impairments in the direct-conversion architecture [9]-[19]. In [9]-[11], adaptive methods were proposed to calibrate transmitter frequency-independent I-Q imbalance and dc offset by using an analog envelope-detector (ED) in the feedback loop. Transmitter frequency-dependent I-Q imbalance was investigated for the continuous frequency-shift-keying (CFSK) systems in [12][13], with no consideration on other impairments; frequency-dependent I-Q imbalance is particularly problematic in a wideband system where it is very challenging to keep the I- and Q-branch analog filters perfectly matched over the entire band. The works in [14]-[16] discussed calibration of the transmitter frequency-independent and dependent I-Q imbalances by using the low-IF architecture in the feedback loop.

So far, most of the calibration techniques in the literature have focused on the transmitter radio impairments, either by employing an ED [9]-[11] or the low-IF radio architecture [14]-[16] in the feedback loop so that the receiver impairments can be neglected safely. However, using an ED or the low-IF radio architecture in the feedback loop increases the complexity in the analog domain. In addition, calibration of the receiver impairments is important in its own right; for example, if one's own receiver has been calibrated, only the transmitter-side impairments (of the transmitting device), rather than the cascaded transmitter and receiver impairments, need to be estimated and compensated for at the receiver, and that reduces the receiver complexity [4]-[8].

Very recently, the issue of joint calibration of a transceiver's own transmitter and receiver radio impairments in the direct-conversion radio architecture was investigated in [17]-[19]. In [17], a two-feedback method was proposed, where the phase of the receiver oscillator is shifted by exact 90 degrees in the second feedback, aiming to separate the transmitter and receiver I-Q imbalances. Unfortunately, it is not practical to have an exact 90 degrees phase rotation in real systems. In [18], a new calibration method was proposed for the OFDM (orthogonal frequency-division multiplexing) type of systems with no dedicated analog circuit in the feedback loop. The method, however, is designed for the OFDM-type of systems and can only calibrate the frequency-independent I-Q imbalance and dc offset.

Manuscript received May 18, 2011; revised August 31, 2011; accepted September 27, 2011. The associate editor coordinating the review of this paper and approving it for publication was G. Colavolpe.

C. J. Hsu is with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 300, Taiwan, R.O.C.

W. H. Sheen is with the Department of Information and Communication Engineering, Chaoyang University of Technology, Wufong, Taichung41349, Taiwan (e-mail: linch.ee89@nctu.edu.tw, whsheen@cyut.edu.tw).

Digital Object Identifier 10.1109/TWC.2011.110811.110939

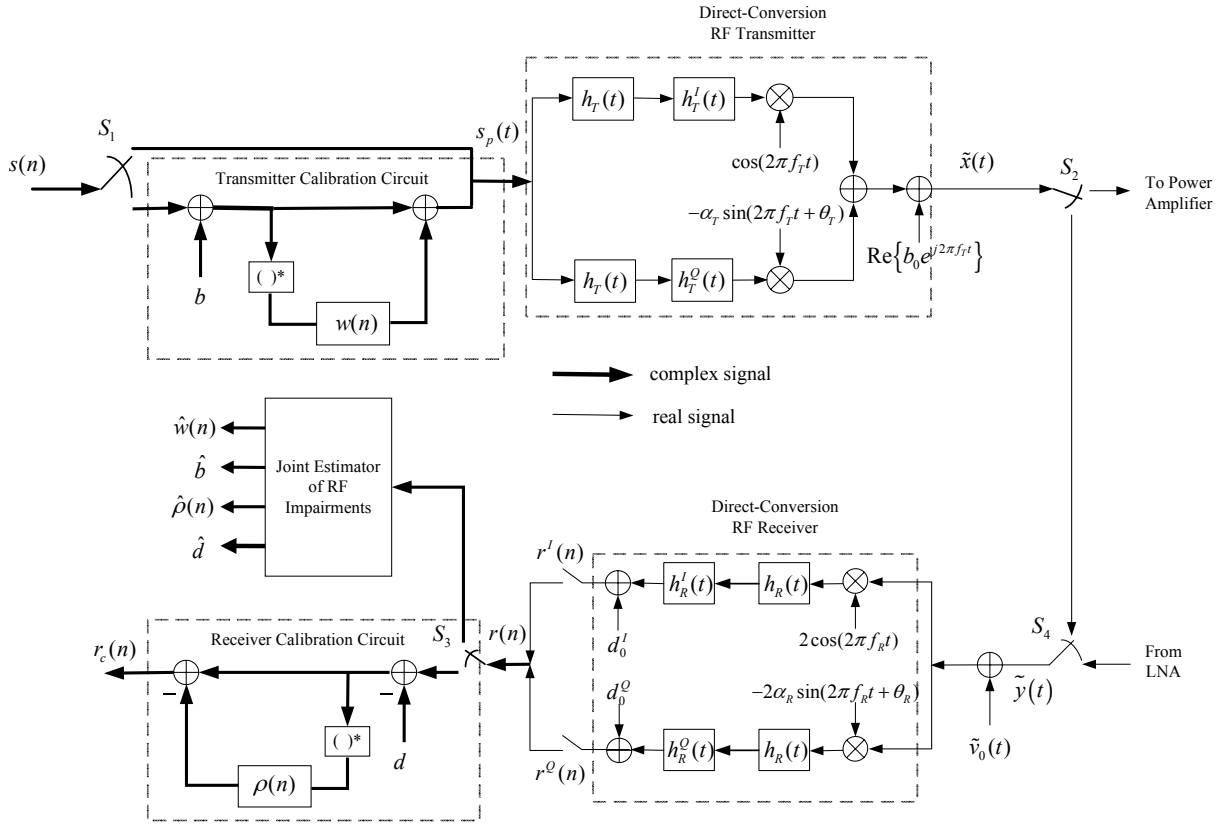

Fig. 1. The system consisting of a direct-conversion RF transceiver, calibration circuits and a joint estimator of the calibration parameters.

In this paper, a new method is proposed to calibrate a transceiver's own transmitter and receiver impairments simultaneously, with no dedicated analog circuit in the feedback loop. Compared to the work in [18], the proposed method is unique in that it is applicable to any type of communication systems and is able to calibrate jointly the frequency-independent I-Q imbalance, frequency-dependent I-Q imbalance, and dc offset, thanks to the proposed unified time-domain approach. This work is an extension of our work in [19], where parts of the results in this paper were firstly reported.

The rest of this paper is organized as follows. Section II describes the models of radio impairments and digital calibration circuits. Section III develops a joint estimator under the principle of nonlinear least-squares. The proposed method is then analyzed in Section IV, and numerical results and conclusions are given in Section V and VI, respectively.

II. Radio Impairments And Calibration Circuits

Figure 1 depicts the considered system that consists of calibration circuits, the direct-conversion radio transceiver, and a joint estimator of the transmitter and receiver radio impairments. The switches S_1, S_2, S_3 and S_4 are used to control the signal flow paths; during calibration training, S_1 , S_3 and S_4 are at the upper positions, and S_2 is at the lower position to form an internal loopback, whereas during normal communication, S_1, S_3 and S_4 are at the lower positions, and S_2 is at the upper position. Since the transceiver's own receiver is used for internal loopback during calibration training, a dedicated calibration period is assumed in this method prior to normal communication where transmitter and receiver are located in different devices.

A. Radio Impairments

At the transmitter, the radio impairments that are investigated include frequency-dependent I-Q imbalance, frequency-independent I-Q imbalance and carrier feedthrough. Frequency-dependent I-Q imbalance is due to mismatch between the in-phase (I) and quadrature-phase (Q) analog filters which are denoted by $h_T(t) \otimes h_T^I(t)$ and $h_T(t) \otimes h_T^Q(t)$. Here, $h_T(t)$ is the common part of the filters, and $\otimes$ denotes the operation of linear convolution. Frequency-independent I-Q imbalances are due to the gain and phase mismatches between the I and Q branches of the mixer circuitry and denoted by α_T and θ_T, respectively. And, the carrier feed-through induces dc offset at the receiving side [1][2] and is characterized by $\mathrm{Re}\left\{b_0 e^{j2\pi f_T t}\right\}$, where $b_0 = b_0^I + jb_0^Q$, f_T is the transmit center frequency, and $j = \sqrt{-1}$. In the following, b_0 is called the transmitter dc offset.

