Novel IQ imbalance and offset compensation techniques for quadrature mixing radio transceivers

Josias J. de Witt, and Gert-Jan van Rooyen

Department of Electric and Electronic Engineering, Stellenbosch University, Stellenbosch 7600.
Tel: (021) 8084315, Fax: (021) 8083951, E-mail: {jdewitt,gvrooyen}@sun.ac.za

Abstract—Despite the advantages that quadrature mixing offers to radio front-ends, its practical use has been limited due to its sensitivity towards gain and phase mismatches between its in-phase and quadrature channels. DC offsets are also a problem when a zero-IF transceiver topology is used. In this paper, novel digital compensation techniques are introduced to extract and correct the quadrature imbalances (gain and phase errors), as well as any offset errors that may exist in the quadrature mixing front-end of the modulator or demodulator. The techniques developed in this paper are not based on iterative search techniques, but rather use spectral measurements to extract the imbalance and offset errors directly. Simulation results show that the imbalance and offset errors of the modulator and demodulator can be suppressed to the noise floor of the transceiver, thus enabling the practical use of quadrature mixing front-ends.

Index Terms—Access network technologies, radio technologies, quadrature imbalance compensation, I/Q imbalance compensation.

I. INTRODUCTION

The quadrature mixing front-ends employed in zero-IF and low-IF transceivers, offer a low-cost, flexible alternative to traditional heterodyne mixing front ends. Its architecture lends itself well to low-cost, low-power, monolithic implementations, while providing theoretically infinite image rejection ratios. This image rejection quality eliminates the need for many off-chip components.

The main drawbacks to quadrature mixing front-ends are their sensitivity towards gain and phase mismatches between the in-phase (I) and quadrature (Q) channels [1]. In zero-IF implementations, DC offsets also become a significant problem.

Digital compensation techniques have been widely suggested in literature to compensate for frequency independent mismatches in quadrature mixing transceivers. The authors of [1]–[3] use envelope detector feedback paths to analyse the imbalances in the modulator frontend. Compensation is then performed with the use of test tones and iterative adjustments to the compensation parameters. [4] and [5] avoid test signals by using adjacent channel power measurements while iteratively searching for the optimal compensation parameters.

This work was supported by the Telkom Centre of Excellence. J. J. de Witt and G.-J. van Rooyen are with the University of Stellenbosch.

Current techniques for demodulator compensation either make use of analogue test signals inserted directly into the demodulator [6], or on test signals or real data signals originating from a modulator. When a modulator is used it is assumed that it does not contribute to the imbalances. For transceivers employing digital modulation schemes, some authors (see e.g. [1], [7]) suggest the use of training sequences generated by a modulator and adaptive algorithms to estimate and correct imbalance errors. [1] also suggests that such adaptive algorithms can be used in a decision directed fashion, resembling an adaptive equalizer.

Another popular approach to demodulator compensation is the use of blind techniques such as interference-canceller (IC) based methods, [8], [9], and blind source separation (BSS) [9], [10]. These methods require no training or calibration signals and are usually implemented on receivers employing a low-IF architecture.

In this paper, novel quadrature imbalance and offset compensation techniques are presented that do not require an iterative search for the optimal parameters. Instead, spectral measurements of the relative sideband ratio and DC spur are used to determine the compensation parameters directly.

To understand the effects of quadrature imbalance and offset errors, a mathematical system model is presented in section II, after which the basic principles of compensation are described in section III. This is followed by a description of the novel imbalance and offset extraction techniques in sections IV and V. Section VI shows how the developed techniques can be used to do complete automatic transceiver calibration. Finally some simulation results are presented in section VII and the paper is concluded in section VIII.

