CX74017

On the Direct Conversion Receiver

Abstract

Increased pressure for low power, small form factor, low cost, and reduced bill of materials in such radio applications as mobile communications has driven academia and industry to resurrect the Direct Conversion Receiver (DCR). Long abandoned in favor of the mature superheterodyne receiver, direct conversion has emerged over the last decade or so thanks to improved semiconductor process technologies and astute design techniques. This paper describes the characteristics of the DCR and the issues it raises.

Introduction

Very much like its well-established superheterodyne receiver counterpart, introduced in 1918 by Armstrong [1], the origins of the DCR date back to the first half of last century when a single down-conversion receiver was first described by F.M. Colebrook in 1924 [2], and the term *homodyne* was applied. Additional developments led to the publication in 1947 of an article by D. G. Tucker [3], which first coined the term *synchrodyne*, in a receiver, which was designed as a precision demodulator for measurement equipment rather than a radio. Another paper by the latter in 1954 [4] reports the various single down-conversion receivers published at the time, and clarifies the difference between the homodyne (sometimes referred to as *coherent detector*) and the synchrodyne receivers: the former obtains the Local Oscillator (LO) directly, for example, from the transmitter, whereas the latter synchronizes a free-running LO to the incoming carrier.

Over the last decade or so, the drive of the wireless market and enabling monolithic integration technology have triggered research activities on DCRs, which integrated with the remaining analog and digital sections of the transceiver, has the potential to reach the “one-chip radio”. Besides, it favors multi-mode, multi-standard applications and constitutes thereby another step towards *software radio*.

The present article often refers to several recent publications [5-6], providing a thorough survey and insight, and displaying the renewed interest for DCRs.

Overcoming some of the problems associated with the traditional superheterodyne and being more prone to integration, DCR has nevertheless an array of inherent challenges. After a brief description of alternative and well-established receiver architectures, this paper presents the direct conversion reception technique and highlights some of the system-level issues associated with DCR.

Traditional Reception Techniques

The Superheterodyne

The superheterodyne, or more generally heterodyne[1], receiver is the most widely used reception technique. This technique finds numerous applications from personal communication devices to radio and TV tuners, and has been tried inside out and is therefore well understood. It comes in a variety of combinations [7-9], but essentially relies on the same idea: the RF signal is first amplified in a frequency selective low-noise stage, then translated to a lower intermediate frequency (IF), with significant amplification and additional filtering, and finally downconverted to baseband either with a phase discriminatory or straight mixer, depending on the modulation format. This is illustrated in the generic line-up of Figure 1.

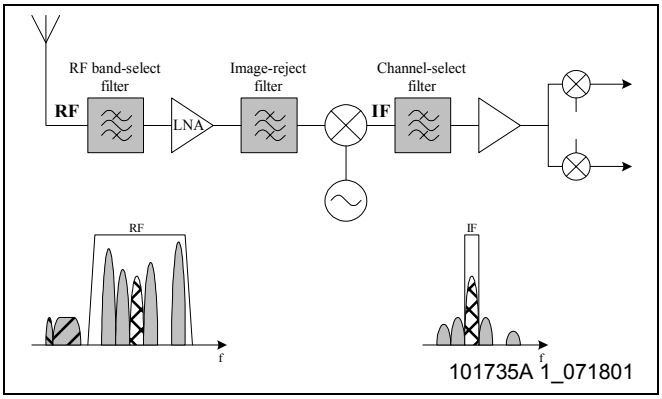


Figure 1. The Superheterodyne Receiver

Superheterodyning entails several trade-offs. Image rejection is a prevailing concern in this architecture. During the first

[1] Homo: Greek from “homos” - same; Hetero: Greek from “heteros” – other; Synchro: Greek from “sunkhronos” – same time; Dyne: Greek from “dunamis” – power.

downconversion to IF, any unwanted activity at a frequency spaced at f_{IF} offset from f_{LO}, on the opposite side of f_{LO} from the desired RF channel, produces a mixing product falling right into the downconverted channel, at f_{IF}. In practice, a RF bandpass filter, usually a Surface Acoustic Wave (SAW) device, is utilized to perform band selection ahead of the Low Noise Amplifier (LNA), while a second filter follows the LNA to perform image rejection. If these filters are identical, then, in fact, they share the burden of the two functions. But some amount of image rejection must particularly follow the LNA, for without it, the LNA noise figure effectively doubles due to the mixing of amplified image noise into the IF channel. Instead of the RF SAW filter, other passive filtering technologies, such as dielectric or ceramic, can also be featured. It can be seen from Figure 2 that the higher the IF, the more relaxed the requirements on the image reject filter cut-off frequency. Once at the IF, the presence of an interfering signal in the vicinity of the channel mandates sharp filtering around the channel; this is performed after the first mixer by the channel select filter, which is also often an IF SAW filter. Essentially, the exercise is that of a carefully engineered balance among several variables:

- Rejection provided by the various filters
- Frequency planning
- Linearity of the active stages

Dual IFs provide additional room to maneuver with filter selectivity, but complicate the frequency planning somewhat.