Define

$$s_p(t) \doteq s_p^I(t) + js_p^Q(t) = \sum_n s_p(n)\,\delta(t - nT_s), \quad (1)$$

be the signal appearing at the input of the transmitter, where $s_p(n)$ is its discrete-time equivalent, T_s is the symbol duration, and $\delta(t)$ is the Dirac delta function. Under the effects of the radio impairments, the pass-band transmit signal is $\tilde{x}(t) =$

$\mathrm{Re}\left\{x(t)e^{j2\pi f_T t}\right\}$ with its base-band equivalent given by [15]

$$x(t) = h_{T,+}(t) \otimes s_p(t) + h_{T,-}(t) \otimes s_p^*(t) + b_0, \quad (2)$$

where

$$h_{T,\pm}(t) = 1/2 \cdot \left[h_T^I(t) \pm \alpha_T e^{j\theta_T} h_T^Q(t)\right] \otimes h_T(t), \quad (3)$$

and a^* denote the complex conjugate of a. In (2), $s_p(t)$ can be viewed as being transmitted through two paths along with a corruption from dc offset; one is the desired path with impulse response $h_{T,+}(t)$, and the other is the mirror-frequency path with impulse response $h_{T,-}(t)$. Clearly, I-Q imbalances incur mirror-frequency interference in the transmitted signal. With no I-Q imbalances and dc offset, i.e., $h_T^I(t) = h_T^Q(t) = \delta(t)$, $\alpha_T = 1$ and $\theta_T = b_0 = 0$, $x(t) = s_p(t) \otimes h_T(t)$ as one might expect.

Likewise, the radio impairments that are investigated at the receiver include frequency-independent I-Q imbalance, characterized by α_R and θ_R, frequency-dependent I-Q imbalance, characterized by the filters $h_R(t) \otimes h_R^I(t)$ and $h_R(t) \otimes h_R^Q(t)$, and dc offset, $d_0 = d_0^I + jd_0^Q$. $h_R(t)$ is the common part of the filters, and $f_R = f_T - \Delta f$ is the receive center frequency. During normal communication, Δf is a real frequency offset between transmitter and receiver which are located in different devices, whereas during calibration training, Δf is an intentional frequency shift[1] that is introduced purposely in our method to help the radio impairments estimation, as is to be detailed in Section III.

Denote $\tilde{y}(t)$ and $\tilde{v}_0(t)$ be the received pass-band signal and additive white Gaussian noise in Figure 1. Under the effects of radio impairments, the received base-band signal is given by [8]

$$\begin{aligned} r(t) = {} & h_{R,+}(t) \otimes \left[e^{j2\pi\Delta f t} y(t) + v_0(t)\right] \\ & + h_{R,-}(t) \otimes \left[e^{j2\pi\Delta f t} y(t) + v_0(t)\right]^* + d_0, \end{aligned} \quad (4)$$

where $y(t)$ and $v_0(t)$ are the low-pass equivalents of $\tilde{y}(t)$ and $\tilde{v}_0(t)$, respectively, and

$$h_{R,\pm}(t) = 1/2 \cdot \left[h_R^I(t) \pm \alpha_R e^{\mp j\theta_R} h_R^Q(t)\right] \otimes h_R(t). \quad (5)$$

Again, (4) says that the receiver I-Q imbalances induce mirror-frequency interference, and in the absence of I-Q imbalance and dc offset, $r(t) = h_R(t) \otimes \left[e^{j2\pi\Delta f t} y(t) + v_0(t)\right]$. It is worthy to note that during calibration training, $y(t) = x(t)$ which is the signal transmitted from its own transmitter because of the internal loopback. During normal communication, on the other hand, $y(t) = x(t) \otimes c(t)$, where $x(t)$ is the signal transmitted from the transmitter in other device, and $c(t)$ is the channel impulse response.

B. Calibration Circuits

At the transmitter, we propose to use a pre-distortion filter, $w(n)$, and a dc correction term, b, to calibrate I-Q imbalances and dc offset, respectively as in Figure 1[2]. After calibration, $s_p(n)$ is given by

$$s_p(n) = [s(n) + b] + w(n) \otimes [s(n) + b]^*, \quad (6)$$

where $s(n)$ is the transmitted symbol. For convenience, the equivalent discrete-time model will be used throughout the rest of the paper. In this way, (2) is rewritten as

$$x(n) = g_{T,+}(n) \otimes s(n) + g_{T,-}(n) \otimes s^*(n) + \Delta b, \quad (7)$$

where $u(n) = u(t)|_{t=nT_s}$ $u \in \{x, h_{T,+}, h_{T,-}\}$,

$$g_{T,+}(n) = h_{T,+}(n) + w^*(n) \otimes h_{T,-}(n), \quad (8)$$

$$g_{T,-}(n) = h_{T,-}(n) + w(n) \otimes h_{T,+}(n), \quad (9)$$

and

$$\Delta b = g_{T,+}(n) \otimes b + g_{T,-}(n) \otimes b^* + b_0. \quad (10)$$

In (8)-(10), $g_{T,+}(n)$ is regarded as the overall impulse response of the desired path after calibration, $g_{T,-}(n)$ is that of the mirror-frequency path, and Δb is the residual dc offset. Ideally, $g_{T,-}(n) = 0$ and $\Delta b = 0$ which lead to

$$w_{opt}(n) = -\left(h_{T,+}(n)\right)^\dagger \otimes h_{T,-}(n), \quad (11)$$

$$g_{T,+,opt}(n) = h_{T,+}(n) + w_{opt}^*(n) \otimes h_{T,-}(n), \quad (12)$$

and

$$b_{opt} = -b_0 \otimes \left(g_{T,+,opt}(n)\right)^\dagger, \quad (13)$$

where the notation $(h(n))^\dagger$ is to denote the inverse filter of $h(n)$. As is expected, $w_{opt}(n) = b_{opt} = 0$ for the case of no I-Q imbalances and dc offsets. The estimation of $w_{opt}(n)$ and b_{opt} is done during calibration training and used during normal communication.

As in [15][16], the image-rejection-ratio (IRR) will be adopted as the performance measure for the I-Q imbalance calibration, which is defined as

$$IRR_T(f) = 10\log_{10}\frac{|G_{T,+}(f)|^2}{|G_{T,-}(f)|^2}\mathrm{dB}, \quad (14)$$

where $U(f) \doteq \mathrm{FT}[u(n)]$ is the Fourier transform (FT) of $u(n)$. In addition, the ratio

$$\varepsilon_T = 10\log_{10}\frac{|\Delta b|^2}{|b_0|^2}\mathrm{dB} \quad (15)$$

will be adopted as the performance measure for the dc offset calibration.

At the receiver, a time-domain calibration filter, $\rho(n)$, is employed to remove the receiver mirror-frequency interference, and a dc correction term, d, is used to remove the dc offset (see Figure 1). Thus, the received signal after calibration is given by

$$\begin{aligned} r_c(n) &= (r(n) - d) - \rho(n) \otimes (r(n) - d)^* \\ &= g_{R,+}(n) \otimes \left[e^{j2\pi\mu n} y(n) + v_0(n)\right] \\ &\quad + g_{R,-}(n) \otimes \left[e^{j2\pi\mu n} y(n) + v_0(n)\right]^* + \Delta d, \end{aligned} \quad (16)$$

where

$$\begin{aligned} r(n) = {} & h_{R,+}(n) \otimes \left[e^{j2\pi\mu n} y(n) + v_0(n)\right] \\ & + h_{R,-}(n) \otimes \left[e^{j2\pi\mu n} y(n) + v_0(n)\right]^* + d_0, \end{aligned} \quad (17)$$

$\mu = \Delta f T_s$ is the normalized frequency offset,

$$g_{R,\pm}(n) = h_{R,\pm}(n) - \rho(n) \otimes h_{R,\mp}^*(n), \quad (18)$$

[1] In practice, the intentional frequency shift can be implemented precisely with two digital frequency synthesizers from a single reference oscillator [20].

[2] The idea of using a pre-distortion filter to remove the mirror-frequency interference was also reported in [15].

$$\Delta d = (d_0 - d) - \rho(n) \otimes (d_0 - d)^*, \tag{19}$$

and $\{v_0(n)\}$ are i.i.d. (independent, and identically distributed) zero-mean Gaussian noise with variance of σ_0^2. In (16), $y(n) = x(n) \otimes c(n)$ with $x(n)$ being the signal transmitted from the other device (normal communication), and μ is the real frequency offset between the transmitter and receiver which are located in different devices. Clearly, it is desirable to have $g_{R,-}(n) = 0$ and $\Delta d = 0$; or equivalently,

$$\rho_{opt}(n) = \left(h_{R+}^*(n)\right)^\dagger \otimes h_{R-}(n), \tag{20}$$

and

$$d_{opt} = d_0. \tag{21}$$

Similar to the transmitter case, the receiver calibration performance is evaluated by

$$IRR_R(f) = 10\log_{10}\frac{|G_{R,+}(f)|^2}{|G_{R,-}(f)|^2}\text{dB} \tag{22}$$

and

$$\varepsilon_R = 10\log_{10}\frac{|\Delta d|^2}{|d_0|^2}\text{dB}, \tag{23}$$

for the I-Q imbalance and dc offset, respectively. Also, $\rho_{opt}(n) = d_{opt} = 0$ for the case of no receiver I-Q imbalances and dc offset. $\rho_{opt}(n)$ and d_{opt} are estimated during calibration training. With perfect calibration in the transmitter and receiver, $r_c(n)$ in (16) is given by,

$$r_c(n) = g_{R,+}(n) \otimes \left[e^{j2\pi\mu n}\left(g_{T,+}(n) \otimes s(n) \otimes c(n)\right) + v_0(n)\right]. \tag{24}$$

In this case, the frequency offset μ and channel $c(n)$ are the radio parameters left to be estimated and compensated at the receiver before making a detection on $s(n)$. Quite a lot of methods have been proposed for channel and frequency-offset estimation, for example in [4]-[5],[7]-[8] and references therein.