II. SYSTEM MODEL

A. Modulator

In a practical quadrature modulator the I and Q channels may have different gains and the local oscillator (LO) signals may not be exactly 90° out of phase. We denote the gain of the I and Q channels as \( g_I \) and \( g_Q \), respectively. We denote the difference in phase from 90° between the two oscillator signals of the mixer by \( \phi \). We may now define the gain imbalance \( g_M \) of the modulator as

\[
g_M = g_I / g_Q.
\]
The oscillator signal at the modulator, \( x_M(t) \), can then be written as
\[
x_M(t) = g_M \alpha_Q \cos(2\pi f_c t + \phi/2) + j \alpha_Q \sin(2\pi f_c t - \phi/2).
\]
(2)

The gain of the Q channel, \( \alpha_Q \), will only have an overall scaling effect and can thus, without loss of generality, be normalised to unity (\( \alpha_Q = 1 \)) for convenience. \( x_M(t) \) can be written in exponential form as
\[
x_M(t) = V_1 e^{j2\pi f_c t} + V_2 e^{-j2\pi f_c t},
\]
where \( V_1 \) and \( V_2 \) were defined as
\[
\begin{align*}
V_1 &= \cos(\phi_s/2) \left( \frac{g_M + 1}{2} \right) + j \sin(\phi_s/2) \left( \frac{g_M - 1}{2} \right) \\
V_2 &= \cos(\phi_s/2) \left( \frac{g_M - 1}{2} \right) - j \sin(\phi_s/2) \left( \frac{g_M + 1}{2} \right).
\end{align*}
\]
(4)

From equation 3, it is seen that the oscillator imbalances cause two frequency translations to take place, instead of only one.

Apart from imperfections in the mixer, there may also exist non-zero DC-offsets in the I and Q signal paths, denoted by \( \kappa_I \) and \( \kappa_Q \) for the two channels respectively. Let \( \kappa = \kappa_I + j\kappa_Q \). In the transmitted signal, there may also be a carrier leak-through component with an unknown amplitude \( \alpha_M \) and phase \( \gamma \), denoted by \( \alpha_M \cos(2\pi f_c t + \gamma) \). Taking these practical effects into account, it can be seen from Fig. 1, the transmitted signal at the output of the imbalanced modulator, \( s(t) \), can be written as
\[
s(t) = \text{Re} \left\{ [m(t) + \kappa] x_M(t) \right\} + \alpha_M \cos(2\pi f_c t + \gamma) \\
= \left[ \kappa_I + m_I(t) \right] g_M \cos(2\pi f_c t + \phi/2) \\
- \left[ \kappa_Q + m_Q(t) \right] \sin(2\pi f_c t - \phi/2) \\
+ \alpha_M \cos(2\pi f_c t + \gamma),
\]
(5)

where \( m(t) = m_I(t) + j m_Q(t) \) is the transmitted baseband message signal.

It is instructive to look at the (complex) baseband equivalent representation of \( s(t) \), denoted by \( \tilde{s}(t) \)
\[
\tilde{s}(t) = \underbrace{m(t)V_1}_{\text{Desired signal}} + \underbrace{m^*(t)V_2^*}_{\text{Image signal}} + \underbrace{(\kappa V_1 + \kappa^* V_2^* + \alpha_M e^{j\gamma})}_{\text{DC offset}}.
\]
(6)

The baseband or lowpass equivalent of the passband signal \( s(t) \) is defined as \( \tilde{s}(t) = [s(t) + j \tilde{u}(t)] e^{-j2\pi f_c t} \) [11, pp. 51], where \( \tilde{s}(t) \) denotes the Hilbert transform of \( s(t) \).

Using a length-2 column vector notation\(^2\) to represent the real and imaginary parts of a complex signal, the complex baseband envelope \( \tilde{s}(t) \) of the passband signal \( s(t) \) can be written as
\[
\tilde{s}(t) = M [m(t) + a] + b,
\]
(7)

where we have defined
\[
\begin{align*}
a &= [\kappa_I \quad \kappa_Q]^T \\
b &= [\alpha_M \cos(\gamma) \quad \alpha_M \sin(\gamma)]^T \\
M &= \begin{bmatrix} g_M \cos(\phi/2) & \sin(\phi/2) \\
g_M \sin(\phi/2) & \cos(\phi/2) \end{bmatrix}.
\end{align*}
\]
(8)

B. Demodulator

Using a similar derivation, we may use \( \beta_I \) and \( \beta_Q \) to represent the gain of the I and Q channels respectively of the practical quadrature demodulator. We define the amplitude imbalance of the demodulator, \( g_D \), as
\[
g_D = \beta_I / \beta_Q.
\]
(9)

