RF Microelectronics

Second Edition

Behzad Razavi

Prentice Hall Communications Engineering and Emerging Technologies Series

Theodore S. Rappaport, Series Editor

RF MICROELECTRONICS

Second Edition

are neglected. Assuming a unity voltage gain for the mixer for simplicity, we write the total output noise as $A_{v1}^2(\overline{V_{n,LNA}^2} + 4kTR_S) + \overline{V_{n,mix}^2}$. The overall noise figure is thus equal to

$$\mathrm{NF_{tot}} = \frac{A_{v1}^2(\overline{V_{n,LNA}^2} + 4kTR_S) + \overline{V_{n,mix}^2}}{A_{v1}^2}\frac{1}{4kTR_S} \tag{5.7}$$

$$= \mathrm{NF_{LNA}} + \frac{\overline{V_{n,mix}^2}}{A_{v1}^2}\cdot\frac{1}{4kTR_S}. \tag{5.8}$$

In other words, for NF calculations, the noise of the second stage is divided by the gain from the input voltage source to the LNA output.

Now consider the same cascade repeated in Fig. 5.3(b) with the nonlinearity of the LNA expressed as a third-order polynomial. From Chapter 2, we have[1]

$$\frac{1}{\mathrm{IP}_{3,\mathrm{tot}}^2} = \frac{1}{\mathrm{IP}_{3,\mathrm{LNA}}^2} + \frac{\alpha_1^2}{\mathrm{IP}_{3,\mathrm{mixer}}^2}. \tag{5.9}$$

In this case, α_1 denotes the voltage gain from the *input of the LNA* to its output. With input matching, we have $R_{in} = R_S$ and $\alpha_1 = 2A_{v1}$. That is, the mixer noise is divided by the *lower* gain and the mixer $\mathrm{IP_3}$ by the *higher* gain—both against the designer's wish.

Input Return Loss The interface between the antenna and the LNA entails an interesting issue that divides analog designers and microwave engineers. Considering the LNA as a *voltage* amplifier, we may expect that its input impedance must ideally be infinite. From the noise point of view, we may precede the LNA with a transformation network to obtain minimum NF. From the signal *power* point of view, we may realize *conjugate* matching between the antenna and the LNA. Which one of these choices is preferable?

We make the following observations. (1) the (off-chip) band-select filter interposed between the antenna and the LNA is typically designed and characterized as a high-frequency device and with a standard termination of 50 Ω. If the load impedance seen by the filter (i.e., the LNA input impedance) deviates from 50 Ω significantly, then the passband and stopband characteristics of the filter may exhibit loss and ripple. (2) Even in the absence of such a filter, the antenna itself is designed for a certain real load impedance, suffering from uncharacterized loss if its load deviates from the desired real value or contains an imaginary component. Antenna/LNA co-design could improve the overall performance by allowing even non-conjugate matching, but it must be borne in mind that, if the antenna is shared with the transmitter, then its impedance must contain a negligible imaginary part so that it *radiates* the PA signal. (3) In practice, the antenna signal must travel a considerable distance on a printed-circuit board before reaching the receiver. Thus, poor matching at the RX input leads to significant reflections, an uncharacterized loss, and possibly voltage attenuation. For these reasons, the LNA is designed for a 50-Ω resistive input impedance. Since none of the above concerns apply to the other interfaces within the RX (e.g., between the LNA and the mixer or between the LO and the mixer), they are typically designed to maximize *voltage* swings rather than power transfer.

1. The $\mathrm{IM_3}$ components arising from second-order terms are neglected.

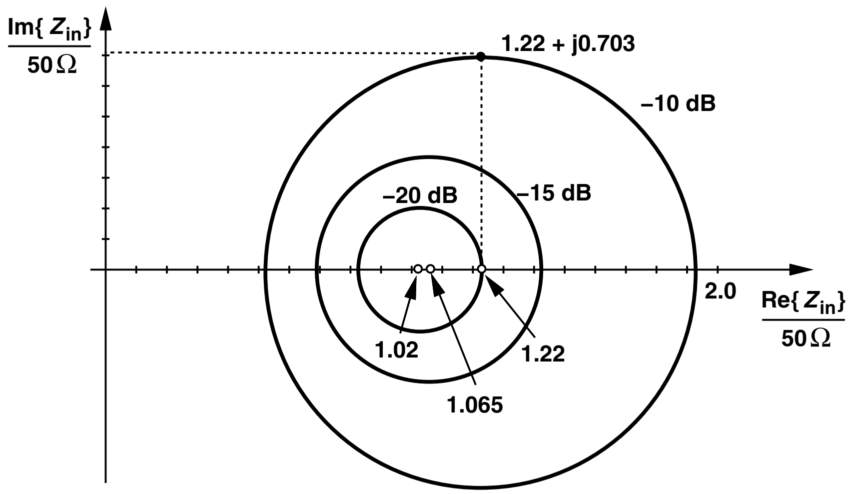

Figure 5.4 *Constant-Γ contours in the input impedance plane.*

The quality of the input match is expressed by the input "return loss," defined as the reflected power divided by the incident power. For a source impedance of R_S, the return loss is given by[2]

