Monolithic RF Active Mixer Design

Keng Leong Fong, *Member, IEEE,* and Robert G. Meyer, *Fellow, IEEE*

***Abstract*—An overview of monolithic radio-frequency (RF) active mixer design is presented. The paper is divided into two parts. The first part discusses the performance parameters that are relevant to the design of downconversion mixers, and how they affect the system performance. The second part presents three common kinds of mixer topologies, namely, unbalanced, single-balanced, and double-balanced designs. This paper concentrates on active mixers only. The advantages and disadvantages, as well as the design and optimization techniques for the three kinds of mixers, are discussed.**

***Index Terms*—Analog integrated circuits, circuit optimization, MIMIC's, MMIC circuits, mixers, nonlinear circuits.**

I. Introduction

The rapid growth of portable wireless communication systems, such as wireless (cordless and cellular) phones, global positioning satellite (GPS), wireless local area network (LAN), etc., has increased the demand for low-cost and high-performance front-end receivers. This presents a challenge to radio frequency (RF) circuit designers to find optimal solutions for the realization of high-frequency (900 MHz to 3 GHz) receivers using low-cost plastic packages and high-volume silicon technologies.

Fig. 1 shows a typical RF receiver front-end architecture. The downconversion mixer is used to convert the RF signal down to an intermediate frequency (IF) by mixing the RF signal from the low-noise amplifier (LNA) with the local oscillator (LO) signal. This allows channel selection and gain control at lower frequencies where high quality-factor (Q) filters and variable-gain amplifiers can be constructed economically. Instead of using an IF filter with tunable passband frequency, an IF filter with fixed passband frequency is used, and the LO frequency is tuned to select the desirable channel. The LNA is used to amplify the RF signal to reduce the noise contribution from the mixer. The RF and image-rejection filters are used to reject undesired out-of-band signals. The downconversion mixer is a very important building block because its performance affects the system performance and the performance requirements of its adjacent building blocks, which include the LNA, LO, RF filter, image-rejection filter, and IF stages.

Manuscript received July 31, 1997; revised February 23, 1998. This material is based on work supported in part by the U.S. Army Research Office under Grant DAAH04-93-F-0200.

K. L. Fong was with the Electronics Research Laboratory, Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720 USA. He is now with Philips Semiconductors, Sunnyvale, CA 94088-3409 (e-mail: Keng.Fong@sv.vc.philips.com).

R. G. Meyer is with the Electronics Research Laboratory, Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720 USA.

Publisher Item Identifier S 1057-7130(99)01774-7.

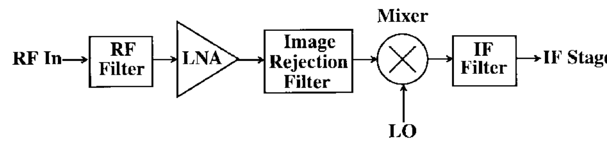


Fig. 1. RF front-end.

Downconversion mixers perform frequency conversion by using nonlinear elements in time-varying circuits. The nonlinear operation is difficult to describe analytically, and hence optimization becomes very difficult. In addition to magnitude changes and phase shifts as in a linear system, signals and noise also undergo frequency shifts in a time-varying nonlinear system. As a result, many active mixers realized in silicon technology are not optimized, and have relatively poor performance (high noise figure and poor linearity).

This paper is divided into two parts. This first part (Section II) discusses performance parameters relevant to the design of downconversion mixers, and how they affect the overall system performance. The second part (Section III) presents common mixer topologies with emphasis on their advantages and disadvantages, as well as the design and optimization techniques.

In terms of conversion gain, mixers can generally be categorized into passive and active mixers. Passive mixers, such as diode mixers [1], [2], and passive field-effect transistor (FET) mixers [3], [4], have no conversion gain. On the other hand, active mixers have conversion gain which acts to reduce the noise contribution from the IF stages. Since passive mixers are well documented, this paper concentrates on active mixers only. The emphasis is on the topologies that can be readily fabricated in integrated circuit technologies.

II. Performance Parameters

The parameters that affect the receiver performance can be divided into four categories, namely, sensitivity, selectivity, overloading, and power consumption. Sensitivity measures the smallest received signal the receiver needs to achieve a specific bit error rate (BER). It depends on the system noise figure of the receiver and the demodulation scheme used. Selectivity, which includes adjacent channel selectivity, image rejection, and out-of-band blocker rejection, measures the ability to detect the desired signal and to reject undesired signals. Adjacent channel selectivity depends on the third-order intermodulation performance of the LNA and downconversion mixer, the selectivities of the IF and baseband filters, and the LO phase noise. Image rejection depends on the selectivities of the RF and image-rejection filters. Out-of-band blocker rejection depends on the selectivity of the RF and image-rejection filters, the desensitization effect [5], [6] on the LNA and downconversion mixer by the blocker, and the LO phase noise. Overloading measures the largest desired signal the receiver can handle while maintaining a specific BER. It depends on the system 1-dB compression point $(P_{-1\,\mathrm{dB}})$ of the receiver. Power consumption determines the usage time of a portable receiver. Although the overall performance of a receiver depends on the performance parameters of all building blocks, this section concentrates on the performance parameters of the downconversion mixer only, and how they affect the system performance and the performance requirement of the adjacent building blocks.

A. Noise Figure

Noise figure (NF) is commonly used in communication systems to specify the noise performance of a circuit. It measures the signal-to-noise ratio (SNR) degradation caused by the circuit [2]. In communication systems where the source impedance is well defined, NF is defined as

$$\mathrm{NF} = \frac{N_s + N_i}{N_s} = 1 + \frac{N_i}{N_s} \tag{1}$$

where N_s and N_i are the noise power of the source impedance and the input-referred noise power of the circuit, respectively. The value of NF is meaningless if the source impedance is not specified. Noise figure is typically expressed in a decibel (dB) scale.