\( \varphi \) is used to represent the difference in phase from 90\(^\circ\) between the two oscillator signals of the mixer. The oscillator signal at the demodulator \( x_D(t) \), can then be written as
\[
x_D(t) = g_D \cos(2\pi f_c t + \varphi/2) - j \sin(2\pi f_c t - \varphi/2) \\
= W_1 e^{-j2\pi f_c t} + W_2 e^{j2\pi f_c t},
\]
(10)

where we have once again normalised the amplitude of the Q channel to unity (\( \beta_Q = 1 \)) for convenience. \( W_1 \) and \( W_2 \) is defined as
\[
\begin{align*}
W_1 &= \cos(\varphi/2) \left( \frac{g_D + 1}{2} \right) + j \sin(\varphi/2) \left( \frac{1 - g_D}{2} \right) \\
W_2 &= \cos(\varphi/2) \left( \frac{g_D - 1}{2} \right) + j \sin(\varphi/2) \left( \frac{g_D + 1}{2} \right)
\end{align*}
\]
(11)

Apart from imperfections in the mixer, there may once again exist non-zero DC offsets in the I and Q signal paths, denoted by \( \chi_I \) and \( \chi_Q \) for the two channels respectively. Let \( \chi = \chi_I + j \chi_Q \). There may also be a carrier leak-through component with unknown amplitude \( \beta_D \) and phase \( \rho \), denoted by \( \beta_D \cos(2\pi f_c t + \rho) \). Combining all these factors, we may write the output of the imbalanced demodulator, \( u(t) \), after a passband signal \( r(t) \) has been received as
\[
u(t) = [r(t) + \beta_D \cos(2\pi f_c t + \rho)] x_D(t) + \chi.
\]
(12)

\(^2\)For instance, the signal \( m(t) = m_I(t) + j m_Q(t) \) is written as \( m(t) = [m_I(t) \quad m_Q(t)]^T \)
If lowpass filters are employed to remove the double frequencies, then the resultant signal \( v(t) \) can be written as

\[
v(t) = [u(t)]_{\text{LPF}} = \tilde{r}(t)W_1 + \tilde{r}^*(t)W_2 + (1/2)\beta_D (W_1 e^{j\rho} + W_2 e^{-j\rho}) + |X|e^{\arg(X)}.
\]

In the above equation, \( \tilde{r}(t) \) denotes the lowpass (complex) representation of the bandpass signal \( r(t) \) and is related to \( r(t) \) as

\[
r(t) = \tilde{r}(t)e^{j2\pi f_1 t} + \tilde{r}(t)e^{-j2\pi f_1 t}.
\]

After trigonometric manipulation, \( v(t) \) can be written in terms of its in-phase and quadrature components, in matrix form as

\[
v(t) = D[\tilde{r}(t) + c] + d,
\]

where we have defined

\[
e = [(1/2)\beta_D \cos(\rho), (1/2)\beta_D \sin(\rho)]^T,
\]

\[
d = [X_1 \chi Q]^T,
\]

\[
D = \begin{bmatrix}
\sin(\varphi_s/2) & \cos(\varphi_s/2) \\
\cos(\varphi_s/2) & \sin(\varphi_s/2)
\end{bmatrix}.
\]

C. Cascaded effect

By combining equations (7) and (15), we can write the cascaded effect of the modulator and the demodulator observed at the output of the demodulator as

\[
v(t) = D\left[ R(M(m(t) + a) + b) + c\right] + d
\]

\[
= DRM + DRDm + DRb + DC + d. \quad (17)
\]

The rotational matrix \( R \) represents the effect when there exists a frequency (\( \Delta f \)) and phase (\( \Delta \sigma \)) difference between the LOs of the modulator and demodulator. \( R \) is given by [11]

\[
R = \begin{bmatrix}
\cos(2\pi \Delta f t + \Delta \sigma) & \sin(2\pi \Delta f t + \Delta \sigma) \\
-\sin(2\pi \Delta f t + \Delta \sigma) & \cos(2\pi \Delta f t + \Delta \sigma)
\end{bmatrix}.
\]

Note the inseparability of the modulator and demodulator’s imbalance and offset errors when they are cascaded.

III. PRINCIPLES OF COMPENSATION

From the previous section it is seen that quadrature mismatches (gain and phase errors) in the analogue front-end can be modelled as a linear transformation on \( I \) and \( Q \) channels. In principle, when this linear transform is invertible it should always be possible to compensate for these imbalances, by applying the inverse of the linear transform to the \( I \) and \( Q \) channels [12]. It is seen that when \( \varphi \neq \pm \pi/2 \) and \( g_M \neq 0 \), the matrix \( M \) is invertible. The same is true for \( \varphi, g_D \) and the imbalance matrix \( D \). If the imbalance parameters could thus be extracted in some way, it would be possible to apply the inverse matrices to counter the effects of the imbalanced modulator and demodulator.

These inverse, or compensation, matrices can computed in the digital domain and applied to the signals in the digital domain of the transceiver. The same is true for offset compensation. If the exact DC offset values can be extracted they can simply be subtracted from the signals in the digital domain.

The compensation networks for the modulator and demodulator is shown in Fig. 3 and Fig. 4 respectively. In these figures subscripts denote individual elements in the specified matrix or vector. The variable \( n \) denotes the discrete time index, since the compensation is done in the digital domain of the transceiver. The matrices \( M^{-1} \) and \( D^{-1} \) denote the inverse of the estimates of the imbalance matrices \( M \) and \( D \), respectively. The vector \( f \) is the estimate of the DC offset vectors \( Ma + b \). The vector \( e \) is the estimate of the DC offset vectors \( c + D^{-1}d \).

The use of these compensation networks assume that imbalance compensation (\( M^{-1} \) and \( D^{-1} \)) is applied before DC offset extraction takes place. Once the DC offset extraction (\( f \) and \( e \)) is done, they are applied as indicated in Fig. 3 and Fig. 4.

IV. IMBALANCE EXTRACTION FROM RELATIVE SIDEBAND MEASUREMENTS

A. The relative sideband ratio

The relative sideband ratio is the ratio of the unwanted image signal, to the desired one. Using equation (6), the relative sideband ratio of the modulator, \( S_M \), is defined as

\[
S_M = \frac{V_2}{V_1}.
\]

Using equation (13), the relative sideband ratio of the demodulator, \( S_D \), is defined as

\[
S_D = \frac{W_2^*}{W_1}.
\]
In this section it will be shown that the imbalance parameters of the modulator and demodulator can be extracted from measuring the relative sideband ratio of each.

To measure this ratio, a signal is needed which could easily be decomposed into a desired and an image component. A single-sideband, single-tone signal is ideal for this purpose. Although it is difficult to distinguish the desired from the unwanted signal in the time domain, it is quite straightforward in the frequency domain. If it is thus possible to transmit or receive what is supposed to be a single sideband tone, a spectral analysis of the signal after it has passed through either the imperfect modulator or demodulator, will render enough information to extract the phase error and gain imbalance.

Once the relative sideband ratio is measured, the contribution of phase imbalance and gain error must be separated. This issue is addressed next.

B. Imbalance extraction with phase information

We begin by observing what happens to the relative sideband ratio when the phase imbalance tends towards zero. Let this be denoted by \( S_{D,\beta} \), which is given by

\[
S_{D,\beta} = \lim_{\varphi \to 0} \frac{W_2}{W_1} = \frac{g_D - 1}{g_D + 1}. \tag{21}
\]

The magnitude and phase of \( S_{D,\beta} \) is given by

\[
|S_{D,\beta}| = \frac{|g_D - 1|}{g_D + 1} \tag{22}
\]

and

\[
\angle S_{D,\beta} = 0, \pi \text{ rad} \tag{23}
\]

respectively.

Now consider the relative sideband ratio when the gain of the I channel approaches that of the Q channel in the limit. Let this be denoted by \( S_{D,\varphi} \), which can be written as

\[
S_{D,\varphi} = \lim_{g_D \to 1} \frac{W_2}{W_1} = -j \tan(\varphi/2). \tag{24}
\]

The magnitude and phase of \( S_{D,\varphi} \) is given by

\[
|S_{D,\varphi}| = \tan(\varphi/2) \tag{25}
\]

and

\[
\angle S_{D,\varphi} = \mp \frac{\pi}{2} \text{ rad} \tag{26}
\]

respectively.

The above results, indicate that when the phase and gain error are small (\( \varphi \to 0 \) and \( g_D \to 1 \)), their contributions toward the relative sideband spur can be considered to be orthogonal to each other. In this case it is possible to separate the effect of the gain imbalance from the phase imbalance, by examining the real and imaginary parts of the relative sideband ratio. Using the real part of the measured relative sideband ratio, an estimate of the gain imbalance can now be computed as

\[
\hat{g}_D = \frac{1 + \text{Re}\{S_D\}}{1 - \text{Re}\{S_D\}}. \tag{27}
\]

An estimate of the phase imbalance, \( \hat{\varphi} \), can be determined from the imaginary part of \( S_D \) as

\[
\hat{\varphi} = -2 \arctan(\text{Im}\{S_D\}). \tag{28}
\]

C. Imbalance extraction without phase information

In the previous section it was shown how the I/Q imbalance parameters can be estimated from measuring the complex relative sideband. The implementation of such a method is practical for a demodulator, since the processing can be done at baseband in the digital domain using the complex FFT.

When one wants to apply the same techniques to a modulator to extract its I/Q imbalances, the procedure becomes more difficult. The transmitter can easily be used to generate a complex single-sideband tone in the digital domain and up-convert it to passband. Although the same desired and image spurs will be visible in the passband, spectral analysis may be limited to only magnitude measurements, thus losing all phase information.

An alternative is to use a receiver to mix the passband signal down to baseband, where it could be converted into the digital domain and analysed there. Even when a perfect demodulator is used (recall the cascading effect of section II-C), they may not share a LO and thus the lack of phase coherence will render any phase measurements useless.

When the phase of the relative sideband ratio cannot be measured or trusted, only the power of the relative sideband ratio can be used to extract the imbalance parameters. The power of the relative sideband ratio of the modulator (which has exactly the same form as that of the demodulator), \( P_M \), is given by

\[
P_M = \frac{|V_2|^2}{|V_1|^2} = \frac{g_M^2 + 1 - 2g_M \cos(\phi)}{g_M^2 + 1 + 2g_M \cos(\phi)}. \tag{29}
\]

The relationship between \( P_M \), \( g_M \) and \( \phi \) is shown graphically in Fig. 5. With only one power measurement, there exists an infinite number of combinations of gain and phase errors which could have resulted in the measured relative sideband ratio.
power. By changing the gain of either the $I$ or $Q$ channel of the modulator by a known quantity and doing another power measurement, the difference in the power of the relative sideband power can be measured. From Fig. 5 it is seen that knowledge of the amount by which the gain imbalance was changed as well as the difference it made to $P_M$, is enough to determine the exact value of $g_M$ as well as the absolute value of $\varphi$. The sign ambiguity of $\varphi$ is due to the symmetry of $P_M$, but can be determined by choosing the sign that results in the best reduction of the power of the relative sideband ratio, when used for compensation.

V. DC OFFSET EXTRACTION FROM SPECTRAL MEASUREMENTS

When complex spectral measurements can be made, then the in-phase and quadrature components of any DC offset that is present in the signal can be read directly from the DC bin value. As with the imbalance extraction without phase information, more than one measurement is needed to determine the in-phase and quadrature components of the DC offset, when the phase of the DC spur cannot be measured.

A simple algorithm to determine the phase of the DC spur measurement, denoted by $S_{DC}$, requiring three measurements is presented here. We begin by making the first observation $O_1$, of the magnitude of the DC bin,

$$O_1 = |S_{DC}|e^{\jmath \angle S_{DC}} = |S_{DC}|. \quad (30)$$

Since no information concerning the phase is available, one may arbitrarily choose to compensate with any phase angle. To simplify the mathematical derivation, a phase angle of zero radians is chosen. This implies that we assume that the DC offset spur is only due to offset in the $I$ channel. The second observation, $O_2$, should now exhibit a change in the magnitude of the DC spur. We thus have

$$O_2 = |S_{DC}|e^{\jmath \angle S_{DC}} - O_1 = \sqrt{O_1^2 - 2O_2O_1 \cos(\angle S_{DC}) + O_2^2} \quad (31)$$

where the fact that $|S_{DC}| = O_1$ was used in the last step. From equation (31) the phase of $S_{DC}$ can be determined as

$$\angle S_{DC} = \pm \arccos \left( \frac{2 - (O_2/O_1)^2}{2} \right). \quad (32)$$

Note that there exists an ambiguity on the sign of the phase angle. The sign which gives the best suppression of the DC spur is to be used.