$$\Gamma = \left|\frac{Z_{in} - R_S}{Z_{in} + R_S}\right|^2, \tag{5.10}$$

where Z_{in} denotes the input impedance. An input return loss of -10 dB signifies that one-tenth of the power is reflected—a typically acceptable value. Figure 5.4 plots contours of constant Γ in the Z_{in} plane. Each contour is a circle with its center shown. For example, $Re\{Z_{in}\} = 1.22 \times 50\,\Omega = 61\,\Omega$ and $Im\{Z_{in}\} = 0.703 \times 50\,\Omega = 35.2\,\Omega$ yield $S_{11} = -10$ dB. In Problem 5.1, we derive the equations for these contours. We should remark that, in practice, a Γ of about -15 dB is targeted so as to allow margin for package parasitics, etc.

Stability Unlike the other circuits in a receiver, the LNA must interface with the "outside world," specifically, a poorly-controlled source impedance. For example, if the user of a cell phone wraps his/her hand around the antenna, the antenna impedance changes.[3] For this reason, the LNA must remain stable for all source impedances at *all frequencies*. One may think that the LNA must operate properly only in the frequency band of interest and not necessarily at other frequencies, but if the LNA begins to oscillate at any frequency, it becomes highly nonlinear and its gain is very heavily compressed.

A parameter often used to characterize the stability of circuits is the "Stern stability factor," defined as

$$K = \frac{1 + |\Delta|^2 - |S_{11}|^2 - |S_{22}|^2}{2|S_{21}||S_{12}|}, \tag{5.11}$$

2. Note that Γ is sometimes defined as $(Z_{in} - R_S)/(Z_{in} + R_S)$, in which case it is expressed in decibels by computing $20 \log \Gamma$ (rather than $10 \log \Gamma$).

3. In the presence of a front-end band-select filter, the LNA sees smaller changes in the source impedance.

where $\Delta = S_{11}S_{22} - S_{12}S_{21}$. If $K > 1$ and $\Delta < 1$, then the circuit is unconditionally stable, i.e., it does not oscillate with any combination of *source* and *load* impedances. In modern RF design, on the other hand, the load impedance of the LNA (the input impedance of the on-chip mixer) is relatively well-controlled, making K a pessimistic measure of stability. Also, since the LNA output is typically not matched to the input of the mixer, S_{22} is not a meaningful quantity in such an environment.

Example 5.2

A cascade stage exhibits a high reverse isolation, i.e., $S_{12} \approx 0$. If the output impedance is relatively high so that $S_{22} \approx 1$, determine the stability conditions.

Solution:

With $S_{12} \approx 0$ and $S_{22} \approx 1$,

$$K \approx \frac{1 - |S_{22}|^2}{2|S_{21}||S_{12}|} > 1 \tag{5.12}$$

and hence

$$|S_{21}| < \frac{1 - |S_{22}|^2}{2|S_{12}|}. \tag{5.13}$$

In other words, the forward gain must not exceed a certain value. For $\Delta < 1$, we have

$$S_{11} < 1, \tag{5.14}$$

concluding that the input *resistance* must remain positive.

The above example suggests that LNAs can be stabilized by maximizing their reverse isolation. As explained in Section 5.3, this point leads to two robust LNA topologies that are naturally stable and hence can be optimized for other aspects of their performance with no stability concerns. A high reverse isolation is also necessary for suppressing the LO leakage to the input of the LNA.

LNAs may become unstable due to ground and supply parasitic inductances resulting from the packaging (and, at frequencies of tens of gigahertz, the on-chip line inductances). For example, if the gate terminal of a common-gate transistor sees a large series inductance, the circuit may suffer from substantial feedback from the output to the input and become unstable at some frequency. For this reason, precautions in the design and layout as well as accurate package modeling are essential.

Linearity In most applications, the LNA does not limit the linearity of the receiver. Owing to the cumulative gain through the RX chain, the latter stages, e.g., the baseband amplifiers or filters tend to limit the overall input IP_3 or P_{1dB}. We therefore design and optimize LNAs with little concern for their linearity.