The system noise factor (noise figure expressed in linear scale) for the downconverter shown in Fig. 1 is

$$\begin{aligned}\mathrm{NF} &= \frac{1}{L_{\mathrm{RF}}} + \frac{\mathrm{NF}_{\mathrm{LNA}} - 1}{L_{\mathrm{RF}}} + \frac{1}{L_{\mathrm{RF}} G_{\mathrm{LNA}}}\left(\frac{1}{L_{\mathrm{IM}}} - 1\right) \\ &\quad + \frac{\mathrm{NF}_{\mathrm{MIX}} - 1}{L_{\mathrm{RF}} G_{\mathrm{LNA}} L_{\mathrm{IM}}} \\ &= \frac{1}{L_{\mathrm{RF}}}\left(\mathrm{NF}_{\mathrm{LNA}} + \frac{\mathrm{NF}_{\mathrm{MIX}} - L_{\mathrm{IM}}}{G_{\mathrm{LNA}} L_{\mathrm{IM}}}\right)\end{aligned} \tag{2}$$

where L_{RF} and L_{IM} are the insertion gain (less than 1 for passive filters) of the RF filter and the image-rejection filter, respectively; $\mathrm{NF}_{\mathrm{LNA}}$ and $\mathrm{NF}_{\mathrm{MIX}}$ are the noise figures of the LNA and the downconversion mixer, respectively; and GLNA is the power gain of the LNA. This equation assumes that the noise figures of the filters are the same as their insertion losses. Noise contribution from the IF stage is not included in this equation. As shown in (2), the LNA needs to have sufficient power gain to reduce the noise contribution from the mixer. Hence, a mixer with low noise figure is highly desirable in order to relax the gain requirement of the LNA.

There are two types of noise figure measures for downconversion mixers, namely, single-sideband (SSB) noise figure and double-sideband (DSB) noise figure. The single-sideband noise figure is applicable to the heterodyne architecture where the RF signal is converted to an IF which is the higher than one-half of the image-rejection filter bandwidth. Fig. 2 shows how the LO signal and its harmonics mix noise at various frequencies to the IF. The term “single-sideband” is derived from the fact that only one of the sidebands (the RF band) of the LO signal is converted to the IF (the image band is rejected). On the other hand, the double-sideband noise figure is applicable to the homodyne (direct conversion) architecture [7] where the RF signal is converted to the baseband directly. Fig. 3 shows how the LO and its harmonics mix noise at various frequencies to the baseband. The term “double sideband” is derived from the fact that two sidebands of the LO signal are converted to the baseband (LO frequency is in the middle of the RF band). Comparing Figs. 2 and 3, it is obvious that the mixer in the heterodyne architecture has twice as many noise contributors as that in the homodyne architecture. Hence, the single-sideband noise power is about 2 times (3 dB) higher than the double-sideband noise power. It is important to notice that a factor of 2 difference in input-referred noise power does not translate to 3-dB difference in noise figures because

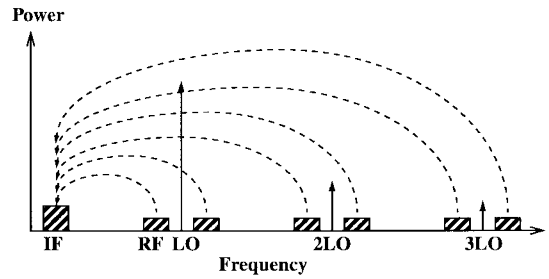


Fig. 2. LO mixes noise to the IF.

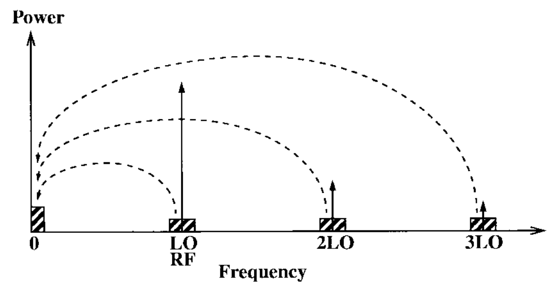


Fig. 3. LO mixes noise to the baseband.

$$\left(1 + \frac{2N_{\mathrm{DSB}}}{N_s}\right) < 2\left(1 + \frac{N_{\mathrm{DSB}}}{N_s}\right) \Rightarrow \mathrm{NF}_{\mathrm{SSB}} < 2\mathrm{NF}_{\mathrm{DSB}}$$

where N_{DSB} is the input-referred double-sideband noise power of the mixer, $\mathrm{NF}_{\mathrm{SSB}}$ and $\mathrm{NF}_{\mathrm{DSB}}$ are the single-sideband and double-sideband noise figures of the mixer, respectively. However, if N_{DSB} is much larger than N_s, $\mathrm{NF}_{\mathrm{SSB}}$ is about 3 dB higher than $\mathrm{NF}_{\mathrm{DSB}}$.

B. Conversion Gain

A downconversion mixer should provide sufficient power gain to compensate for the IF filter loss, and to reduce the noise contribution from the IF stages. However, this gain should not be too large as a strong signal may saturate the output of the mixer. Typically, power gain, instead of voltage or current gains, is specified. The reason is that noise figure is a power quantity, and hence it is easier to translate the NF of the IF stages to the system NF using power gain. Power gain (G) is

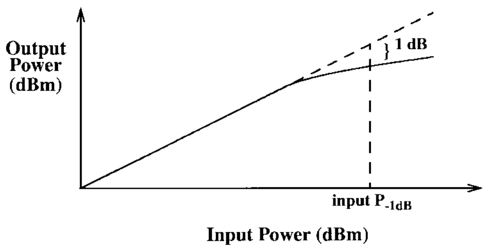


Fig. 4. Magnitude response of a mixer.

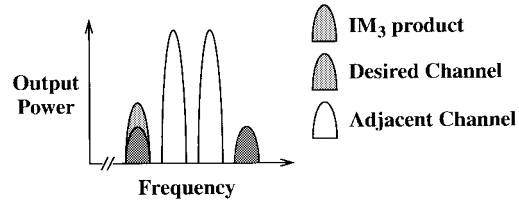


Fig. 5. Third-order intermodulation product corrupts desired channel.

related to voltage or current gain by

$$G = \left(\frac{V_O}{V_I}\right)^2 \frac{R_S}{R_L} = \left(\frac{I_O}{I_I}\right)^2 \frac{R_L}{R_S} \tag{3}$$

where V_O and V_I are output and input voltages, respectively; I_O and I_I are output and input currents, respectively; R_L and R_S are load and source resistance, respectively. Although increasing the load resistance by a factor of 2 can increase the voltage gain by 6 dB, the power gain is increased by only 3 dB.