The selectivity required of the two aforementioned filters, in terms of fractional bandwidth, makes them unsuitable candidates in the foreseeable future for integration. This is because of low Qs of current silicon processes and the need to be implemented by bulky, off-chip components. The IF channel filter especially requires high-Q resonators for its implementation: the higher the IF, the lesser the filter's fractional bandwidth, that is, its ratio of bandwidth to center frequency, necessitating ever-higher Q. This high-Q requirement is most commonly met by the use of a piezoelectric SAW and crystal filters. This introduces additional constraints, as those filters require often-inconvenient terminating impedances, and matching may impinge on such issues as noise, gain, linearity, and power dissipation of the adjoining active stages. The narrower the fractional bandwidth, the more likely that the filter's passband shape will exhibit an extreme sensitivity to variations in matching element values. Additionally, the specificity of the IF filter to the signal bandwidth and hence the standard used, makes superheterodyne receivers unsuitable for multi-standard operation.

Nonetheless, superheterodyne is praised for its high selectivity and sensitivity.

Image-Reject Receivers

Alternatively, by a smart use of trigonometric identities, the image can be removed without the need of any post-LNA image-reject filtering. This is the principle of image-reject receivers [8] and [10]. The first is the Hartley architecture, introduced in [11] in 1928, and shown in Figure 3. It uses two mixers with their local oscillators in a quadrature phase relationship. This separates the IF signal into in-phase (I) and quadrature (Q) components. It then shifts the Q component by 90° before recombining the two paths. This is where the desired signal, present in both paths with identical polarities, is reinforced, while the image, present in both paths with opposite polarities, is cancelled out. The dual of the Hartley architecture, known as the Weaver image-reject receiver [12], achieves the relative phase shift of one path by 90º by the use of a second LO enroute to another IF or to baseband, see Figure 4. The same result is achieved.

However, the reliability of these receivers depends heavily on the accuracy of the I/Q paths, that is, the gain and phase imbalance between the two branches.

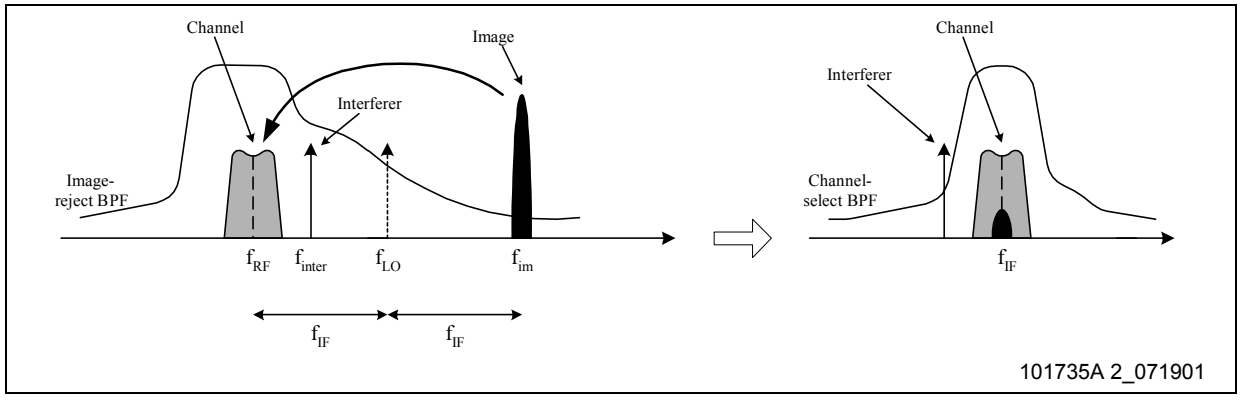


Figure 2. Image-Rejection and Selectivity in a Superheterodyne Receiver (High-Side LO Injection)