An exception to the above rule arises in "full-duplex" systems, i.e., applications that transmit and receive simultaneously (and hence incorporate FDD). Exemplified by the

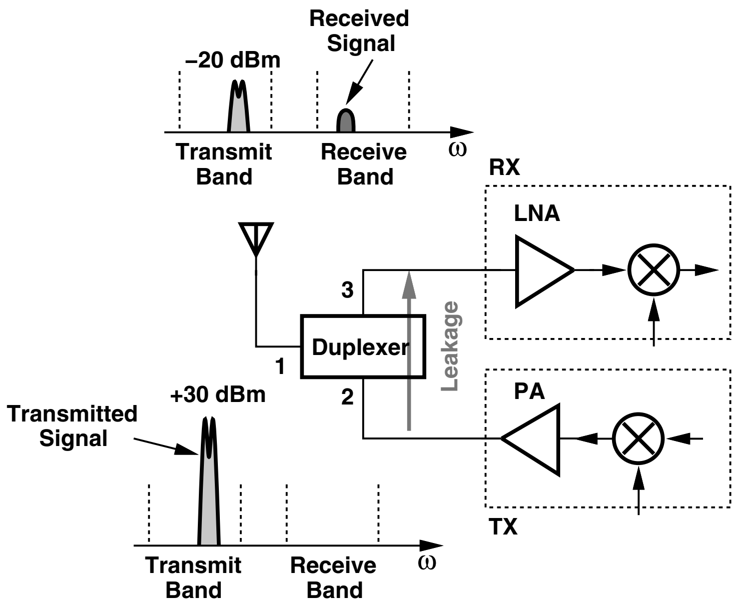

Figure 5.5 *TX leakage to RX in a full-duplex system.*

CDMA systems studied in Chapter 3, full-duplex operation must deal with the leakage of the strong transmitted signal to the receiver. To understand this issue, let us consider the front end shown in Fig. 5.5, where a duplexer separates the TX and RX bands. Modeling the duplexer as a three-port network, we note that S_{31} and S_{21} represent the losses in the RX and TX paths, respectively, and are about 1 to 2 dB. Unfortunately, leakages through the filter and the package yield a finite isolation between ports 2 and 3, as characterized by an S_{32} of about −50 dB. In other words, if the PA produces an average output power of +30 dBm (1 W), then the LNA experiences a signal level of −20 dBm in the TX band while sensing a much smaller received signal. Since the TX signal exhibits a variable envelope, its peak level may be about 2 dB higher. Thus, the receiver must remain uncompressed for an input level of −18 dBm. We must therefore choose a P_{1dB} of about −15 dBm to allow some margin.

Such a value for P_{1dB} may prove difficult to realize in a receiver. With an LNA gain of 15 to 20 dB, an input of −15 dBm yields an output of 0 to +5 dBm (632 to 1124 mV_{pp}), possibly compressing the LNA at *its output*. The LNA linearity is therefore critical. Similarly, the 1-dB compression point of the downconversion mixer(s) must reach 0 to +5 dBm. (The corresponding mixer IP_3 is roughly +10 to +15 dBm.) Thus, the mixer design also becomes challenging. For this reason, some CDMA receivers interpose an off-chip filter between the LNA and the mixer(s) so as to remove the TX leakage [1].

The linearity of the LNA also becomes critical in wideband receivers that may sense a large number of strong interferers. Examples include "ultra-wideband" (UBW), "software-defined," and "cognitive" radios.

Bandwidth The LNA must provide a relatively flat response for the frequency range of interest, preferably with less than 1 dB of gain variation. The LNA −3-dB bandwidth must therefore be substantially larger than the actual band so that the roll-off at the edges remains below 1 dB.

In order to quantify the difficulty in achieving the necessary bandwidth in a circuit, we often refer to its "fractional bandwidth," defined as the total −3-dB bandwidth divided by the center frequency of the band. For example, an 802.11g LNA requires a fractional bandwidth greater than 80 MHz/2.44 GHz = 0.0328.

Example 5.3

An 802.11a LNA must achieve a −3-dB bandwidth from 5 GHz to 6 GHz. If the LNA incorporates a second-order *LC* tank as its load, what is the maximum allowable tank *Q*?

Solution:

As illustrated in Fig. 5.6, the fractional bandwidth of an *LC* tank is equal to $\Delta\omega/\omega_0 = 1/Q$. Thus, the *Q* of the tank must remain less than 5.5 GHz/1 GHz = 5.5.