C. Gain Compression

A strong signal can saturate a mixer and reduce its power gain. The input 1 dB compression point $(P_{-1\,\mathrm{dB}})$ measures the input power level that causes the mixer to deviate from its linear magnitude response by 1 dB. Fig. 4 shows the magnitude response of a mixer as a function of input signal power. The dotted line shows the linear magnitude response of an ideal mixer. Due to odd-order nonlinearities and limiting (current limiting and/or voltage headroom limiting), the conversion gain of an actual mixer is reduced at high input power level as shown by the solid line. The conversion gain of the mixer is the ratio of output power to input power. The point where the large-signal gain is 1 dB below the small-signal gain is the $P_{-1\,\mathrm{dB}}$. In the case where gain compression is caused by limiting, the gain drops abruptly and the output power stays constant as the input signal power exceeds the input $P_{-1\,\mathrm{dB}}$. In the case where gain compression is caused by the odd-order nonlinearities in the transfer functions of the devices used, the gain decreases more gradually as the input signal power exceeds the input $P_{-1\,\mathrm{dB}}$.

If the input power of the desired signal is larger than the input $P_{-1\,\mathrm{dB}}$, the desired signal can be distorted at the output of the mixer. This distortion causes amplitude modulation (AM) to phase modulation (PM) conversion. No information is lost if the desired signal is frequency modulated. If the desired signal is phase modulated, the unwanted phase shift due to AM-to-PM conversion may result in detection error, which increases the BER. On the other hand, if the input power of the undesired signal exceeds the input $P_{-1\,\mathrm{dB}}$, distortion of the undesired signal does not affect the system performance. However, a strong undesired signal (known as a blocker or interferer) can overload a mixer and cause gain compression of the small desired signal if the mixer does not have sufficiently high input $P_{-1\,\mathrm{dB}}$ [5], [6].

The blocker should not reduce the gain of the small desired signal by more than 1 dB to avoid increasing the noise contribution from the IF stages significantly. Unfortunately, there is no simple relationship between the gain compression of the small desired signal and that of the large undesired signal. The relationship derived in [5] assumes a weakly nonlinear condition where the gain compression is solely caused by the third-order term in the transfer function of the circuit. If this were the case, the input blocker power that caused 1-dB gain compression to the small desired signal would be 3.1 dB less than the input $P_{-1\,\mathrm{dB}}$ of the circuit. This describes many practical mixers, but higher odd-order terms can also be important in the presence of large signals. The small desired signal can be viewed as amplitude modulation on top of the large blocker which functions like a carrier. Typically, the modulation signal (the small desired signal) is compressed more than the carrier (the large blocker). The actual value of input $P_{-1\,\mathrm{dB}}$ is not the true design criterion in many receiver systems. Alternately, a new performance parameter, the blocking $P_{-1\,\mathrm{dB}}$, can be defined as the input power of the blocker that causes 1-dB gain compression to the small desired signal.

D. Third-Order Intermodulation Distortion

Due to the odd-order nonlinearities in the transfer function of the mixer, two undesired signals in the adjacent channels generate third-order intermodulation (IM3) products at the output of the mixer. As illustrated in Fig. 5, one of the IM3 product can corrupt the desired signal if it falls within the desired channel. If the two adjacent channel frequencies are ω_a and ω_b, respectively, two IM_3 products are generated at frequencies $(2\omega_a - \omega_b)$ and $(2\omega_b - \omega_a)$, respectively. At low input power level, the IM_3 product is dominated by the third-order nonlinearity. As the input power increases, higher-order nonlinearities become more important.

The third-order intercept point (IP_3) measures only the third-order nonlinearity. Fig. 6 shows the magnitude responses of the desired signal and the IM3 product. The solid lines are the actual responses. At low input power levels, the gain of the desired signal is constant, and the power of the IM_3 product increases with the cube of the input power. At high input power levels, the gain of the desired signal is typically compressed (gain expansion may happen in class-AB and class-C circuits), and the IM_3 is no longer dominated by the third-order nonlinearity. Depending on the phase relationship among odd-order nonlinearities, the IM_3 may increase at a rate higher or lower than the cube of the input signal power. The dotted lines are the linear extrapolations of the small-signal magnitude responses of the desired signal and the IM_3 product, respectively. The point where the two extrapolated curves meet is the IP_3. Given the input signal power in the adjacent

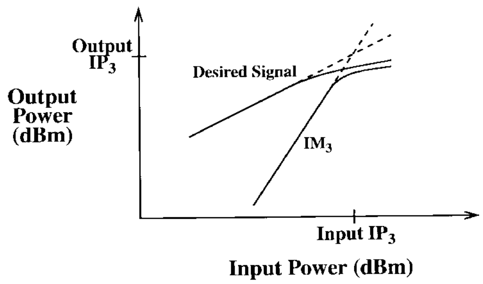


Fig. 6. Magnitude responses of desired signal and IM_3.

channels, the power of the IM_3 product can be calculated by using the IP_3 value. However, the calculated value only applies to the small-signal IM3 product since the IP_3 value is the result of extrapolation from the small-signal condition.

The system input IP_3 (power quantity in linear scale) for the downconverter shown in Fig. 1 is

$$\mathrm{IP}_3 = L_{\mathrm{RF}}\left[\frac{1}{\mathrm{IP}_{3(\mathrm{LNA})}} + \frac{1}{\mathrm{IP}_{3(\mathrm{MIX})}/(G_{\mathrm{LNA}}L_{\mathrm{IM}})}\right]^{-1} \tag{4}$$

where $\mathrm{IP}_{3(\mathrm{LNA})}$ and $\mathrm{IP}_{3(\mathrm{MIX})}$ are the input IP_3 (in power unit) of the LNA and downconversion mixer, respectively. This equation assumes that the IM_3 contributions from the filters are negligible since they are passive components, and that the IM_3 products from the LNA and the mixer add coherently (in phase). On the other hand, if the IM_3 products add incoherently (out-of-phase), the system input IP_3 becomes

$$\mathrm{IP}_3 = L_{\mathrm{RF}}\left[\frac{1}{\mathrm{IP}^2_{3(\mathrm{LNA})}} + \frac{1}{\mathrm{IP}^2_{3(\mathrm{MIX})}/(G^2_{\mathrm{LNA}}L^2_{\mathrm{MIX}})}\right]^{1/2}. \tag{5}$$

As shown in (4) and (5), increasing the gain of the LNA decreases the system input IP_3.