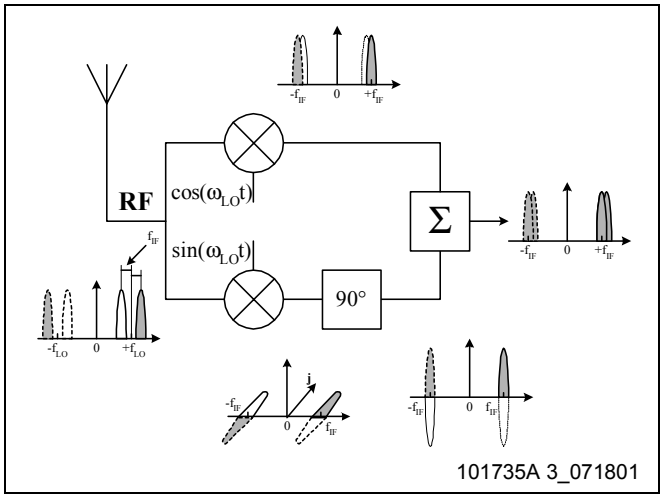


Figure 3. Hartley Image-Reject Architecture

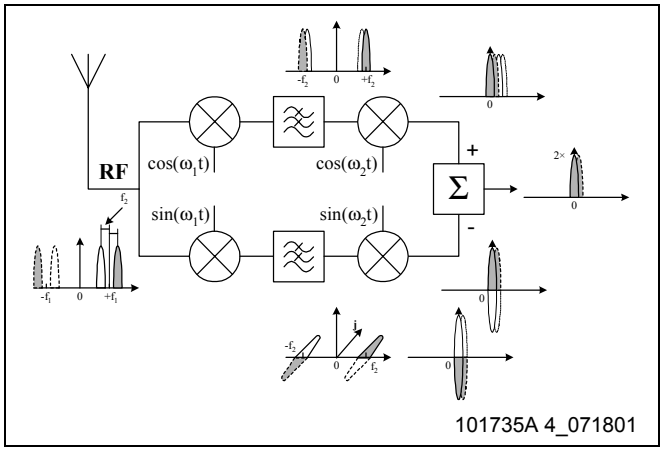


Figure 4. Weaver Image-Reject Architecture

Low IF Single Conversion Receiver

Low IF single conversion, see Figure 5, is an offspring of the DCR, and is covered in the following paragraphs. Its main purpose is to protect the receiver from all the direct current (DC)-related obstacles that pertain to DCR, while retaining the DCR's benefit of the elimination of high-Q IF filters. As its name indicates, instead of directly converting the signal to baseband, the LO is slightly offset from the RF carrier, typically one to two channels. The low IF means that the fractional bandwidth of the IF bandpass filtering is large, making it possible to implement it with low-Q components. The IF SAW, or crystal filter needed in the high IF case, can be replaced with an active resistive capacitor (RC) filter or other filter suitable for low frequency operation that is also conducive to silicon integration. The low IF signal may be translated to baseband through another mixer, or preferably, in the digital domain following analog-to-digital (A/D) conversion. Of course, this comes at the expense of faster and higher resolution analog-to-digital converters (ADCs). If the IF frequency is equal to only one or two channel widths, then it is not possible to provide image-rejection at radio frequency (RF), as the RF filter must be wide enough to pass all channels of the system. In this case, all image rejection must come from the quadrature downconversion to the low IF, which itself resembles the Hartley architecture, once the baseband conversion is added.

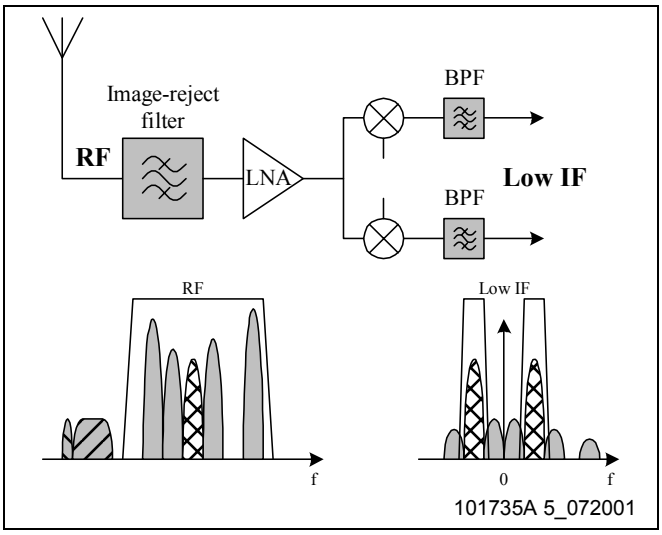


Figure 5. Low IF Single-Conversion Receiver

Wideband IF with double conversion

This architecture, shown in Figure 6, is very similar to the superheterodyne. In this case, the first mixer utilizes an LO that is at a fixed frequency, and all RF band channels are translated to IF, retaining their positions relative to one another. The second mixer utilizes a tunable LO, thus selecting the desired channel to be translated to baseband. A subsequent lowpass filter suppresses adjacent channels.

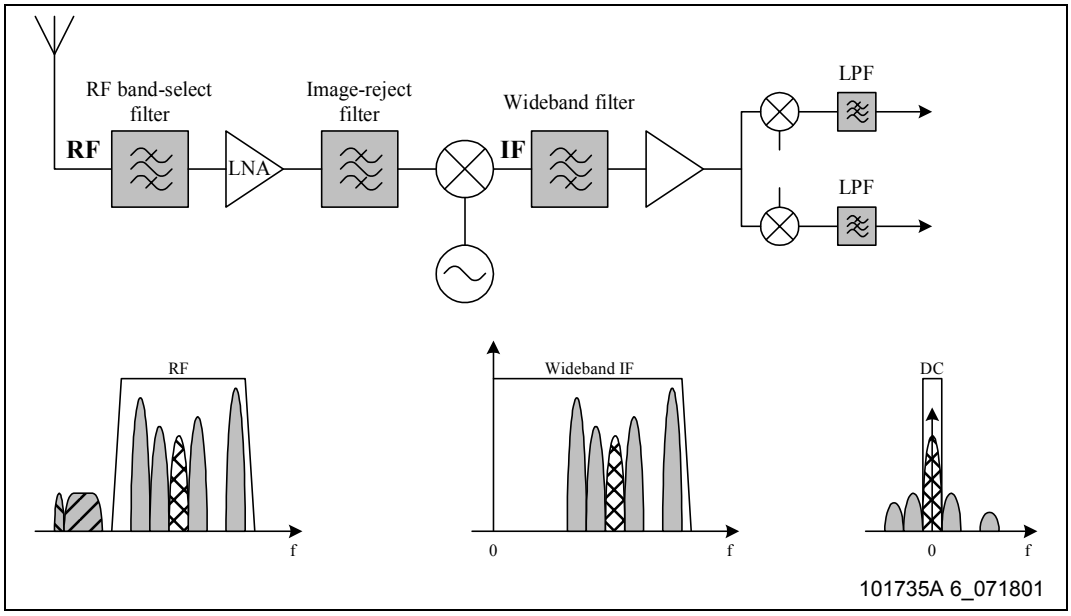


Figure 6. Wideband IF with Double Conversion.