III. Joint Estimation Of Calibration Parameters

In this section, a joint estimation of the calibration parameters, $w_{opt}(n)$, b_{opt}, $\rho_{opt}(n)$, and d_{opt} is developed first, followed by the design of training sequence and frequency shift μ to optimize the calibration performance. Recall that the estimation is done during calibration training (internal loopback), where $s_p(n) = s(n)$ is the training sequence, $y(t) = x(t)$ with $x(t)$ being the signal transmitted from its own transmitter, and μ is the intentional frequency shift introduced to help estimation of the calibration parameters.

A. Non-linear Least-Squares Estimation

Using (7), the received signal in (17) can be rewritten as

$$\begin{aligned} r(n) = {} & e^{j2\pi\mu n}\left[f_{1,+}(n) \otimes s(n) + f_{1,-}(n) \otimes s^*(n) + b_1\right] \\ & + e^{-j2\pi\mu n}\left[f_{2,+}(n) \otimes s(n) + f_{2,-}(n) \otimes s^*(n) + b_2\right] \\ & + d_0 + v(n), \end{aligned} \tag{25}$$

where

$$f_{1,\pm}(n) = \left(h_{R,+}(n)\, e^{-j2\pi\mu n}\right) \otimes h_{T,\pm}(n), \tag{26}$$

$$f_{2,\pm}(n) = \left(h_{R,-}(n)\, e^{j2\pi\mu n}\right) \otimes h_{T,\mp}^*(n), \tag{27}$$

$$b_1 = \left(h_{R,+}(n)\, e^{-j2\pi\mu n}\right) \otimes b_0, \tag{28}$$

$$b_2 = \left(h_{R,-}(n)\, e^{j2\pi\mu n}\right) \otimes b_0^*, \tag{29}$$

and

$$v(n) = h_{R,+}(n) \otimes v_0(n) + h_{R,-}(n) \otimes v_0^*(n). \tag{30}$$

Our goal here is to estimate $w_{opt}(n)$, b_{opt}, $\rho_{opt}(n)$ and d_{opt} from $r(n)$, given the training sequence $s(n)$ and the intentional frequency shift μ. Obviously, one possible way to do this is to estimate $h_{T,\pm}(n)$, $h_{R,\pm}(n)$, b_0 and d_0 directly from (25)-(30) and then apply them to (11)-(13), (20) and (21). Direct estimation of $h_{T,\pm}(n)$, $h_{R,\pm}(n)$, b_0 and d_0, however, is very complex as can be seen from (25)-(29). Instead, a simpler method is proposed here based on the following observations:

$$\begin{aligned} w_{opt}(n) &= -\left(h_{T,+}(n)\right)^\dagger \otimes h_{t,-}(n) \\ &= -\left(f_{1,+}(n)\right)^\dagger(n) \otimes f_{1,-}(n), \end{aligned} \tag{31}$$

$$\begin{aligned} b_{opt} &= -b_0 \otimes \left(g_{T,+,opt}(n)\right)^\dagger \\ &= -b_1 \otimes \left[f_{1,+}(n) + {w_{opt}}^*(n) \otimes f_{1,-}(n)\right]^\dagger, \end{aligned} \tag{32}$$

and

$$\begin{aligned} \rho_{opt}(n) &= \left(h_{R,+}^*(n)\right)^\dagger \otimes h_{R,-}(n) \\ &= \left(f_{1,+}^*(n)\, e^{-j2\pi\mu n}\right)^\dagger \otimes \left(f_{2,-}(n)\, e^{-j2\pi\mu n}\right). \end{aligned} \tag{33}$$

Therefore, $w_{opt}(n)$, b_{opt}, and $\rho_{opt}(n)$ can be calculated through $f_{1,\pm}(n)$, $f_{2,-}(n)$, and b_1 which along with d_0 can be estimated from (25)-(30) in a much easier way, as to be discussed below. In the proposed method, $f_{1,\pm}(n)$, $f_{2,\pm}(n)$, b_1, b_2 and d_0 will all be estimated under the principle of least-squares with $f_{2,+}(n)$ and b_2 serving as auxiliary variables which are not needed in the final evaluation (see (31)-(33)). To this end, firstly let $f_{1,\pm}(n)$ and $f_{2,\pm}(n)$ abe modeled as FIR (finite impulse response) filters,

$$\mathbf{f}_{i,\pm} = \left[f_{i,\pm}(0), f_{i,\pm}(1), \ldots, f_{i,\pm}(L_f - 1)\right]^T, i = 1, 2, \tag{34}$$

where L_f is the filters' length and usually not known in advance. In Section V, it will be shown that the estimation performance is quite insensitive to the value of L_f if it is of sufficient length.

Consider a training sequence $\{s(n)\}_{n=-K}^{N-1}$, where $K \geq L_f$, and $s(n) = s(n+N)$ $n = -K, \cdots, -1$ is the cyclic-prefix[3]. Define $\mathbf{S}$ be the $N \times L_f$ signal matrix with $[\mathbf{S}]_{i,j} = s(i-j)$, $0 \leq i \leq N-1$, $0 \leq j \leq L_f - 1$, and $\mathbf{f} = \left[\mathbf{f}_{1,+}^T, \mathbf{f}_{1,-}^T, b_1, \mathbf{f}_{2,+}^T, \mathbf{f}_{2,-}^T, b_2, d_0\right]^T$. Then, (25) can be rearranged into the following vector-matrix form

$$\mathbf{r} = \mathbf{\Phi}\mathbf{f} + \mathbf{v}, \tag{35}$$

where $\mathbf{r} = [r(0), r(1), \ldots, r(N-1)]^T$, $\mathbf{\Phi} = \left[\begin{array}{cccc} \mathbf{\Gamma}_N(\mu)\mathbf{T} & \mathbf{\Gamma}_N(-\mu)\mathbf{T} & \mathbf{1}_N \end{array}\right]$, $\mathbf{T} = \left[\begin{array}{ccc} \mathbf{S} & \mathbf{S}^* & \mathbf{1}_N \end{array}\right]$, $\mathbf{\Gamma}_N(\mu) = \text{diag}\left\{1, e^{j2\pi\mu}, \ldots, e^{j2\pi\mu(N-1)}\right\}$ is the diagonal

[3]Generally, N is selected based on a tradeoff between calibration performance and complexity. Given a design of a radio transceiver where the worst values of radio impairments are specified, N can be selected according to the designer's own tradeoff on the performance vs. complexity.

matrix with elements $\left\{1, e^{j2\pi\mu}, \cdots, e^{j2\pi\mu(N-1)}\right\}$, $\mathbf{1}_N$ is the all 1 vector with dimension N, and $\mathbf{v} = [v(0), v(1), \ldots, v(N-1)]^T$. From (35), the least-squares estimate $\hat{\mathbf{f}}$ is given by

$$\hat{\mathbf{f}} = \mathbf{\Upsilon}\mathbf{r}, \tag{36}$$

where $\mathbf{\Upsilon} = \left(\mathbf{\Phi}^H\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^H$ is the pseudo inverse of $\mathbf{\Phi}$. Note that $\mathbf{\Phi}$ has to have full-rank in order to assure identifiability. After obtaining $\hat{\mathbf{f}}$, $w_{opt(n)}$, b_{opt},and $\rho_{opt}(n)$ can be evaluated as in (31), (32), and (33), respectively. Substitute (36) into (35), one has

$$\hat{\mathbf{f}} = \mathbf{f} + \mathbf{\Upsilon}\mathbf{v} \tag{37}$$

which is an unbiased estimate of $\hat{\mathbf{f}}$ with the mean-square error (MSE) given below,

$$\mathrm{E}\left[\left\|\hat{\mathbf{f}} - \mathbf{f}\right\|^2\right] = tr\left\{\mathbf{\Upsilon}\mathrm{E}\left[\mathbf{v}\mathbf{v}^H\right]\mathbf{\Upsilon}^H\right\} = tr\left\{\mathbf{\Upsilon}\mathbf{C}_\mathbf{v}\mathbf{\Upsilon}^H\right\}, \tag{38}$$

where $\mathrm{E}[\cdot]$ denotes the operation of taking expectation, $\mathbf{C}_\mathbf{v} = \mathrm{E}\left[\mathbf{v}\mathbf{v}^H\right]$ is the noise correlation matrix, and $tr\{\mathbf{X}\}$ denotes the trace of the square matrix $\mathbf{X}$. Notice that with no frequency shift, i.e., $\mu = 0$, (35) becomes

$$\mathbf{r} = \mathbf{S}\left(\mathbf{f}_{1,+} + \mathbf{f}_{2,+}\right) + \mathbf{S}^*\left(\mathbf{f}_{1,-} + \mathbf{f}_{2,-}\right) + (b_1 + b_2 + d_0)\,\mathbf{1}_N. \tag{39}$$

In such an undesirable case, $\mathbf{f}_{1,+}$ and $\mathbf{f}_{2,+}$ are not identifiable, so are $\mathbf{f}_{1,-}$ and $\mathbf{f}_{2,-}$, and b_1, b_2, and d_0, and, therefore, it is not possible to estimate $w_{opt}(n)$, $\rho_{opt}(n)$ and b_{opt} as in (31)-(33), respectively. This explains the necessity of the introduction of the frequency shift μ during calibration training.