The numerical value of the input IP_3 is not directly related to that of input $P_{-1\,\mathrm{dB}}$ because IP_3 measures the small-signal nonlinearity which is dominated by the third-order nonlinearity, whereas $P_{-1\,\mathrm{dB}}$ measures the large-signal nonlinearity which includes all odd-order nonlinearities. Furthermore, IP_3 depends on the magnitude of the third-order nonlinearity only, but $P_{-1\,\mathrm{dB}}$ depends on both magnitude and phase of the third-order nonlinearity [8], [9]. If both IP_3 and $P_{-1\,\mathrm{dB}}$ were dominated by the third-order nonlinearity, the value of input IP_3 would be 9.6 dB (true for low-frequency case where the third-order nonlinearity is exactly out of phase with the fundamental signal) higher than that of input $P_{-1\,\mathrm{dB}}$. In many practical designs (one exception is the class AB mixer described in [6]), the numerical values of IP_3 are more than 9.6 dB higher than those of $P_{-1\,\mathrm{dB}}$.

E. Power Consumption

The power consumption of other building blocks within a receiver system is as important as that of the downconversion mixer. While optimizing the power consumption of the mixer, care has to be taken to avoid increasing the power consumption of other building blocks. For instance, a downconversion mixer with high NF increases the gain requirement of the LNA. This increases the power consumption of the LNA. As shown in (4) and (5), increasing the gain of the LNA also increases the input IP_3 requirement of the mixer in order to meet the system input IP_3 specification. This in turn increases the power consumption of the mixer [9]. A mixer which requires high LO drive increases the power consumption of the LO. It may take up to 10 mA of bias current in an LO output buffer to supply 0 dBm of LO power into the 50 Ω LO port of the mixer.

F. Port Return Loss

When a port impedance is not matched to that of the source resistance, some of the power delivered to the port is reflected back to the source. Return loss is defined as the fraction of incident power reflected. The impedance of the RF and LO input ports is typically matched to 50 Ω, while the impedance of the IF output port is matched to that of the IF filter. Impedance matching at the RF and IF ports is necessary to avoid signal reflection and excessive passband ripple in the frequency responses of the filters. Typically, return losses of less than -10 dB (voltage wave standing ratio of less than 2) are required. On the other hand, the return loss specification on the LO port can be more relaxed. However, excessive return loss requires the LO to deliver high power which would increase the power consumption of the overall system. Furthermore, excessive LO signal reflected back to the LO may cause LO-pulling problem.

G. Port Isolation

The isolation between LO and RF ports of the mixer is important as LO-to-RF feedthrough results in LO signal leaking through the antenna. The leaked LO signal should be small enough to avoid corrupting the desired signals of other RF systems. If the downconversion mixer is in a different package from the LNA, the amount of LO-to-RF feedthrough that is allowed depends on the reverse isolation of the LNA, and the stopband attenuation of the RF and image-rejection filters at the LO frequency. On the other hand, if the LNA is in the same package as the mixer, the LO signal can feed through to RF input port (due to capacitive and inductive coupling among package pins) of the LNA, bypassing the RF filter and LNA.

LO-to-IF and RF-to-IF isolations are not important because the high-frequency feedthrough signals can be rejected by the high-Q IF filter easily. However, large LO and RF feedthrough signals at the IF output port may saturate the IF output port, and decrease the $P_{-1\,\mathrm{dB}}$ of the mixer.

III. Mixer Topologies

Mixer topologies can divided into three categories, namely, double-balanced, single-balanced, and unbalanced designs. In this section, the advantages and disadvantages, as well as the design and optimization techniques for these three different kinds of mixers, are presented. The discussion starts with the double-balanced mixer. Although the circuit topology looks more complicated than the other two kinds of mixers, it is the easiest to design and to optimize.

A. Double-Balanced Mixer

Fig. 7 shows the basic circuit topology of a double-balanced active mixer realized in bipolar technology. The mixer comprises a differential-pair driver stage (Q_1 and Q_2) and a differential switching quad (Q_3, Q_4, Q_5, and Q_6). The driver stage amplifies the RF signal to compensate for the attenuation due to the switching process, and to reduce the noise contribution from the switching quad. If the RF input signal is single-ended, one side of the driver stage can be ac grounded. To reduce the voltage headroom required in a low-supply design, the tail current source $(2I_T)$ can be implemented by a resistor [10] or an inductor, and inductors can be used to bias the IF output ports to the supply voltage [6]. The switching quad perform the mixing function which converts the RF signal down to the IF as illustrated in the following equation:

$$\begin{aligned} I_O &= V_{\rm RF}\cos\omega_{\rm RF}t \times G_M \\ &\quad\times\left(\frac{2}{\pi}\cos\omega_{\rm LO}t - \frac{2}{3\pi}\cos 3\omega_{\rm LO}t + \cdots\right) \\ &= \frac{1}{\pi}G_M V_{\rm RF}\cos(\omega_{\rm LO}-\omega_{\rm RF})t \\ &\quad + \frac{1}{\pi}G_M V_{\rm RF}\cos(\omega_{\rm LO}+\omega_{\rm RF})t + \cdots \end{aligned} \tag{6}$$

where I_O is the differential output signal current, $\omega_{\rm RF}$ and $\omega_{\rm LO}$ are the RF and LO frequencies, respectively, $V_{\rm RF}$ is the RF input signal, and G_M is the transconductance of the driver stage. This equation assumes instantaneous switching (multiplying the RF signal with square wave) of the switching quad. If high-side mixing (LO frequency is higher than RF frequency) is used, the $(\omega_{\rm LO}-\omega_{\rm RF})$ term is the IF signal, and the $(\omega_{\rm LO}+\omega_{\rm RF})$ term is the unwanted signal. The $1/\pi$ factor is due to the power lost in the $(\omega_{\rm LO}+\omega_{\rm RF})$ term and other higher-frequency terms. If low-side mixing (LO frequency is lower than RF frequency) is used, the $(\omega_{\rm RF}-\omega_{\rm LO})$ and $(\omega_{\rm LO}+\omega_{\rm RF})$ terms are the wanted and unwanted signals, respectively. Equation (6) applies to the case where differential output is taken. If single-ended output is taken, the mixing process is represented by the following equation:

$$\begin{aligned} I_O &= \left(I_T + V_{\rm RF}\frac{G_M}{2}\cos\omega_{\rm RF}t\right) \\ &\quad\times\left(\frac{1}{2}+\frac{2}{\pi}\cos\omega_{\rm LO}t - \frac{2}{3\pi}\cos 3\omega_{\rm LO}t + \cdots\right) \\ &\quad+\left(I_T - V_{\rm RF}\frac{G_M}{2}\cos\omega_{\rm RF}t\right) \\ &\quad\times\left(\frac{1}{2}-\frac{2}{\pi}\cos\omega_{\rm LO}t + \frac{2}{3\pi}\cos 3\omega_{\rm LO}t + \cdots\right) \\ &= I_T + \frac{1}{\pi}G_M V_{\rm RF}\cos(\omega_{\rm LO}-\omega_{\rm RF})t \\ &\quad + \frac{1}{\pi}G_M V_{\rm RF}\cos(\omega_{\rm LO}+\omega_{\rm RF})t + \cdots \end{aligned} \tag{7}$$

where I_O is the single-ended output signal current. The LO and RF feedthrough signals are cancelled at the mixer output, but the dc component I_T remains.

To improve the linearity of the differential-pair driver stage, it can be degenerated by the impedance Z_e, which can be implemented by using either resistor, inductor, or capacitor.

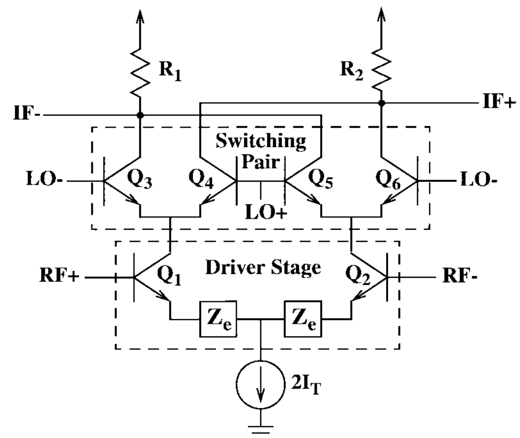


Fig. 7. Double-balanced active mixer.

It can be easily shown that a driver stage with reactive (inductive or capacitive) degeneration has lower NF than that with resistive degeneration since the degeneration reactance (apart from its loss resistance) does not introduce an additional noise source. To reduce the noise contribution from Q_1 and Q_2, large devices with small base resistance (r_b) should be used. The bias current $(2I_T)$ should be optimized to reduce the sum of base and collector shot noise contributions from Q_1 and Q_2. The input-referred noise due to the collector shot noise decreases with bias current, while that due to the base shot noise increases with bias current [11]. Ideally, the gain of the driver stage should be maximized (by minimizing the degeneration impedance) to minimize the noise contribution from the switching quad. However, linearity (IP_3 and $P_{-1\,\rm dB}$) sets the lower limit on the degeneration impedance. It can be shown that the differential-pair driver stage using inductive degeneration is more linear than that using resistive or capacitive degeneration (with the same bias current and transconductance) [9]. In other words, inductive degeneration is more current efficient than both resistive and capacitive degeneration.

The differential impedance looking into the bases of the driver stage is given by

$$Z_{\rm in} = 2\left(r_b + Z_e + \frac{1}{sC_\pi} + \frac{\omega_T}{s}Z_e\right) \tag{8}$$

where C_π is base–emitter capacitance of Q_1 and Q_2, ω_T is the unity current-gain frequency of Q_1 and Q_2. This equation neglects the effect of collector–base junction capacitance (C_μ) of Q_1 and Q_2. With resistive and inductive degeneration, the real part of the $Z_{\rm in}$ is supplied by the (r_b+Z_e) term and the $(r_b+(\omega_T/s)Z_e)$ term, respectively. Matching networks may be needed to match the real part of $Z_{\rm in}$ to the source resistance (R_s). The imaginary part of $Z_{\rm in}$ has to be cancelled. For stability reasons, capacitive degeneration is not recommended since the $((\omega_T/s)Z_e)$ term is a negative real number, and the real part of $Z_{\rm in}$ may be negative (negative resistance may cause oscillation).

If the gain of the driver stage and its output noise power were constant across all frequencies, the instantaneous-switching process would increase the input-referred noise

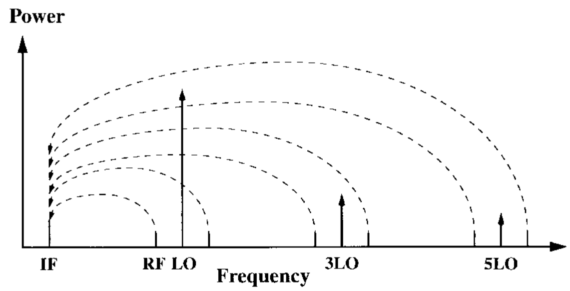


Fig. 8. Switching process increases input-referred noise power of driver stage.

contribution from the driver stage by a factor of $(\pi/2)^2$ (or 3.9 dB) as illustrated in Fig. 8 [12]. The LO and its harmonics (square wave has no even harmonics) mix noise at various frequencies down to the IF. In this case, the overall input-referred noise power (in linear scale) of the mixer would be

$$\text{input-referred noise of driver stage} \times \left(\frac{\pi}{2}\right)^2 + \text{noise contribution from switching pairs.} \quad (9)$$

With inductive degeneration, the gain of the driver stage decreases with frequency [13]. If high-side mixing (LO frequency is higher than RF) is used, the RF signal (and associated noise) has higher gain than the noise at the image frequency. Also, noise at higher frequencies is attenuated by the degeneration inductance. In this case, the mixing process increases the input-referred noise power of the driver stage by a factor of less than $(\pi/2)^2$. On the other hand, if low-side mixing is used, the mixing process increases the input-referred noise power of the driver stage by a factor of more than $(\pi/2)^2$ since noise at the image frequency has higher gain than the RF signal. Therefore, high-side mixing is recommended if the driver stage uses inductive degeneration. Similarly, low-side mixing is recommended if the driver stage uses capacitive degeneration.