VI. A SOLUTION TO TRANSCEIVER COMPENSATION

The compensation techniques presented in the previous sections, can be used to accomplish automatic transceiver imbalance and offset compensation. For transceiver compensation, a topology is used that was suggested in [12, pp. 113] and is shown in Fig. 6. This topology facilitates the use of a test tone to calibrate the demodulator and then proceeds to use the compensated demodulator as a measuring device for modulator compensation. Although the hardware topology corresponds to that of [12], the compensations and extraction algorithms that are used, are novel.

The assumption for using this topology is that there exists some common frequency band between the modulator and demodulator, since the demodulator serves as a measuring platform for extracting modulator imbalance and offset errors. Frequency and phase coherence between the LOs of the modulator and demodulator do not need to be established and are not assumed here.

The switch positions for the different stages of transceiver compensation are listed in Table I. The demodulator’s imbalance and offset compensation is done using a calibration test tone in the frequency band of interest, along with the techniques discussed in section IV-B. Once the demodulator has been corrected, it is used to down-convert a single sideband tone generated by the modulator. Modulator compensation is accomplished using the techniques discussed in section IV-C. If coherence between the LOs of the modulator and demodulator can be established, then the techniques discussed in section IV-B are also applicable to modulator compensation.

<table>
<thead>
<tr>
<th>Switch positions for transceiver compensation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demodulator compensation</td>
</tr>
<tr>
<td>Switch 1</td>
</tr>
<tr>
<td>Switch 2</td>
</tr>
</tbody>
</table>

VII. SIMULATION RESULTS

In order to test the imbalance and offset extraction and compensation techniques presented in this paper, a simulation platform was developed using MATLAB Simulink. The topology was the same as that presented in section VI. A phase error of 2.5° and a gain error of 1.2 were introduced in both the modulator and demodulator. DC offsets were also added to each. A random phase difference was introduced between the LOs of the modulator and demodulator.

The digital to analogue conversion was done with 12 bits precision, while the analogue to digital conversion used 14 bits. Internally, the algorithms were performed using IEEE double-precision floating-point numbers. Additive white Gaussian noise was added to simulate measurement noise. The FFT length used in the spectral analyses was 1024 samples.
Fig. 7 shows the resultant baseband spectrum before and after demodulator compensation is performed. It can be seen that the image frequency of what is supposed to be a single sideband tone, is not sufficiently suppressed before compensation. The unwanted DC offset can also be observed. After two iterations of demodulator compensation, the DC offset and unwanted image component are completely suppressed.

Fig. 8 shows the resultant baseband spectrum before and after modulator compensation is performed. The spectrum is measured by the fully compensated demodulator. To be able to observe the true performance of the modulator compensation, the demodulator’s imbalance and DC offsets were removed in the simulation, during modulator compensation. As with the demodulator, it is seen that the unwanted DC offset and image frequency components are completely suppressed, but this time after only one compensation iteration.

VIII. CONCLUSION

In this paper, novel compensation techniques were presented which may be used to compensate for modulator and demodulator imbalance and offset errors. These techniques rely on direct spectral measurements of the relative sideband ratio and DC spurs and not on iterative search techniques, to find the optimal compensation parameters. Simulation results show that the image and DC spur can be suppressed to the point where they can no longer be distinguished from noise. It can thus be concluded that the quadrature mixing architecture’s use in modern radio transceivers does not have to be dismissed due to its sensitivity towards gain, phase or DC offset errors.

REFERENCES


Josias J. de Witt (main author) was born in Pretoria, South Africa, in 1982. He obtained his B.Eng. degree from the University of Pretoria in 2004 (with distinction). He is presently studying towards an M.Sc.Eng degree at the University of Stellenbosch and is part of Telkom’s Centre of Excellence (CoE) program.

Gert-Jan van Rooyen obtained his Ph.D. (Eng) degree from Stellenbosch University, in 2005. He is currently a lecturer at the same university.