Direct Conversion Receivers

Direct conversion reception, also referred to as homodyne, or zero-IF, shown in Figure 7, is the most natural solution to receiving information vehicled by a carrier. However, it has only been a decade or so that this type of reception has found applications other than pagers, for example [13]. For it has several qualities which make it very suitable for integration as well as multi-band, multi-standard operation, but severe inherent obstacles that had for long kept it in the shadow of the superheterodyne.

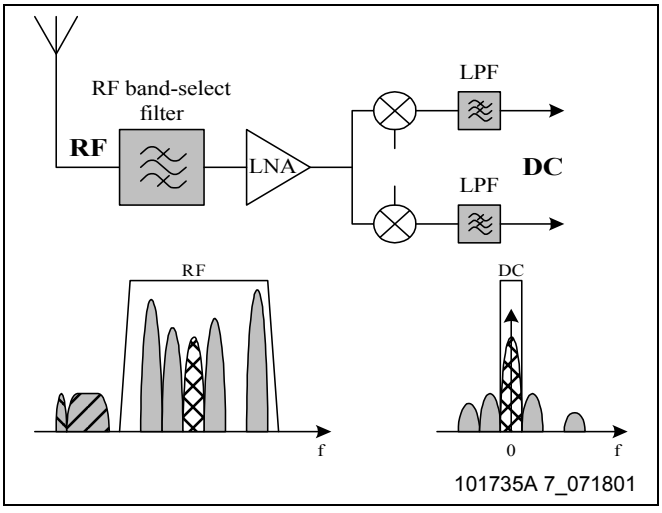


Figure 7. The Direct Conversion Receiver

First, the image problem is no longer present, since the IF is zero and the image to the desired channel, for all but single-sideband signals, is the channel itself. Then, only one LO is needed, which means only one phase noise contribution. The need for the bulky, off-chip filters is consequently removed. Filtering now only occurs at low, that is, baseband, frequencies with some amplification. This means less current consumption than at higher frequencies (to drive device parasitics), fewer components, lower cost, etc. Practically, however, strong out-of-band interferers or blockers may need to be removed prior to downconversion in order to avoid desensitizing the receiver by saturating subsequent stages, as well as producing harmonics and intermodulation terms, which then appear in baseband. Such a filter may be placed after the LNA, for example.

DCR, however, brings its own set of issues. The following paragraphs describe those in more detail.

DC offsets

In direct conversion, as the signal of interest is converted to baseband very early in the receive chain, without any filtering other than RF band-selection, various phenomena contribute to the creation of DC signals, which are directly appearing as interfering signals in the band of interest.

The LO may be conducted or radiated through an unintended path to the mixer's RF input port, thus effectively mixing with itself, producing an unwanted DC component at the mixer output, see Figure 8.

Worse still, this LO leakage may reach the LNA input, producing an even stronger result. This effect presents a high barrier against the integration of LO, mixer, and LNA on a single silicon substrate, where numerous mechanisms can contribute to poor isolation. These include substrate coupling, ground bounce, bond wire radiation, and capacitive and magnetic coupling.

Conversely, a strong in-band interferer, once amplified by the LNA, may find a path to the LO input port of the mixer, see Figure 9, thus producing self-mixing, once again.

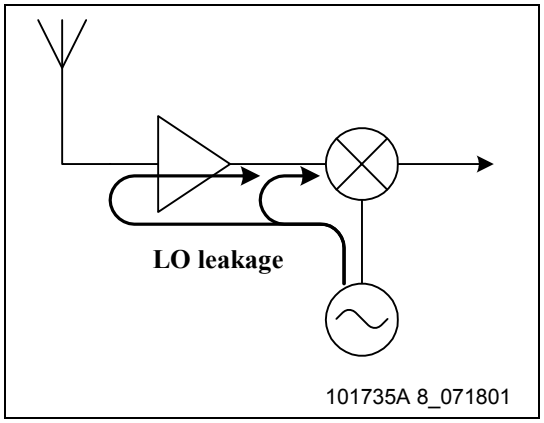


Figure 8. LO Leakage

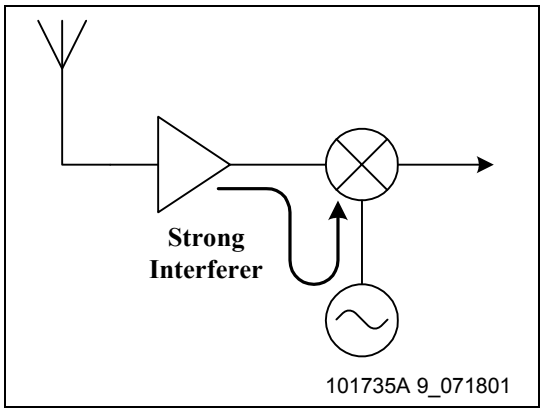


Figure 9. Strong In-Band Interferer, Amplified by the LNA

Some amount of LO power is conducted through the mixer and LNA, due to their non-ideal reverse isolation, to the antenna. The radiated power, appearing as an interferer to other receivers in the corresponding band, may violate emissions standards of the given system.

It is important to note that since the LO frequency is inside the receive band, the front-end filters do nothing to suppress this LO emission. Additionally, the radiated LO signal can then be reflected by buildings or moving objects and re-captured by the antenna, as shown in Figure 10. This effect, however, is not of significant importance compared to the previously mentioned LO self-mixing and blocker self-mixing.

LO or RF signal leakage to the opposite mixer port is not the only way in which unwanted DC can be produced. Any stage that exhibits even-order non-linearity also generates a DC output. This is covered in more detail later.

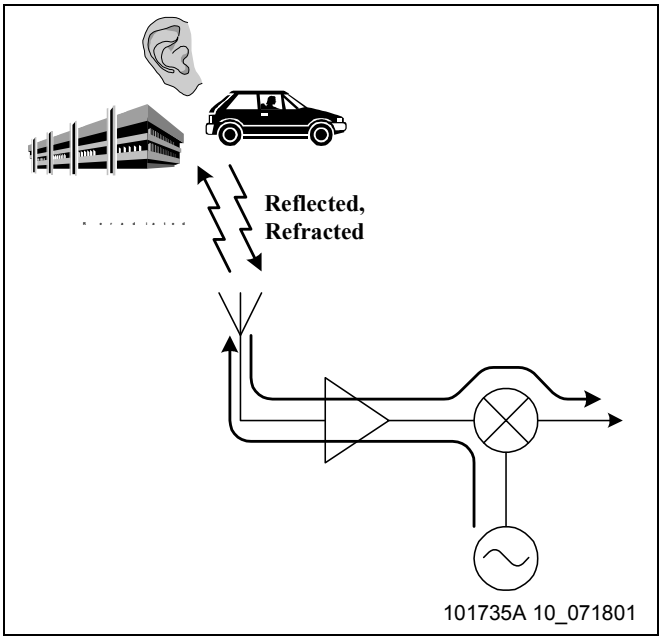


Figure 10. Emission, Reflection, and Refraction of LO Signal via Antenna

Whether or not the DC product desensitizes the receiver depends on the system type. Obviously, it is preferable to alternate current (AC)-couple at the mixer output to eliminate the DC. Some modulation schemes used in paging applications, such as Frequency-Shift Keying (FSK), show little degradation if low frequency spectrum components are filtered out, see Figure 11.

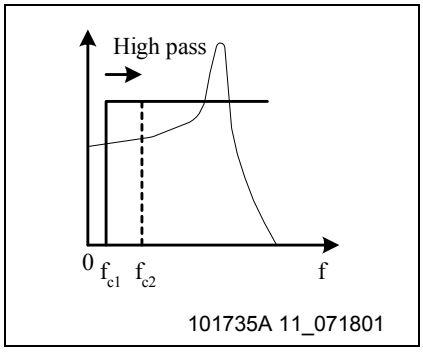


Figure 11. High-Pass Filtering of the Modulation Spectrum

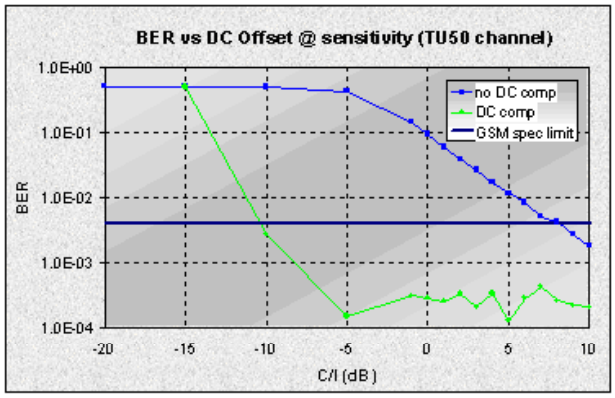
However, other modulation schemes present a peak at DC, and capacitive AC-coupling will infer significant information loss, hence considerably degrading the BER. In TDMA systems such as GSM, although there is no significant low frequency spectral peak, it still becomes impossible to AC couple. This is because of the conflicting requirements on an AC-coupling capacitor in a Time-Division Multiple Access (TDMA) system: the capacitor must be large enough to avoid causing a wide notch at DC, but it must be small enough that all transients settle out upon receiver power-up (every frame) before data reception begins.

In TDMA receivers that cannot be AC coupled, the idle time slot, that is, just before reception, can still be put to good use by storing the value of the offset on a capacitor and then subtracting it from the signal path during the burst. This is exactly the same method which is normally used to correct DC offsets occurring at the second mix of superheterodyne TDMA receivers, where this mix goes to baseband. In that case the only problem causing DC is LO self-mixing. In this method, the DC value produced by the receiver is obtained in a pre-measurement prior to the receive burst.

It is important when using this method, that the signal path prior to the mixer be opened during the DC pre-measurement, to prevent any large blocking signals from affecting the result. Blocking signals, which can appear at any time, most often induce variable or *wandering* offsets. The measurement-and-subtraction process cannot correct these offsets, because the blocking signals may appear during the measurement and not during the burst, or vice-versa. For blocking-induced DC, the most effective measures are the elimination of self-mixing paths and the maximizing of linearity to prevent the DC to start. Failing these, there is still the possibility of DC-correction after-the-fact in the digital signal processing (DSP) occurring at baseband.

DSP techniques can be used to remove the DC offset in TDMA systems, in a way that cannot be duplicated in the analog domain: a full timeslot of the received signal can be buffered, the mean of which is determined and then removed from each data point of the signal. The resulting signal has zero mean. For systems such as Global System for Mobile communications® (GSM®), an unwanted result of this is that any DC that is part of the signal is lost, but the typical effect of this is minimal. Figure 12 illustrates the use of such a method for a typical GSM receiver. This technique can be further refined by tracking the mean over portions of the burst, allowing the detection of sudden interferers or blockers and canceling their DC product only where it occurs.

Careful layout can also improve isolation.



101735A 12_071801

Figure 12. BER Improvement with DSP-Based DC Offset Cancellation

Non-Linearities

As mentioned previously, another DCR problem is non-linearity. Just as with the superheterodyne receiver, the DCR exhibits spurious responses. For the superheterodyne, these occur at RF input frequencies where

$$N(RF) \pm M(LO) = IF,$$

while for the DCR they occur where

$$N(RF) - M(LO) = IF$$

When a blocking signal's carrier falls on one of these spurious frequencies, the signal is translated to baseband with an attendant shift in its bandwidth, dependent on the spurious order.

However, more importantly, large blocking signals also cause DC in the DCR, whether on a spurious frequency or not. The DC is produced at the mixer output and amplified by the baseband stages. It is due primarily to second order mixer non-linearity, characterized by IP2 (second order intercept point), IM2 (second order intermodulation.) It can be alleviated by extremely well-balanced circuit design. However, only a short time ago, the mixer and LNA used to require a single-ended design because the antenna and a hypothetical preselect filter were usually single-ended.

In most systems, third order intermodulation is important, as it usually falls in-band, in the vicinity of the signals of interest, and is characterized by IP3 (third order intercept point). In direct conversion, the second order intermodulation becomes critical, as it produces baseband signals, which now appear as interfering signals in the down-converted desired signal. The second order non-linearity is measured by the IP2. IP2 is defined in the same manner as IP3 Figure 13.

Either a 2-tone, or 1-tone test can be performed, and the IP2 is defined by extrapolating the low-frequency beat tone in the former or the DC component in the latter, until it intercepts the fundamental curve. To illustrate in the case of a single tone test, the input signal is:

$$x(t) = A\cos(\omega t)$$

Assuming a non-linearity modeled by a polynomial:

$$y(x) = a_1 x + a_2 x^2 + a_3 x^3 + \ldots$$

$$y(x) = a_1 A\cos(\omega t) + a_2 A^2\left(\frac{\cos(2\omega t)+1}{2}\right) + \ldots$$

$$= \underbrace{\frac{a_2 A^2}{2}}_{DC} + \underbrace{a_1 A\cos(\omega t)}_{fundamental} + \frac{a_2 A^2}{2}\cos(2\omega t) + \ldots$$

It can be seen from these equations and in Figure 13 that the DC component due to the second order non-linearity is growing with twice the slope of the fundamental on a logarithmic scale. At the intercept point,

$$\frac{a_2 A^2}{2} = a_1 A \Leftrightarrow A = \frac{2a_1}{a_2} = IIP2$$

Due to the doubled slope of the second-order product,

$$IIP2 = Pin + \Delta \quad with \quad \Delta = Pout - IM2$$

Noise

Low frequency noise becomes a great concern in DCR [14], as significant gain is allocated to baseband stages after the mixer. Weak baseband signal levels of a few millivolts are still very vulnerable to noise. This requires stronger RF stage gain to alleviate the poor noise figure of baseband blocks, but of course this must be traded against the linearity problems, just described, that accompany higher RF gain.

Flicker noise, that is, 1/f noise, is the major baseband noise contributor. Associated with DC flow, it has a spectral response proportional to 1/f. In RF circuits, 1/f noise tends to be modulated onto the RF signal. In the case of a mixer with baseband output, 1/f noise sees especially high conversion gain. In practice, flicker noise becomes an issue for Metal Oxide Semiconductor (MOS) devices more than bipolar, and is modeled as a voltage source in series with the gate. 1/f noise complicates the use of MOS transistors for RF circuits, since the main method of reducing it in MOS is to increase the transistor's size, which increases the device capacitance. This adversely affects RF gain. For this reason, it is preferable to use bipolar transistors for DCR mixer designs. In the first baseband stages after the mixer, it becomes possible to use MOS devices, as the transistor-size tradeoff is feasible at low frequencies.

I/Q mismatches

Due to the high frequency of the LO, it is not possible to implement the IQ demodulator digitally. An analog IQ demodulator exhibits gain and phase imbalances between the two branches, as well as the introduction of DC offsets. Such imperfections distort the recovered constellation. Assuming α, φ being the amplitude and phase mismatch respectively between the demodulator quadrature ports, and the complex signal incident upon it having in-phase and quadrature components I and Q:

$$Iout = \left(I\cos(\omega t) + Q\sin(\omega t)\right)\cdot 2\cos(\omega t)$$
$$Qout = \left(I\cos(\omega t) + Q\sin(\omega t)\right)\cdot 2\left(1+\alpha\right)\sin(\omega t + \varphi)$$

Filtering out the high frequency terms:

$$Iout = I$$
$$Qout = \left(1+\alpha\right)\left(-I\sin(\varphi) + Q\cos(\varphi)\right)$$