As shown in (6) and (7), there are little RF and LO feedthrough signals at the IF output ports of the double-balanced mixer. In other words, double-balanced mixers reject LO and RF feedthrough. Since the switching quad is a differential structure, a double-balanced mixer also reject LO-to-RF feedthrough if the switching quad is driven differentially. If the LO input signal is single-ended, an LO buffer can be used to convert the single-ended LO signal into a differential one.

The switching quad should be driven by a large LO signal to minimize its noise contribution. The switching quad contributes noise to the mixer output when all transistors (Q_3, Q_4, Q_5, Q_6) are active [12], and a large LO amplitude is needed to reduce the duration of this condition. Linearity, head room, and power consumption considerations set the upper limit on the LO amplitude. A very large LO amplitude results in excessive current being pumped into the common-emitter points of the switching quad through the base–emitter junction capacitance (C_{je}), and thus generates additional third-order intermodulation [12]. Large LO amplitudes also decrease the voltage headroom at the mixer output. Another disadvantage of using large LO amplitudes is increased power consumption.

In bipolar transistor technology, differential LO signals larger than 300 mV are typically used to achieve a low noise figure [6], [14]. If the switching quad is driven directly by an external LO, 300 mV of sinusoidal signal is equivalent to 0 dBm of LO power (assuming LO port is matched to 50 Ω). It might take up to 10 mA of bias current in an external LO driver to supply this LO power. Therefore, an LO buffer is recommended to reduce the LO input power requirement [6], [14]. The LO buffer needs to have low noise to avoid increasing the phase noise of the LO signal.

Reasonably large devices should be used to reduce the r_b noise contribution from the switching quad. Small r_b also reduces the ac voltage drop across r_b which would decrease the effective LO signal amplitude driving the switching devices. However, if the C_{je} is too large, the switching of C_{je} can pump additional current into the common-emitter point of the switching quad, and decrease the linearity [12].

In the double-balanced design, the IF output can be taken either single-endedly or differentially with little impact on the linearity performance. The IF output can be taken single-endedly by using either the IF+ or IF− port. However, taking the output differentially increases the output power and conversion gain of the mixer. Furthermore, noise from the tail current source $(2I_T)$ at the IF can feed through to the IF output port. This common-mode noise would increase the IF output noise power significantly if single-ended output is taken. Since this noise is common-mode, it can be cancelled by taking the IF output differentially. If a differential IF filter is available, both of the IF output ports can be connected directly to the filter. On the other hand, if the IF filter is single-ended, differential to single-ended conversion is needed. This can be achieved by using either a transformer or some kind of narrow-band current-combining networks [15], [16]. Alternately, an IF output buffer can be used (linearity and headroom issues should be considered carefully).

The basic topology of the double-balanced mixer shown in Fig. 7 can also be implemented in FET technologies (GaAs MESFET [16] or silicon MOSFET [17]). Since the inherent linearity of an FET is good enough for many applications, the driver stage does not have to be degenerated. Furthermore, the linearity of the driver stage can be improved by increasing $(V_{\mathrm{GS}} - V_t)$, where V_{GS} is the gate–source voltage and V_t is the threshold voltage. Compared to the bipolar design, the FET switching quad needs to be driven by a larger LO signal to minimize its noise contribution. The reason is that larger LO voltage swing is needed to turn off one side of the FET switching quad.

B. Single-Balanced Mixer

Fig. 9 shows the basic circuit topology of a single-balanced active mixer. Due to the simplicity of the circuit, single-balanced mixers have lower noise figure than double-balanced mixers. This is because there are fewer noise contributors in the single-balanced design. The mixer comprises a common-emitter driver stage (Q_7) and a differential switching pair $(Q_8$

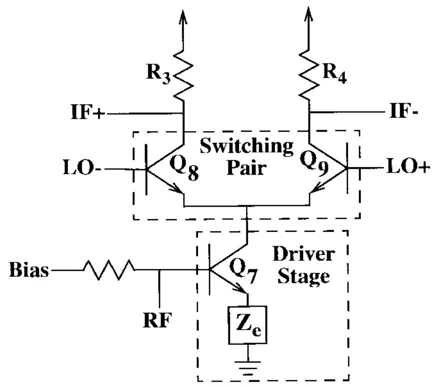


Fig. 9. Single-balanced active mixer.

and Q_9). The driver stage is degenerated by an impedance Z_e to improve its linearity. It can be shown that reactive degeneration has better noise performance than resistive degeneration, and that inductive degeneration is more current efficient than both resistive and capacitive degeneration [9]. Inductively degenerated common-emitter driver stages also exhibit class AB behavior, which can be exploited to reduce the bias current required to meet the $P_{-1\,\mathrm{dB}}$ specification [6], [14]. Compared to the differential-pair driver stage, the common-emitter driver stage requires less bias current for the same transconductance and linearity [9]. The noise performance of the common-emitter driver stage can be optimized in a similar way as the differential-pair driver stage of the double-balanced mixer. The impedance looking into the base of the driver stage is given by

$$Z_{\mathrm{in}} = r_b + Z_e + \frac{1}{sC_\pi} + \frac{\omega_T}{s} Z_e. \tag{10}$$

With capacitive degeneration, the $((\omega_T/s)Z_e)$ is a negative real number. Negative resistance may cause oscillation.

Alternately, the driver stage can be implemented by common-base transconductance stage [12]. Common-base transconductance stage tends to be more linear than common-emitter transconductance stage. However, the lack of current gain (or low power gain) makes the design very noisy because the noise contribution from the switching pair is not attenuated.

The switching pair performs the mixing operation. The noise and linearity performance of the switching pair can be optimized in a similar way as the switching quad of the double-balanced mixer. In contrast to the double-balanced mixer which has no dc components in the LO and RF signals, the single-balanced design has a dc component in the RF signal. If the IF output is taken differentially, the mixing process can be represented by the following equation:

$$\begin{aligned} I_O &= (I_Q + V_{\mathrm{RF}} G_M \cos \omega_{\mathrm{RF}} t) \\ &\quad \times \left(\frac{2}{\pi} \cos \omega_{\mathrm{LO}} t - \frac{2}{3\pi} \cos 3\omega_{\mathrm{LO}} t + \cdots \right) \\ &= \frac{2}{\pi} I_Q \cos \omega_{\mathrm{LO}} t + \frac{1}{\pi} G_M V_{\mathrm{RF}} \cos(\omega_{\mathrm{LO}} - \omega_{\mathrm{RF}}) t \\ &\quad + \frac{1}{\pi} G_M V_{\mathrm{RF}} \cos(\omega_{\mathrm{LO}} + \omega_{\mathrm{RF}}) t + \cdots \end{aligned} \tag{11}$$