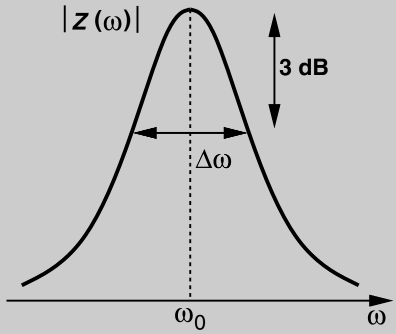

Figure 5.6 *Relationship between bandwidth and Q of a tank.*

LNA designs that must achieve a relatively large fractional bandwidth may employ a mechanism to *switch* the center frequency of operation. Depicted in Fig. 5.7(a) is an

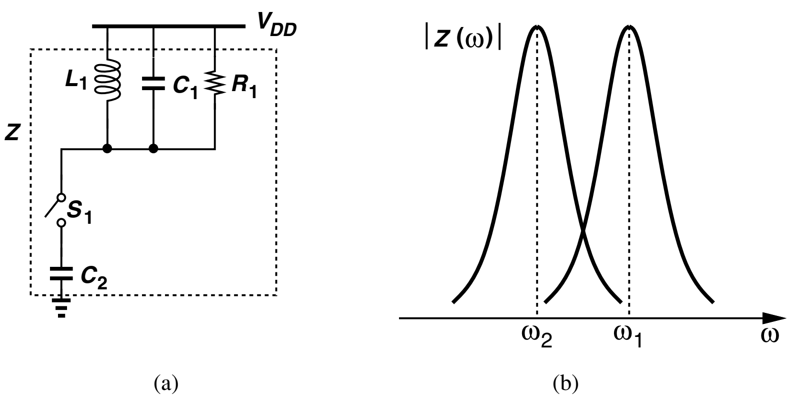

Figure 5.7 *(a) Band switching, (b) resulting frequency response.*

example, where an additional capacitor, C_2, can be switched into the tank, thereby changing the center frequency from $\omega_1 = 1/\sqrt{L_1C_1}$ to $\omega_2 = 1/\sqrt{L_1(C_1 + C_2)}$ [Fig. 5.7(b)]. We return to this concept in Section 5.5.

Power Dissipation The LNA typically exhibits a direct trade-off among noise, linearity, and power dissipation. Nonetheless, in most receiver designs, the LNA consumes only a small fraction of the overall power. In other words, the circuit's noise figure generally proves much more critical than its power dissipation.

5.2 PROBLEM OF INPUT MATCHING

As explained in Section 5.1, LNAs are typically designed to provide a 50-Ω input resistance and negligible input reactance. This requirement limits the choice of LNA topologies. In other words, we cannot begin with an arbitrary configuration, design it for a certain noise figure and gain, and then decide how to create input matching.

Let us first consider the simple common-source stage shown in Fig. 5.8, where C_F represents the gate-drain overlap capacitance. At very low frequencies, R_D is much smaller than the impedances of C_F and C_L and the input impedance is roughly equal to $[(C_{GS} + C_F)s]^{-1}$. At very high frequencies, C_F shorts the gate and drain terminals of M_1, yielding an input *resistance* equal to $R_D||(1/g_m)$. More generally, the reader can prove that the real and imaginary parts of the input admittance are, respectively, equal to

$$Re\{Y_{in}\} = R_D C_F \omega^2 \frac{C_F + g_m R_D (C_L + C_F)}{R_D^2 (C_L + C_F)^2 \omega^2 + 1} \tag{5.15}$$

$$Im\{Y_{in}\} = C_F \omega \frac{R_D^2 C_L (C_L + C_F) \omega^2 + 1 + g_m R_D}{R_D^2 (C_L + C_F)^2 \omega^2 + 1}. \tag{5.16}$$

Is it possible to select the circuit parameters so as to obtain $Re\{Y_{in}\} = 1/(50\,\Omega)$? For example, if $C_F = 10\,\text{fF}$, $C_L = 30\,\text{fF}$, $g_m R_D = 4$, and $R_D = 100\,\Omega$, then $Re\{Y_{in}\} = (7.8\text{ k}\Omega)^{-1}$ at 5 GHz, far from $(50\,\Omega)^{-1}$. This is because C_F introduces little feedback at this frequency.

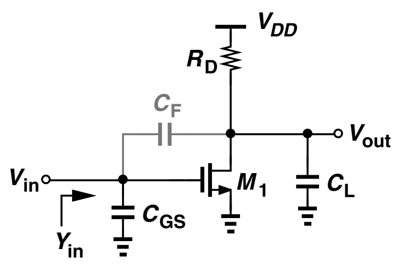

Figure 5.8 *Input admittance of a CS stage.*