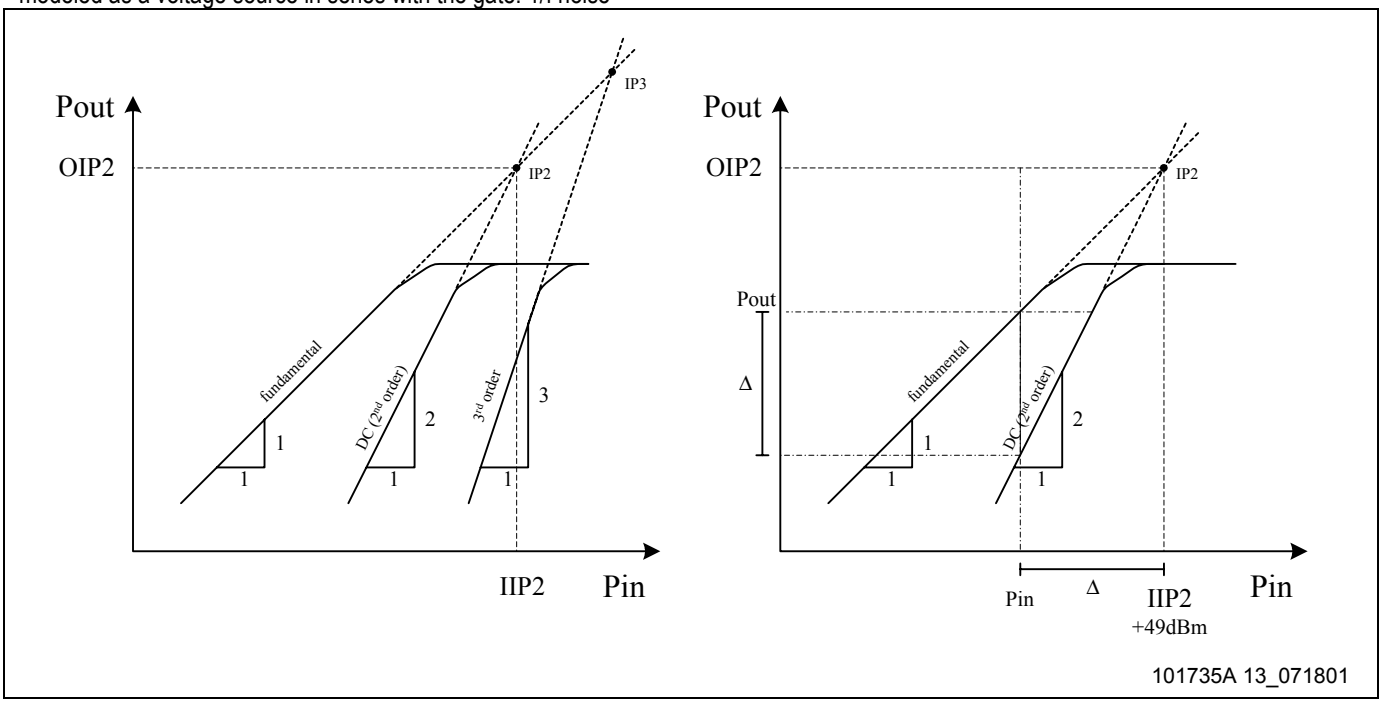


Figure 13. Second Order Intercept Point (IP2)

Figure 14 and Figure 15 sketch how this affects a given constellation diagram. In DCR systems however, the IQ matching is not as critical as in image-rejection architectures. Rather, it is only important insofar as the accuracy of the modulation is concerned.

Analog and digital (DSP based) calibration and adaptation methods have been described so as to correct for these imbalances, for example in [15].

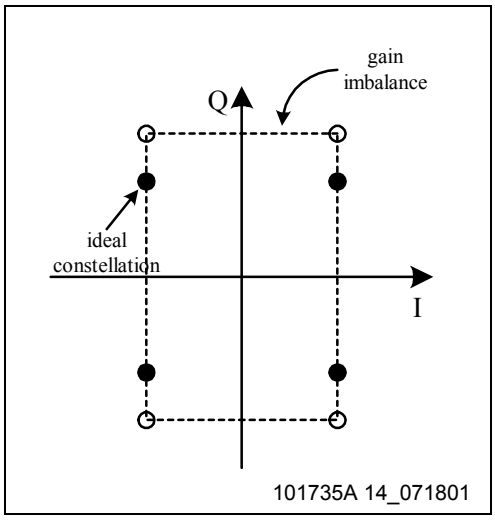


Figure 14. IQ Demodulator Imperfections (Gain Imbalance)

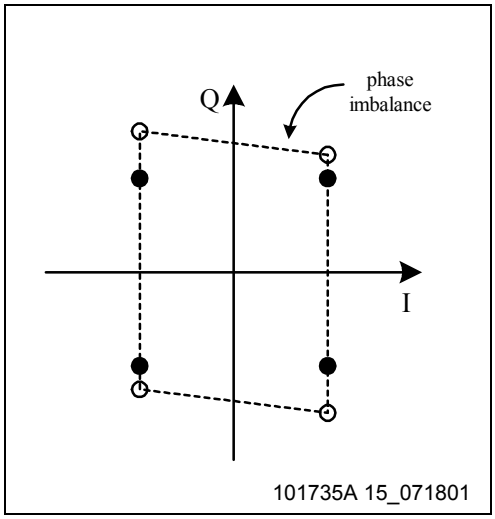


Figure 15. IQ Demodulator Imperfections (Phase Imbalance)

Conclusion

The direct conversion receiver is an attractive yet challenging receiving technique. It has recently been applied successfully to devices such as pagers, mobile phones, PC and internet wireless connectivity cards, satellite receivers etc. in a variety of process technologies and increasing integration levels. It is poised to appear in many more applications in the near future.

References

[1] L. Lessing, *"Man of High Fidelity: Edwin Howard Armstrong, a Biography"*, Bantam Books, New York, 1969.

[2] F.M. Colebrook, *"Homodyne"*, Wireless World and Radio Rev., 13, p. 774, 1924.

[3] D.G. Tucker, *"The Synchrodyne"*, Electronic Engng, 19, p. 75-76, March 1947.

[4] D.G. Tucker, *"The History of the Homodyne and the Synchrodyne"*, Journal of the British Institution of Radio Engineers, April 1954.

[5] A.A. Abidi, *"Direct-Conversion Radio transceivers for Digital Communications"*, IEEE Journal of Solid-State Circuits, vol. 30, no. 12, December 1995.

[6] B. Razavi, *"Design Considerations for Direct-Conversion Receivers"*, IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing, vol. 44, no. 6, June 1997.

[7] S.J. Franke, *"ECE 353 – Radio Communication Circuits"*, Department of Electrical and Computer Engineering, University of Illinois, Urbana, 1994.

[8] B. Razavi, *"RF Microelectronics"*, Prentice Hall, Upper Saddle River, NJ, 1998.

[9] J.C. Rudell et al., *"Recent Developments in High Integration Multi-Standard CMOS Transceivers for Personal Communication Systems"*, International Symposium on Low Power Electronics and Design, 1998.

[10] J.C. Rudell, *"Issues in RFIC Design"*, lecture notes, University of California Berkeley/National Technological University, 1997.

[11] R. Hartley, *"Single-Sideband Modulator"*, U.S. Patent no. 1666206, April 1928.

[12] D.K. Weaver, *"A Third Method of Generation and Detection of Single Sideband Signals"*, Proceedings of the IRE, vol. 44, p. 1703-1705, December 1956.

[13] I.A.W. Vance, *"Fully Integrated Radio Paging Receiver"*, IEE Proc., vol. 129, no. 1, p. 2-6, 1982.

[14] P.R. Gray, R.G. Meyer, *"Analysis and Design of Analog Integrated Circuits"*, Third edition, John Wiley & Sons, New York, 1993.

[15] J.K. Cavers, M. W. Liao, *"Adaptive Compensation for Imbalance and Offset Losses in Direct Conversion Transceivers"*, IEEE Transactions on Vehicular Technology, vol. 42, p. 581-588, November 1993.

© 2002, Skyworks Solutions, Inc. All Rights Reserved.

Information in this document is provided in connection with Skyworks Solutions, Inc. ("Skyworks") products. These materials are provided by Skyworks as a service to its customers and may be used for informational purposes only. Skyworks assumes no responsibility for errors or omissions in these materials. Skyworks may make changes to its products, specifications, and product descriptions at any time, without notice. Skyworks makes no commitment to update the information and shall have no responsibility whatsoever for conflicts, incompatibilities, or other difficulties arising from future changes to its products and product descriptions.

No license, express or implied, by estoppel or otherwise, to any intellectual property rights is granted by this document. Except as may be provided in Skyworks' Terms and Conditions of Sale for such products, Skyworks assumes no liability whatsoever.

THESE MATERIALS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, RELATING TO SALE AND/OR USE OF SKYWORKS™ PRODUCTS INCLUDING WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, PERFORMANCE, QUALITY OR NON-INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. SKYWORKS FURTHER DOES NOT WARRANT THE ACCURACY OR COMPLETENESS OF THE INFORMATION, TEXT, GRAPHICS OR OTHER ITEMS CONTAINED WITHIN THESE MATERIALS. SKYWORKS SHALL NOT BE LIABLE FOR ANY SPECIAL, INDIRECT, INCIDENTAL, OR CONSEQUENTIAL DAMAGES, INCLUDING WITHOUT LIMITATION, LOST REVENUES OR LOST PROFITS THAT MAY RESULT FROM THE USE OF THESE MATERIALS.

Skyworks™ products are not intended for use in medical, lifesaving or life-sustaining applications. Skyworks' customers using or selling Skyworks™ products for use in such applications do so at their own risk and agree to fully indemnify Skyworks for any damages resulting from such improper use or sale.

The following are trademarks of Skyworks Solutions, Inc.: Skyworks™, the Skyworks symbol, and "Breakthrough Simplicity"™. Product names or services listed in this publication are for identification purposes only, and may be trademarks of third parties. Third-party brands and names are the property of their respective owners.

Additional information, posted at www.skyworksinc.com, is incorporated by reference.

General Information:

Skyworks Solutions, Inc.
4311 Jamboree Rd.
Newport Beach, CA 92660-3007

www.skyworksinc.com

SKYWORKS™
BREAKTHROUGH SIMPLICITY