where I_Q is the bias current of the driver stage. The $((2/\pi)I_Q \cos \omega_{\mathrm{LO}} t)$ term represents the LO feedthrough. In other words, single-balanced mixer rejects RF-to-IF feedthrough (if the IF output is taken differentially), but not LO-to-IF feedthrough. If the differential pair is driven differentially, single-balanced mixer also rejects LO-to-RF feedthrough. On the other hand, if the IF output is taken single-endedly, there are dc components in both RF and LO signals. In this case, the mixing process is represented by the following equation:

$$\begin{aligned} I_O &= (I_Q + V_{\mathrm{RF}} G_M \cos \omega_{\mathrm{RF}} t) \\ &\quad \times \left(\frac{1}{2} + \frac{2}{\pi} \cos \omega_{\mathrm{LO}} t - \frac{2}{3\pi} \cos 3\omega_{\mathrm{LO}} t + \cdots \right) \\ &= \frac{I_Q}{2} + \frac{1}{2} V_{\mathrm{RF}} G_M \cos \omega_{\mathrm{RF}} t + \frac{2}{\pi} I_Q \cos \omega_{\mathrm{LO}} t \\ &\quad + \frac{1}{\pi} G_M V_{\mathrm{RF}} \cos(\omega_{\mathrm{LO}} - \omega_{\mathrm{RF}}) t \\ &\quad + \frac{1}{\pi} G_M V_{\mathrm{RF}} \cos(\omega_{\mathrm{LO}} + \omega_{\mathrm{RF}}) t + \cdots. \end{aligned} \tag{12}$$

The $((1/2)V_{\mathrm{RF}} G_M \cos \omega_{\mathrm{RF}} t)$ and $((2/\pi) I_Q \cos \omega_{\mathrm{LO}} t)$ terms represent the RF and LO feedthrough signals, respectively. Normally, the RF and LO feedthrough signals at the IF output ports do not cause problems since the IF filter has high enough stopband attenuation to filter out the unwanted signals at high frequencies. However, these feedthrough signals can produce large signal swings at the IF output ports, and degrade the $P_{-1\,\mathrm{dB}}$ by saturating the output ports. Hence, capacitors are needed at the IF output ports to attenuate these high-frequency feedthrough signals [6], [14].

Taking the IF output single-endedly or differentially does not affect the linearity performance. However, taking the IF output single-endedly would increase the input-referred noise contribution from the driver stage by a factor of $[2(\pi/2)^2]$ (or 6.9 dB) if the output noise power of the driver stage were constant across all frequencies. Since there is a dc component in the LO signal, noise from the driver stage at the IF can mix with this dc component and increase the noise power at the IF output ports. If the driver stage is inductively degenerated, it has high gain and noise power at the IF. In this case, taking the IF output single-endedly would increase the NF significantly. Therefore, the IF output of the single-balanced mixer has to be taken differentially (LO signal has no dc component) in order to minimize the NF [6], [14].

Since the RF signal has a dc component, noise from the LO at the IF can mix with this dc component and increase the noise power at the IF output port. Therefore, the LO signal should have low noise power at the IF. If an LO buffer is used, bandpass or highpass load can be used at the output of the LO buffer to reduce its noise at the IF [6], [14].

C. Unbalanced Mixer

Figs. 10 and 11 show two different circuit topologies of unbalanced mixers. The circuits are the simplest among the active mixers, and hence the unbalanced designs have the lowest noise figures. In both topologies, the mixing operation

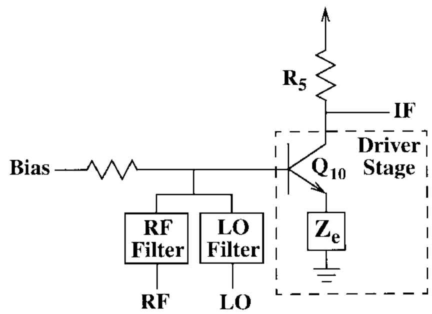


Fig. 10. Single-transistor active mixer.

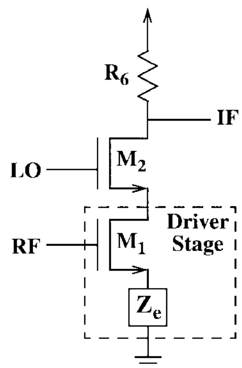


Fig. 11. Dual-gate FET mixer.

is performed by modulating transconductances of the driver stages with the LO signals. The single-transistor active mixer shown in Fig. 10 can also be implemented in FET technologies. The LO signal modulates the transconductance of the driver stage by varying the base–emitter voltage $(V_{\rm BE})$ of Q_{10}. In the dual-gate FET mixer shown in Fig. 11, the LO signal modulates the transconductance of the driver stage by varying the drain–source voltage $(V_{\rm DS})$ of M_1. The drain of M_1 is typically biased at the edge between triode and saturation regions to maximize the transconductance variation due to the LO signal. This design cannot be implemented in bipolar technology since the frequency response of bipolar transistors is greatly degraded in saturation.

Since there are dc components in both RF and LO signals, unbalanced mixers do not reject RF-to-IF and LO-to-IF feedthrough. Hence, a capacitor is needed at the IF output port to suppress these high-frequency feedthrough signals. Since the LO signals are unbalanced, unbalanced mixers do not reject LO-to-RF feedthrough. In the single-transistor active mixer shown in Fig. 10, the LO signal is injected into the RF port through the RF filter. Hence, the magnitude of LO-to-RF feedthrough depends on the stopband attenuation of the RF filter at the LO frequency. In the dual-gate FET mixer shown in Fig. 11, the magnitude of LO-to-RF feedthrough depends on the gate-to-drain $(C_{\rm GD})$ capacitance of M_1.

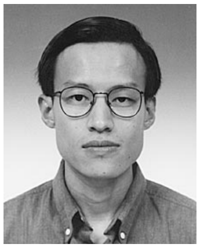
In unbalanced designs, noise from the driver stage at the IF can mix with the dc component of the LO signal, and thus increase the noise power at the IF output port. To reduce this IF noise, the driver stage has to be degenerated by a capacitor. In this case, the driver stage has little gain and noise at the IF. Similarly, noise from the LO at the IF can mix with the dc component of the RF signals, and increase the noise power at the IF output port. Hence, the LO signal needs to have low noise at the IF.