B. Training Sequence Design

Theoretically, the optimal training sequence is the one that minimizes MSE $\mathrm{E}\left[\left\|\hat{\mathbf{f}} - \mathbf{f}\right\|^2\right] = tr\left\{\mathbf{\Upsilon}\mathbf{C}_\mathbf{v}\mathbf{\Upsilon}^H\right\}$. As is seen in (30), however, $\mathbf{C}_\mathbf{v}$ is a function of $h_{R,+}(n)$ and $h_{R,-}(n)$, and therefore the optimal training sequence differs from one transceiver to another and there is no way to design it. In our method, the simplified measure

$$tr\left\{\mathbf{\Upsilon}\mathbf{\Upsilon}^H\right\} = tr\left\{\left(\mathbf{\Phi}^H\mathbf{\Phi}\right)^{-1}\right\} \tag{40}$$

is adopted in search of good training sequences. The measure is optimal only if $v(n), n = 0, \cdots, N-1$ are white Gaussian noises.

Let $\sum_{n=0}^{N-1}|s(n)|^2 \Big/ N = 1$. It can be shown that

$$tr\{\mathbf{\Phi}^H\mathbf{\Phi}\} = (4L_f + 3)\cdot N, \tag{41}$$

and from [21] the minimum MSE in (40) is achieved provided that

$$\mathbf{\Phi}^H\mathbf{\Phi} = N\cdot\mathbf{I}_{4L_f+3} \tag{42}$$

which in turns leads to

$$\mathbf{T}^H\mathbf{T} = N\cdot\mathbf{I}_{2L_f+1}, \tag{43}$$

$$\mathbf{T}^H\mathbf{\Gamma}(-2\mu)\,\mathbf{T} = \mathbf{0}_{2L_f+1\times 2L_f+1}, \tag{44}$$

and

$$\mathbf{T}^H\mathbf{\Gamma}(\pm\mu)\,1_N = \mathbf{0}_{2L_f+1\times 1}, \tag{45}$$

where $\mathbf{I}_m$ is the identity matrix with dimension m, and $\mathbf{0}_{m\times n}$ is the all zero matrix of size $m\times n$. Clearly, from (43)-(45), $s(n)$ has to be designed jointly with frequency shift μ in order to have the best performance, but that, unfortunately, complicates the design significantly. In the following, a simpler method is proposed.

Consider a periodic training sequence that consists of $P+1$ periods with K samples in each period, i.e., $s(n) = s(n+K)$, $n = -K, \cdots, 0, \cdots, N-K-1$, where $N = KP$. Define $\mathbf{S}_1$ be the signal matrix for one period (from the second period), with its $i-j$ element given by $[\mathbf{S}_1]_{i,j} = s(i-j)$, $0\le i\le K-1$, and $0\le j\le L_f-1$, $\mathbf{T}_1 = [\mathbf{S}_1, \mathbf{S}_1^*, \mathbf{1}_K]$, and $\mathbf{T} = \underbrace{[\mathbf{T}_1^T, \ldots, \mathbf{T}_1^T]}_{P}^T$, then the matrix $\mathbf{\Phi}$ can be decomposed as follows.

$$\begin{aligned}\mathbf{\Phi} &= \left[\begin{array}{ccc}\mathbf{\Gamma}_N(\mu)\,\mathbf{T} & \mathbf{\Gamma}_N(-\mu)\,\mathbf{T} & 1_N\end{array}\right]\\ &= \left[\begin{array}{ccc}\mathbf{\Gamma}_K(\mu)\,\mathbf{T}_1 & \mathbf{\Gamma}_N(-\mu)\,\mathbf{T}_1 & \mathbf{1}_K\\ e^{j2\pi\mu K}\mathbf{\Gamma}_K(\mu)\,\mathbf{T}_1 & e^{-j2\pi\mu K}\mathbf{\Gamma}_K(-\mu)\,\mathbf{T}_1 & \mathbf{1}_K\\ \vdots & \vdots & \vdots\\ e^{j2\pi\mu(P-1)K}\mathbf{\Gamma}_K(\mu)\,\mathbf{T}_1 & e^{-j2\pi\mu(P-1)K}\mathbf{\Gamma}_K(-\mu)\,\mathbf{T}_1 & \mathbf{1}_K\end{array}\right]\\ &\doteq \mathbf{\Phi}_2\mathbf{\Phi}_1,\end{aligned} \tag{46}$$

where

$$\mathbf{\Phi}_1 = \left[\begin{array}{ccc}\mathbf{\Gamma}_K(\mu)\,\mathbf{T}_1 & \mathbf{0}_{K\times 2L_f+1} & \mathbf{0}_{K\times 1}\\ \mathbf{0}_{K\times 2L_f+1} & \mathbf{\Gamma}_K(-\mu)\,\mathbf{T}_1 & \mathbf{0}_{K\times 1}\\ \mathbf{0}_{1\times 2L_f+1} & \mathbf{0}_{1\times 2L_f+1} & \sqrt{K}\end{array}\right], \tag{47}$$

$$\mathbf{\Phi}_2 = \left[\begin{array}{ccc}\mathbf{I}_K & \mathbf{I}_K & \frac{1}{\sqrt{K}}\cdot\mathbf{1}_K\\ e^{j2\pi\mu K}\mathbf{I}_K & e^{-j2\pi\mu K}\mathbf{I}_K & \frac{1}{\sqrt{K}}\cdot\mathbf{1}_K\\ \vdots & \vdots & \vdots\\ e^{j2\pi\mu(P-1)K}\mathbf{I}_K & e^{-j2\pi\mu(P-1)K}\mathbf{I}_K & \frac{1}{\sqrt{K}}\cdot\mathbf{1}_K\end{array}\right]. \tag{48}$$

Thus, $\mathbf{\Phi}_1^H\mathbf{\Phi}_1 = K\cdot\mathbf{I}_{4L_f+3}$ and $\mathbf{\Phi}_2^H\mathbf{\Phi}_2 = P\cdot\mathbf{I}_{2K+1}$ are to constitute a sufficient condition of (42). Furthermore, the condition $\mathbf{\Phi}_1^H\mathbf{\Phi}_1 = K\cdot\mathbf{I}_{4L_f+3}$, called Condition-A, can be split into the following three sub-conditions:

$$Condition - A.1: \qquad \mathbf{S}_1^H\mathbf{S}_1 = K\cdot\mathbf{I}_{L_f}, \tag{49}$$

$$Condition - A.2: \qquad \mathbf{S}_1^T\mathbf{S}_1 = \mathbf{0}_{L_f\times L_f}, \tag{50}$$

and

$$Condition - A.3: \qquad \mathbf{S}_1^T\mathbf{1}_N = \mathbf{0}_{L_f\times 1}. \tag{51}$$

In [22], methods were given to design sequences that satisfy Condition-A.1 and Condition-A.2, while Condition-A.3 just demands that the designed sequence has a zero mean. As an example, using the frequency-domain nulling (FDN) method in [22], the sequence ($K = 64$)

$$s(n) = \frac{1}{N}\sum_{k=0}^{63} S(k)\, e^{\frac{j2\pi nk}{64}}, \quad n = 0, \cdots, 63, \tag{52}$$

with

$$S(k) = \begin{cases} e^{j\phi_k}, & \text{arbitrary } \phi_k, \text{ for } k\in J,\\ & \text{and } J = [1, 5, 9, \ldots, 61]\\ 0, & \text{for } k\notin J\end{cases} \tag{53}$$

can be shown to satisfy Conditions-A.1, A.2 and A.3.

The condition $\mathbf{\Phi}_2^H \mathbf{\Phi}_2 = P \cdot \mathbf{I}_{2K+1}$, called Condition-B, amounts to

$$\begin{bmatrix} P\cdot\mathbf{I}_K & \gamma_1(-\mu)\cdot\mathbf{I}_K & \gamma_2(-\mu)\cdot\mathbf{1}_K \\ \gamma_1(\mu)\cdot\mathbf{I}_K & P\cdot\mathbf{I}_K & \gamma_2(\mu)\cdot\mathbf{1}_K \\ \gamma_2(\mu)\cdot\mathbf{1}_K^H & \gamma_2(-\mu)\cdot\mathbf{1}_K^H & P \end{bmatrix} = P\cdot\mathbf{I}_{2K+1}. \tag{54}$$

That is,

$$\gamma_1(\mu) = \frac{1-\left(e^{j4\pi\mu K}\right)^P}{1-e^{j4\pi\mu K}} = 0, \tag{55}$$

and

$$\gamma_2(\mu) = \frac{1}{\sqrt{K}} \cdot \frac{1-\left(e^{j2\pi\mu K}\right)^P}{1-e^{j2\pi\mu K}} = 0. \tag{56}$$

From (55) and (56), it is concluded that

$$\mu_{opt} = \frac{k}{PK},\ \{k \in Z \,|\, k \notin iP/2, i \in Z\}. \tag{57}$$

In (49)-(51) and (57), we have successfully separated the design of $s(n)$ from that of μ.

IV. Performance Analysis

In this section, we aim to analyze the calibration performance with the estimates given in (36). Specifically, we aim to analyze the probability density functions (pdfs) of calibrated $IRR_T(f)$, $IRR_R(f)$, ε_T and ε_R. For brevity, only $IRR_T(f)$ and ε_T will be treated explicitly here; similar procedures can be applied to analyze $IRR_R(f)$ and ε_R, as were detailed in [23]. Numerical results will be given in Section V to verify the accuracy of the analyses.

To begin with, define $\mathbf{\Upsilon} \doteq \left[\begin{array}{cccccc} \mathbf{\Upsilon}_{\mathbf{f}_{1,+}}^T & \mathbf{\Upsilon}_{\mathbf{f}_{1,-}}^T & \mathbf{\Upsilon}_{b_1}^T & \mathbf{\Upsilon}_{\mathbf{f}_{2,+}}^T & \mathbf{\Upsilon}_{\mathbf{f}_{2,-}}^T & \mathbf{\Upsilon}_{b_2}^T \quad \mathbf{\Upsilon}_{d_0}^T \end{array}\right]^T$. From (37), then we have $\hat{\mathbf{f}}_{1,+} = \mathbf{f}_{1,+} + \mathbf{\Upsilon}_{\mathbf{f}_{1,+}}\mathbf{v}$, $\hat{\mathbf{f}}_{1,-} = \mathbf{f}_{1,-} + \mathbf{\Upsilon}_{\mathbf{f}_{1,-}}\mathbf{v}$, $\hat{b}_1 = b_1 + \mathbf{\Upsilon}_{b_1}\mathbf{v}$, $\hat{\mathbf{f}}_{2,-} = \mathbf{f}_{2,-} + \mathbf{\Upsilon}_{\mathbf{f}_{2,-}}\mathbf{v}$, and $\hat{d}_0 = d_0 + \mathbf{\Upsilon}_{d_0}\mathbf{v}$. During the internal loopback, the signal-to-noise ratio ($\mathrm{SNR}_{loopback}$) is usually very high, say $\mathrm{SNR}_{loopback} > 30\mathrm{dB}$, and, therefore, it is reasonable to assume that $\hat{\mathbf{f}}_{1,+} \approx \mathbf{f}_{1,+}$. Recall that $\mathbf{f}_{1,+}$ is the desired channel response from transmitter to receiver, as is given in (26). Using this approximation, the estimated calibration filters $\hat{w}(n)$ in (31) is

$$\hat{w}(n) = -\left(\hat{f}_{1,+}(n)\right)^\dagger \otimes \hat{f}_{1,-}(n) \approx -\left(f_{1,+}(n)\right)^\dagger \otimes \hat{f}_{1,-}(n). \tag{58}$$

Without loss of generality, $\hat{w}(n)$ will be modeled as an FIR filter in this analysis and denoted by $\hat{\mathbf{w}} = [\hat{w}(0), \hat{w}(1), \ldots, \hat{w}(L-1)]^T$, where L can be selected as long as one wishes for the desirable analysis accuracy. With this modeling, (58) can be rearranged in the following vector-matrix form

$$\hat{\mathbf{w}} \approx -\mathbf{F}_{1,+} \times \hat{\mathbf{f}}_{1,-} = -\mathbf{F}_{1,+} \times \left(\mathbf{f}_{1,-} + \mathbf{\Upsilon}_{\mathbf{f}_{1,-}}\mathbf{v}\right) = \mathbf{w} + \mathbf{v}_w, \tag{59}$$

where $\mathbf{F}_{1,+}$ is the $L \times L_f$ sized convolution matrix of the truncated inverse filter $\left(f_{1,+}(n)\right)^\dagger$,

$$\mathbf{w} \doteq [w(0), w(2), \cdots, w(L-1)]^T = -\mathbf{F}_{1,+}\mathbf{f}_{1,-}, \tag{60}$$

(see (31)), and

$$\mathbf{v}_w \doteq [v_w(0), v_w(1), \ldots, v_w(L-1)]^T = -\mathbf{F}_{1,+}\mathbf{\Upsilon}_{\mathbf{f}_{1,-}}\mathbf{v}. \tag{61}$$

In addition, $g_{T,+}(n)$ in (8) and $g_{T,-}(n)$ in (9) are approximated as

$$\hat{g}_{T,+}(n) \doteq h_{T,+}(n) + \hat{w}^*(n) \otimes h_{T,-}(n) \approx h_{T,+}(n), \tag{62}$$

and (see (59))

$$\begin{aligned} \hat{g}_{T,-}(n) &\doteq h_{T,-}(n) + \hat{w}(n) \otimes h_{T,+}(n) \\ &= \underbrace{h_{T,-}(n) + w(n) \otimes h_{T,+}(n)}_{\approx 0} + v_w(n) \otimes h_{T,+}(n) \\ &\approx v_w(n) \otimes h_{T,+}(n). \end{aligned} \tag{63}$$

The approximation in (62) is good because $|h_{T,+}(n)| \gg |h_{T,-}(n)|$ in real systems (see Section V), and (63) is good because ideally $w(n)$ is sought to make $h_{T,-}(n) + w(n) \otimes h_{T,+}(n) = 0$ in our method. Using (62) and (63), the calibrated $IRR_T(f)$ in (14) is evaluated by

$$\begin{aligned} IRR_T(f) &\approx 10\log_{10}\frac{|H_{T,+}(f)|^2}{|H_{T,+}(f)|^2|V_w(f)|^2} \\ &= -10\log_{10}|V_w(f)|^2(\mathrm{dB}). \end{aligned} \tag{64}$$

Furthermore, $v(n)$ in (30) is approximated by $h_{R,+}(n) \otimes v_0(n)$ because $|h_{R,+}(n)| \gg |h_{R,-}(n)|$. Using this approximation in (61), we have

$$V_w(f) \doteq \mathrm{FT}[\mathbf{v}_w] \approx -\mathbf{\Psi}^H(f)\mathbf{F}_{1,+}\mathbf{\Upsilon}_{\mathbf{f}_{1,-}}\mathbf{G}\mathbf{v}_0 = \boldsymbol{\chi}_w^H(f)\mathbf{v}_0, \tag{65}$$

where $\mathbf{\Psi}(f) = \left[1, e^{j2\pi f}, \ldots, e^{j2\pi f(L-1)}\right]^T$, $\mathbf{G}$ is the $N \times N$ sized convolution matrix of $h_{R,+}(n)$, $\mathbf{v}_0 = [v_0(0), v_0(1), \ldots, v_0(N-1)]^T$, and $\boldsymbol{\chi}_w^H(f) = -\mathbf{\Psi}^H(f)\mathbf{F}_{1,+}\mathbf{\Upsilon}_{\mathbf{f}_{1,-}}\mathbf{G}$. Since $\{v_0(n)\}_{n=0}^{N-1}$ are zero mean i.i.d complex circular symmetric Gaussian variables with variance σ_0^2, $|V_w(f)|^2$ is an exponentially distributed random variable with the pdf

$$\begin{aligned} p\left(|V_w(f)|^2\right) &= \frac{1}{\boldsymbol{\chi}_w^H(f)\boldsymbol{\chi}_w(f)\sigma_0^2} e^{-\frac{1}{\boldsymbol{\chi}_w^H(f)\boldsymbol{\chi}_w(f)\sigma_0^2}|V_w(f)|^2}, \\ &|V_w(f)|^2 \geq 0, \end{aligned} \tag{66}$$

and the pdf of $IRR_T(f)$ is

$$\begin{aligned} p(IRR_T(f)) &= \frac{\log_e 10}{10\cdot\boldsymbol{\chi}_w^H(f)\boldsymbol{\chi}_w(f)\sigma_0^2} 10^{\frac{-IRR_T(f)}{10}} e^{-\frac{10^{\frac{-IRR_T(f)}{10}}}{\boldsymbol{\chi}_w^H(f)\boldsymbol{\chi}_w(f)\sigma_0^2}}, \\ &-\infty < IRR_T(f) < \infty, \end{aligned} \tag{67}$$

Finally, it can be shown from (66) and [24] that

$$\begin{aligned} \mathrm{E}\{IRR_T(f)\} &= -10\cdot\mathrm{E}\left\{\log_{10}|V_w(f)|^2\right\} \\ &= -10\log_{10}\frac{\boldsymbol{\chi}_w^H(f)\boldsymbol{\chi}_w(f)\sigma_0^2}{e^{\mathbf{C}}}, \end{aligned} \tag{68}$$

$$\begin{aligned} \mathrm{E}\left\{(IRR_T(f))^2\right\} &= 100\cdot\mathrm{E}\left\{\left(\log_{10}|V_w(f)|^2\right)^2\right\} \\ &= \left(10\log_{10}e^{\frac{\pi}{\sqrt{6}}}\right)^2 + (\mathrm{E}\{IRR_T(f)\})^2, \end{aligned} \tag{69}$$

and

$$\mathrm{VAR}\{IRR_T(f)\} = \left(10\log_{10}e^{\frac{\pi}{\sqrt{6}}}\right)^2, \tag{70}$$

where $e^{\mathbf{C}} = 1.781072...$ is the Euler's constant.

TABLE I
THE RF IMPAIRMENTS

RF Impairments	Parameter Value
Frequency independent I-Q imbalance (α_T, θ_T), (α_R, θ_R)	$(\alpha_T = 1.05, \theta_T = -5^o)$, $(\alpha_R = 1.08, \theta_R = 5^o)$
Frequency dependent I-Q imbalance $\{h_T^I(n), h_T^Q(n)\}, \{h_R^I(n), h_R^Q(n)\}$	I part : [1 0.2 0.1 0.05] Q part : [0.9 0.1 0.08 0.12]
DC offset b_0 and d_0, with signal power normalized to 1	$b_0 = -0.1 \times (1+j)/\sqrt{2}$, $d_0 = 0.1 \times (1+j)/\sqrt{2}$

To analyze ε_T, first note that $\hat{b} \approx -\hat{b}_1 \otimes \left(f_{1,+}(n)\right)^{\dagger}$ in (32) because $f_{1,+}(n) \gg f_{1,-}(n)$ in real systems. Using $\hat{b}_1 = b_1 + \boldsymbol{\Upsilon}_{b_1}\mathbf{v}$, then $\hat{b}$ is approximated as

$$\hat{b} \approx \frac{-b_1 - \boldsymbol{\Upsilon}_{b_1}\mathbf{v}}{\sum\limits_n f_{1,+}(n)} \approx b + v_b, \tag{71}$$

where

$$v_b = -\frac{\boldsymbol{\Upsilon}_{b_1}\mathbf{v}}{\sum\limits_n f_{1,+}(n)}. \tag{72}$$

Furthermore, from (10), (71) and (62), Δb is approximated as

$$\begin{aligned}
\Delta b &\approx \hat{g}_{T,+}(n) \otimes \hat{b} + \hat{g}_{T,-}(n) \otimes \hat{b}^* + b_0 \\
&= \underbrace{\hat{g}_{T,+}(n) \otimes b + \hat{g}_{T,-}(n) \otimes b^* + b_0}_{\approx 0} + \hat{g}_{T,+}(n) \otimes v_b \\
&\quad + \hat{g}_{T,-}(n) \otimes v_b^* \\
&\approx \hat{g}_{T,+}(n) \otimes v_b \\
&\approx h_{T,+}(n) \otimes v_b \\
&= -\frac{\left(\sum\limits_n h_{T,+}(n)\right) \cdot \boldsymbol{\Upsilon}_{b_1}\mathbf{v}}{\sum\limits_n f_{1,+}(n)} \\
&\approx \boldsymbol{\chi}_{\Delta b}^H \mathbf{v}_0,
\end{aligned} \tag{73}$$

where

$$\boldsymbol{\chi}_{\Delta b}^H = -\frac{\left(\sum\limits_n h_{T,+}(n)\right) \cdot \boldsymbol{\Upsilon}_{b_1}\mathbf{G}}{\sum\limits_n f_{1,+}(n)} \tag{74}$$

Similar to (67)-(70), we have

$$\begin{aligned}
p(\varepsilon_T) &= \frac{|b_0|^2 \log_e 10}{10 \cdot \boldsymbol{\chi}_{\Delta_b}^H \boldsymbol{\chi}_{\Delta_b} \sigma_0^2} 10^{\frac{\varepsilon_T}{10}} e^{-\frac{|b_0|^2 10^{\frac{\varepsilon_T}{10}}}{\boldsymbol{\chi}_{\Delta_b}^H \boldsymbol{\chi}_{\Delta_b} \sigma_0^2}}, \\
&\quad -\infty < \varepsilon_T < \infty,
\end{aligned} \tag{75}$$

$$\begin{aligned}
\mathrm{E}\{\varepsilon_T\} &= \mathrm{E}\left\{10\log_{10}\frac{|\Delta b|^2}{|b_0|^2}\right\} \\
&\approx 10\log_{10}\frac{\boldsymbol{\chi}_{\Delta b}^H \boldsymbol{\chi}_{\Delta b} \sigma_0^2}{|b_0|^2 e^{\mathbf{C}}},
\end{aligned} \tag{76}$$

and

$$\mathrm{VAR}\{\varepsilon_T\} = \left(10\log_{10} e^{\frac{\pi}{\sqrt{6}}}\right)^2. \tag{77}$$

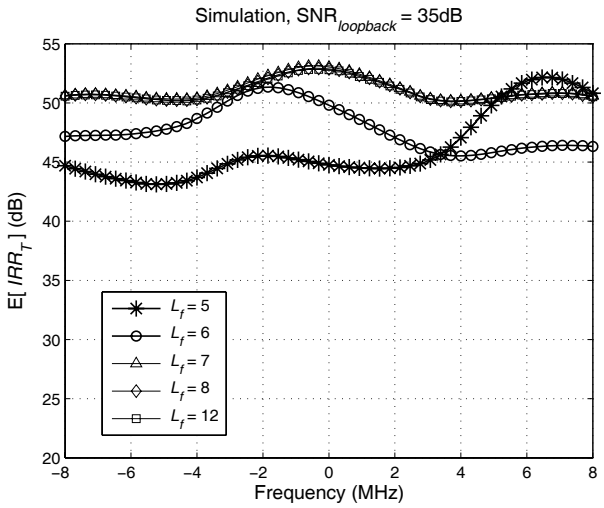

Fig. 2. Performance of the calibrated $\mathrm{E}[IRR_T]$ with different $L_f's$.

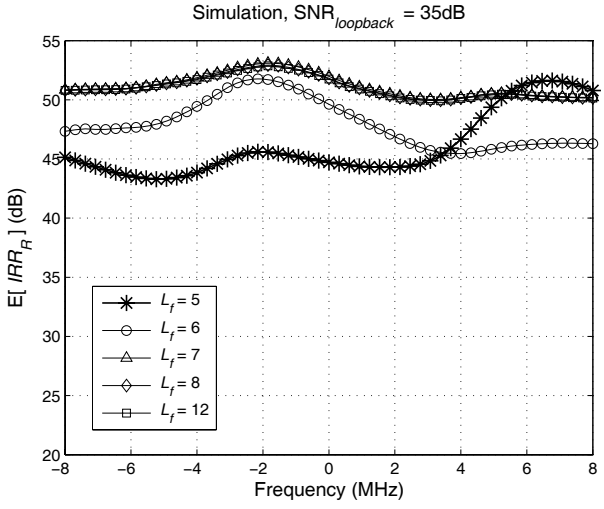

Fig. 3. Performance of the calibrated $\mathrm{E}[IRR_R]$ with different $L_f's$.

V. NUMERICAL RESULTS

In this section, the performance of the proposed calibration method is evaluated through analysis and computer simulations. Table 1 summarizes the transmitter and receiver RF impairments which are typical values in real systems [1]-[18],[25]. In all results, $1/T_s = 20$MHz, $\mathrm{SNR} \doteq (1/N)\sum_{n=0}^{N}|s(n)|^2 \Big/ \sigma_0^2$, and each simulation point is obtained with 10^6 realizations. In addition, $K = 64$, $P = 3$, and $N = KP = 192$. Note that two types of SNR need to be differentiated in the proposed method: $\mathrm{SNR}_{loopback}$ and $\mathrm{SNR}_{channel}$. $\mathrm{SNR}_{loopback}$ is the SNR defined for calibration training (internal loopback) whereas $\mathrm{SNR}_{channel}$ defined for normal communication. In real systems, $\mathrm{SNR}_{loopback} \gg \mathrm{SNR}_{channel}$ because there is no propagation loss during the internal loopback.

Figures 2 and 3 investigate the effects of L_f on the performance of the calibrated $\mathrm{E}[IRR_T(f)]$ and $\mathrm{E}[IRR_R(f)]$, respectively, by computer simulations. The training sequence is the one in (53), and $\mu = 23/(3 \cdot 64)$. Recall that L_f is the length of the filters $f_{i,\pm}(n), i = 1, 2$. As can be seen, the

TABLE II
EXAMPLE MEANS AND STANDARD DEVIATIONS OF THE CALIBRATED IRR_T, IRR_R, ε_T, AND ε_R.

Parameters	$\text{SNR}_{loopback}$	Mean (dB)		Standard deviation (dB)	
		Simulation	Analysis	Simulation	Analysis
IRR_T (4MHz)	35	50.8	50.9	5.59	5.57
	45	60.8	60.9	5.57	5.57
	55	70.8	70.9	5.56	5.57
IRR_R (4MHz)	35	50.8	50.8	5.53	5.57
	45	60.8	60.8	5.58	5.57
	55	70.8	70.8	5.58	5.57
ε_T	35	-40	-40.3	5.57	5.57
	45	-50	-50.3	5.57	5.57
	55	-60	-60.3	5.56	5.57
ε_R	35	-38	-38	5.6	5.57
	45	-48	-48	5.57	5.57
	55	-58	-58	5.59	5.57

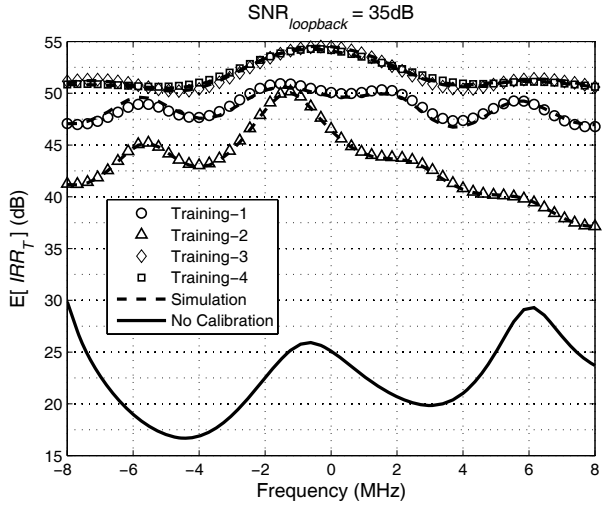

Fig. 4. Performance of the calibrated $\mathrm{E}[IRR_T]$ with different training designs.

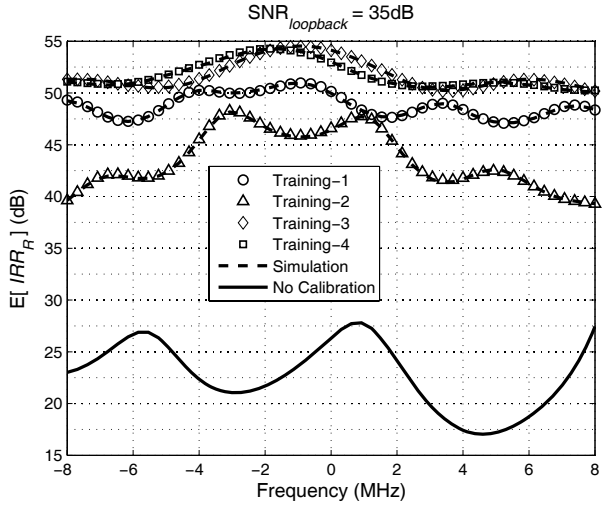

Fig. 5. Performance of the calibrated $\mathrm{E}[IRR_R]$ with different training designs.

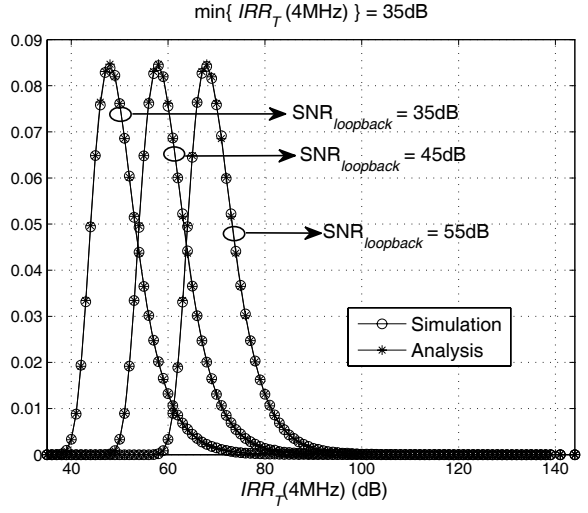

Fig. 6. Analytical and simulated pdfs of the calibrated IRR_T at frequency 4MHz.

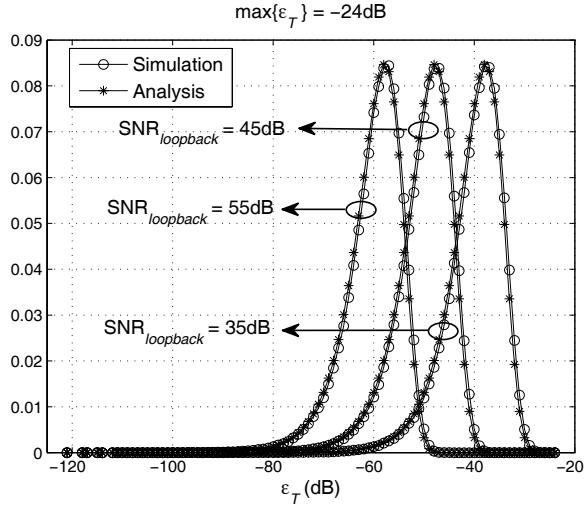

Fig. 7. Analytical and simulated pdfs of the calibrated ε_T.

calibration performances are quite insensitive to the values of L_f as long as it is larger than 6 in this case; similar results are observed for the dc offset calibration. Since L_f may not be known exactly in advance, it is advisable to use a sufficiently large L_f to avoid performance degradation. $L_f = 7$ is used for all the results that follow.

In Figures 4 and 5, the calibrated $\mathrm{E}[IRR_T(f)]$ and $\mathrm{E}[IRR_R(f)]$ are investigated with four periodic training designs. Training-1 uses the sequence in (53) but with $S(k) = e^{j\phi_k}\forall k$, and $\mu = 23/(3 \cdot 64)$, and Training-2, Training-3 and Training-4 use the sequence in (53) with $\mu = 24/(3 \cdot 64)$,

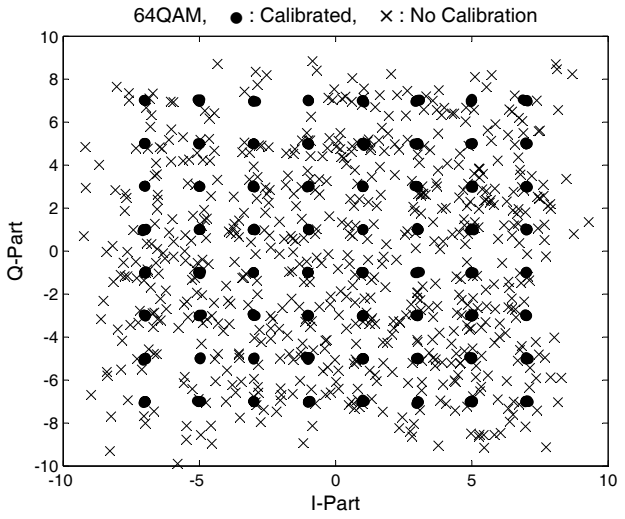

Fig. 8. Sample signal constellation with and without calibrations ($\mathrm{SNR}_{loopback} = 35\mathrm{dB}$, $\mathrm{SNR}_{channel} = \infty$).

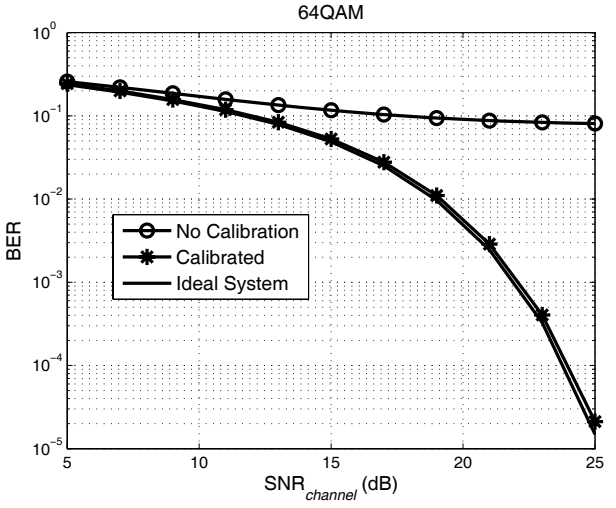

Fig. 9. Bit error rate performance with and without calibration ($\mathrm{SNR}_{loopback} = 35\mathrm{dB}$).

$\mu = 11.4/(3 \cdot 64)$ and $\mu = 23/(3 \cdot 64)$, respectively. For comparison purpose, Training-1 is selected to violate Conditions-A.2 and Condition-A.3, Training-2 and Training-3 are selected to violate Condition-B while Training-4 is the optimal training under the simplified criterion of (40) that satisfies both Condition-A and Condition-B. As can be seen in the figures, violation of Condition-A or Condition-B may incur a large performance loss. In addition, the proposed method provides around 20-35 dB performance improvement over the whole frequency band, as compared to the case of no calibration. The figures also show a nearly perfect match between simulation and analytical results. In the rest of this section, Training-4 is used for the calibration training.

Figure 6 shows the analytical and simulated pdfs of the calibrated $IRR_T(f)$ at f= 4MHz, where it shows very good match between simulation and analysis. The smallest simulated IRR_T is at 36 dB which are around 16 dB better than the cases of no calibration. Figure 7 shows the analytical and simulated pdfs of the calibrated ε_T, where the largest ε_T is at -24 dB. Very significant improvements are observed with the proposed method. Table 2 gives example simulated and analytical means and standard deviations of IRR_T, IRR_R, ε_T and ε_R under different SNRs. Again, it shows very good match between simulation and analysis.

Figures 8 and 9 show a sample received signal constellation and bit error rate performance respectively for an un-coded 64-QAM OFDM (orthogonal frequency-division multiplexing) system with and without calibration. The simulations are obtained under normal communication where transmitter and receiver are located at different devices with $\mu = 0$ and $c(t) = \delta(t)$. A one-tap equalizer is employed at the receiver for the simulated OFDM system that uses 64-point FFT (fast Fourier transform) with 52 subcarriers carrying data. In Fig. 8, $\mathrm{SNR}_{channel} = \infty$ is adopted because, in doing so, the sole effect of residual radio impairments on the constellation points can be investigated. As are shown in the figures, the adverse effects due to radio impairments are removed almost completely by the proposed calibration.

VI. Conclusion

A digital calibration method is proposed for the direct-conversion radio transceiver to calibrate its own transmitter and receiver radio impairments, including frequency-independent I-Q imbalance, frequency-dependent I-Q imbalance, and dc offset. By introducing a shift between transmit and receive frequencies, the radio impairments appearing at the transmitter and receiver can be calibrated simultaneously without a dedicated analog circuitry in the feedback loop. The calibration parameters are estimated based on the nonlinear least-squares principle, and the calibration performance is analyzed that agrees very well with the simulations. The issue of training design is also investigated; sufficient conditions for optimal training are provided under a simplified criterion, and an example of optimal training is given for the periodic training structure. Analytical and simulation results show significant improvement is obtained with the proposed method, as compared to the non-calibrated systems.

References

[1] B. Razavi, *RF Microelectronics*. Prentice-Hall, 1998.
[2] A. A. Abidi, "Direct-conversion radio transceivers for digital communications," *IEEE J. Solid-State Circuits*, vol. 30, pp. 1399–1410, Dec. 1995.
[3] M. Valkama, M. Renfors, and V. Koivunen, "Advanced methods for I/Q imbalance compensation in communication receivers," *IEEE Trans. Signal Process.*, vol. 49, pp. 2335–2344, Oct. 2001.
[4] G. Xing, M. Shen, and H. Liu, "Frequency offset and I/Q imbalance compensation for direct conversion receivers," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 673–680, Mar. 2005.
[5] G. T. Gil, I. H. Sohn, J. K. Park, and Y. H. Lee, "Joint ML estimation of carrier frequency, channel, I/Q mismatch, and DC offset in communication receivers," *IEEE Trans. Veh. Technol.*, vol. 54, no. 1, pp. 338–349, Jan. 2005.
[6] A. Tarighat and A. H. Sayed, "Joint compensation of transmitter and receiver impairments in OFDM systems," *IEEE Trans. Wireless Commun*, vol. 6, pp. 240–246, Jan. 2007.
[7] D. Tandur and M. Moonen, "Joint adaptive compensation of transmitter and receiver IQ imbalance under carrier frequency offset in OFDM-based systems," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5246–5252, Nov. 2007.
[8] C.-J. Hsu, R. Cheng, and W.-H. Sheen, "Joint least-squares estimation of frequency, DC-offset, I-Q imbalance and channel in MIMO receivers," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 2201–2213, June 2009.

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

[10] D. Hilborn, S. Stapleton, and J. Cavers, "An adaptive direct conversion transmitter," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 223–233, May 1994.

[11] R. Marchesani, "Digital pre-compensation of imperfections in quadrature modulators," *IEEE Trans. Commun.*, vol. 48, no. 4, pp. 552–556, Apr. 2000.

[12] A. G. K. C. Lim, V. Sreeram, and G.-Q. Wang, "Digital compensation in IQ modulators using adaptive FIR filters," *IEEE Trans. Veh. Technol.*, vol. 53, no. 6, pp. 1809–1817, Nov. 2004.

[13] J. Tuthill and A. Cantoni, "Efficient compensation for frequency-dependent errors in analog reconstruction filters used in I/Q modulators," *IEEE Trans. Commun.*, vol. 53, pp. 489–496, Mar. 2005.

[14] M. Windisch and G. Fettweis, "Adaptive I/Q imbalance compensation in low-IF transmitter architectures," in *Proc. 2004 IEEE Veh. Technol. Conf.*, pp. 2096–2100.

[15] L. Anttila, M. Valkama, and M. Renfors, "Frequency-selective I/Q mismatch calibration of wideband direct-conversion transmitters," *IEEE Trans. Circuits and Systems II*, vol. 55, pp. 359–363, Apr. 2008.

[16] Y. Zou, M. Valkama, and M. Renfors, "Pilot-based compensation of frequency-selective I/Q imbalances in direct-conversion OFDM transmitters," in *Proc. 2008 IEEE Veh. Technol. Conf.*

[17] S. A. Bassam, S. Boumaiza, and F. M. Ghannouchi, "Block-wise estimation of and compensation for IQ imbalance in direct-conversion transmitters," *IEEE Trans. Signal Process.*, vol. 57, pp. 4970–4973, Dec. 2009.

[18] B. Debaillie, P. V. Wesemael, G. Vandersteen, and J. Craninckx, "Calibration of direct-conversion transceivers," *IEEE J. Sel. Topics Signal Process.*, vol. 3, pp. 488–498, June 2009.

[19] C.-J. Hsu and W.-H. Sheen, "A new self-calibration method for transmitter and receiver radio impairments in direct-conversion architecture," in *Proc. 2010 IEEE Globecom*, pp. 1–6.

[20] U. L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design*. John Wiley Sons, 1997.

[21] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1158–1168, May 2006.

[22] H. Minn and D. Munoz, "Pilot designs for channel estimation of OFDM systems with frequency-dependent I/Q imbalances," in *Proc. 2009 IEEE WCNC*, pp. 713–717.

[23] C.-J. Hsu, "On the radio impairment estimation and compensation techniques for wideband MIMO direct-conversion transceivers," Ph.D. dissertation, Dept. of Elect. Eng., National Chiao Tung University, Hsinchu, Taiwan, 2010.

[24] I. S. Gradshteyn, I. M. Ryzhik, A. Jeffrey, and A. Zwillinger, editors, *Table of Integrals, Series, and Products*, 7th edition. Academic Press, 2007.

[25] M. Mailand, R. Richter, and H.-J. Jentschel, "IQ-imbalance and its compensation for non-ideal analog receivers comprising frequency-selective components," *Adv. Radio Sci.*, vol. 4, pp. 189–195, 2006.

Chen-Jui Hsu was born in HsinChu, Taiwan, in 1982. He received the B.S. degree in electronics engineering in 2004, the M.S. degree in communication engineering in 2006, and the Ph.D degree in electrical engineering in 2010, all from the National Chiao Tung University, Taiwan.

Since 2010, he has been a principal engineer with digital communication division, MStar Semiconductor, Inc., Hsinchu, Taiwan. His research interest is in the inner receiver design of wireless communications, particularly in RF front-end impairments estimation and compensation algorithms.

Wern-Ho Sheen (M'91) Prof. Wern-Ho Sheen received his Ph.D. degree from the Georgia Institute of Technology, Atlanta, USA in 1991. From 1991 to 1993, he was with Chunghwa Telecom Labs as an associate researcher. From 1993 to 2001, he was with the National Chung Cheng University, where he held positions as Professor in the Department of Electrical Engineering and the Managing Director of the Center for Telecommunication Research. From 2001 to 2009, he was a Professor in the Department of Communications Engineering, National Chiao Tung University. Currently he is with the Department of Information and Communication Engineering, Chaoyang University of Technology. Prof. Sheen has been an active researcher in the areas of communication theory, mobile cellular systems, signal processing for wireless communications, and chip implementation of wireless communications systems.