IV. Conclusions

The definition of performance parameters that are relevant to the design of RF downconversion mixers has been presented. Their relationship with the system performance was discussed. Three kinds of mixer topologies have been presented. The unbalanced mixers have the best noise performance, but their unbalanced nature prevents them from general use in monolithic forms. The double-balanced mixers are the easiest to design due to their fully balanced structure, but they have the highest noise figure. The single-balanced mixers are a compromise between the unbalanced and double-balanced designs.

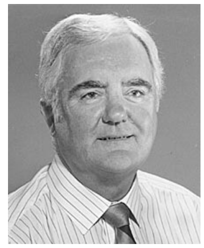
References

[1] S. A. Mass, *Microwave Mixers*. Norwood, MA: Artech House, 1986.
[2] G. D. Vendelin, A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990.
[3] S. Weiner, D. Neuf, and S. Spohrer, “2 to 8 GHz double balanced MESFET mixer with +30 dBm input 3rd order intercept,” in *IEEE MTT-S Int. Microwave Symp. Dig.*, 1988, vol. 11, pp. 1097–1100.
[4] S. A. Maas, “A low distortion GaAs MESFET resistive mixer,” *Microwave J.*, vol. 31, no. 3, pp. 213–220, Mar. 1988.
[5] R. G. Meyer and A. K. Wong, “Blocking and desensitization in RF amplifiers,” *IEEE J. Solid-State Circuits*, vol. 30, pp. 944–946, Aug. 1995.
[6] K. L. Fong, C. D. Hull, and R. G. Meyer, “A class AB monolithic mixer for 900 MHz applications,” *IEEE J. Solid-State Circuits*, pp. 1166–1172, Aug. 1997.
[7] C. D. Hull, J. L. Tham, and R. R. Chu, “A direct-conversion receiver for 900 MHz (ISM band) spread-spectrum digital cordless telephone,” *IEEE J. Solid-State Circuits*, vol. 31, pp. 1955–1963, Dec. 1996.
[8] D. D. Weiner and J. F. Spina, *Sinusoidal Analysis and Modeling of Weakly Nonlinear Circuits*. New York: Van Nostrand Reinhold, 1980.
[9] K. L. Fong and R. G. Meyer, “High-frequency nonlinearity analysis of common-emitter and differential-pair transconductance stages,” *IEEE J. Solid-State Circuits*, vol. 33, pp. 548–555, Apr. 1998.
[10] R. G. Meyer and W. D. Mack, and J. J. E. M. Hageraats, “A 2.5 GHz BiCMOS transceiver for wireless LAN,” in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 1997, pp. 310–311.
[11] K. L. Fong, “Design and optimization techniques for monolithic RF downconversion mixers,” Ph.D. dissertation, Univ. Calif. Berkeley, Berkeley, 1997. Available WWW: http://kabuki.eecs.berkeley.edu/∼fong/.
[12] C. D. Hull, “Analysis and optimization of monolithic RF downconversion receivers,” Ph.D. dissertation, Univ. Calif. Berkeley, Berkeley, 1992.
[13] R. G. Meyer and W. D. Mack, “A 1-GHz BiCMOS RF front-end,” *IEEE J. Solid-State Circuits*, vol. 29, pp. 350–355, Mar. 1994.
[14] K. L. Fong and R. G. Meyer, “A 2.4 GHz monolithic mixer for wireless LAN applications,” in *Proc. IEEE Custom Integrated Circuits Conf.*, May 1997, pp. 9.4.1–9.4.4.
[15] S. H. Lee, “Resonance Balun with arbitrary impedance,” U.S. Patent no. 5 412 355.
[16] L. E. Larson, *RF and Microwave Circuit Design for Wireless Communications*. Norwood, MA: Artech House, 1996.
[17] A. Rofougaran, J. Y-C. Chan, M. Rofougaran, S. Khorram, and A. A. Abidi, “A 1 GHz CMOS RF front-end IC with wide dynamic range,” in *Proc. 21st European Solid-State Circuits Conf.*, Sept. 1995, pp. 250–253.
[18] D. O. Pederson and K. Mayaram, *Analog Integrated Circuits for Communication*. Norwell, MA: Kluwer Academic, 1991.



Keng Leong Fong (S'93–M'93) was born in Kuala Lumpur, Malaysia, on March 6, 1970. He received the B.A.Sc. degree in engineering science (computer engineering option), and the M.A.Sc. degree in electrical engineering, both from the University of Toronto, Canada, in 1992 and 1993, respectively. He received the Ph.D. degree in electrical engineering from the University of California at Berkeley in 1997.

During the summer of 1995, he worked at Rockwell International Corporation, Newport Beach, CA, where he was involved in designing class-AB mixers and evaluating CAD tools for nonlinear noise analysis. During the summer of 1996, he worked in the same company, designing class-AB power amplifiers. Since 1997, he has been working with the Advanced Development Group of Philips Semiconductors, Sunnyvale, CA, where he explores new circuit techniques and supports the next-generation BiCMOS process development. His current research interest is in the areas of analog integrated circuit design for RF applications, BiCMOS process development, and wireless system architecture.



Robert G. Meyer (S'64–M'68–SM'74–F'81) was born in Melbourne, Australia, on July 21, 1942. He received the B.E., M.Eng.Sci., and Ph.D. degrees in electrical engineering from the University of Melbourne in 1963, 1965, and 1968, respectively.

In 1968, he was employed as an Assistant Lecturer in Electrical Engineering at the University of Melbourne. Since September 1968, he has been employed in the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, where he is now a Professor. His current research interests are high-frequency analog integrated-circuit design and device fabrication. He has acted as a consultant on electronic circuit design for numerous companies in the electronics industry. He is coauthor of the book, *Analysis and Design of Analog Integrated Circuits* (New York: Wiley, 1993), and Editor of the book, *Integrated Circuit Operational Amplifiers* (Piscataway, NJ: IEEE Press, 1978).

Dr. Meyer was President of the IEEE Solid-State Circuits Council and was an Associate Editor of the IEEE JOURNAL OF SOLID-STATE CIRCUITS and of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS.