Chapter 2 Fundamentals of Low Noise Amplifier

The first stage of the wireless receivers is typically a low noise amplifier (LNA), whose main function is to provide enough gain to overcome the noises of subsequent stages and amplify the weak input signal from antenna, but adding as little noise as possible. At the same time, good linearity and low power consumption are also important requirements. In this chapter, We fist introduce the types of noise sources and popular noise models of MOSFET, then provide the design concepts, considerations and steps of low noise amplifier further.

2.1 Noise in radio receiver

Noise is broadly defined as “everything except the desired signal”. The sensitivity of a radio receiver is limited by the presence of noise. Many types of noise sources shown as Fig. 2-1 exist everywhere, but only three major electric noise forms are discussed in following: thermal noise, flicker noise and shot noise since they play important roles in electric circuits and devices. Thermal noise comes from Brownian motion of charge carriers and has a dependence on the absolute temperature. The flicker noise usually termed as 1/f noise has a dependence on frequency, but no universal mechanism about it has been identified. Another noise source known as shot noise was first described and explained by Schottky in 1918. More detail discussions and explanations about the noise types and mechanisms can be found in [1] [2], so we ignore these parts here.

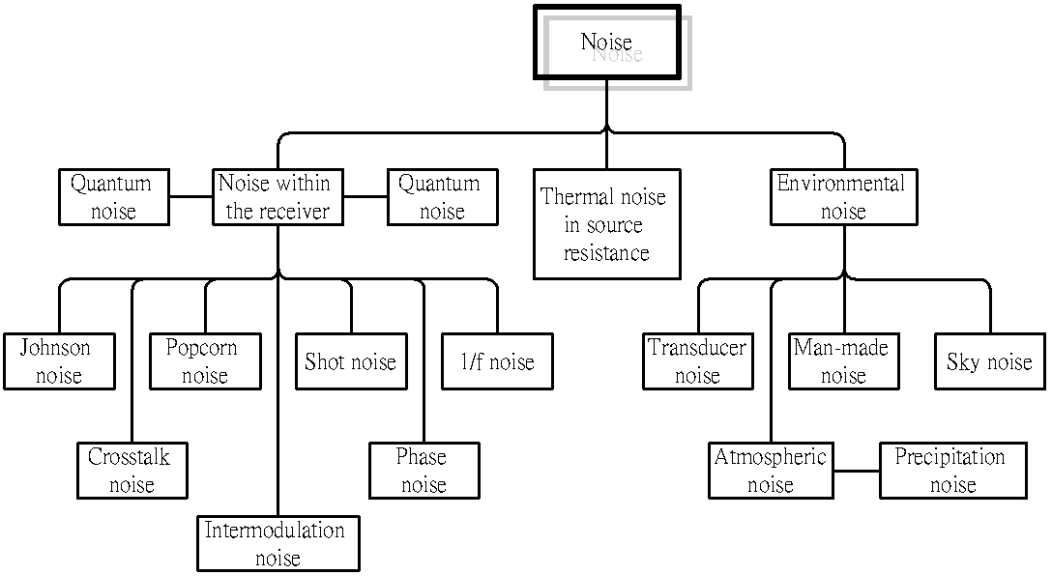

Fig. 2-1 Types of noise sources.

2.2 Noise model of MOSFET

The modified noise model of MOSFET suggested by Thomas H. Lee and Derek K. Shaeffer in [3] can be shown as Fig. 2-2. It concludes several main noise sources: channel thermal noise ($\overline{i_d^2}$), gate induced noise ($\overline{i_g^2} = \overline{i_{g,c}^2} + \overline{i_{g,u}^2}$), the noise due to the gate resistor ($\overline{i_{rg}^2}$) and the noise due to the resistor of the gate series inductor ($\overline{i_\ell^2}$). They can be written as the following expressions and discussed in each section:

- Channel thermal noise ($\overline{i_d^2}$)

$\overline{i_d^2} = 4kT\gamma g_{d0}\Delta f$, where γ is the coefficient of channel thermal noise.For long channel device, $\frac{2}{3} \le \gamma \le 1$. γ may be great than 1 in short channel device due to hot electrons effect, typically 2~3. For instance, γ=0.27 in typical 0.25 μm TSMC 1p5m CMOS technology.

■ Gate induced noise ($\overline{i_g^2} = \overline{i_{g,c}^2} + \overline{i_{g,u}^2}$)

$\overline{i_g^2} = 4kT\delta g_g \Delta f = \overline{i_{g,c}^2} + \overline{i_{g,u}^2} = 4kT\delta g_g \left(1 - |c|^2\right)\Delta f + 4kT\delta g_g |c|^2 \Delta f$, where δ is the coefficient of gate induced noise and $g_g = \frac{\omega^2 C_{gs}^2}{5 g_{d0}}$. The denoted c and u represent correlated and uncorrelated gate induced noise with channel thermal noise respectively. The value of c is about 0.395j for long channel device, but c=0.47j in typical TSMC 0.25 μm 1p5m CMOS technology.

■ Gate resistor ($\overline{i_{rg}^2}$)

The noise of gate resistor can be eliminated by using multi-finger device that parallel many resistor to form a smaller resistance that can be ignored.

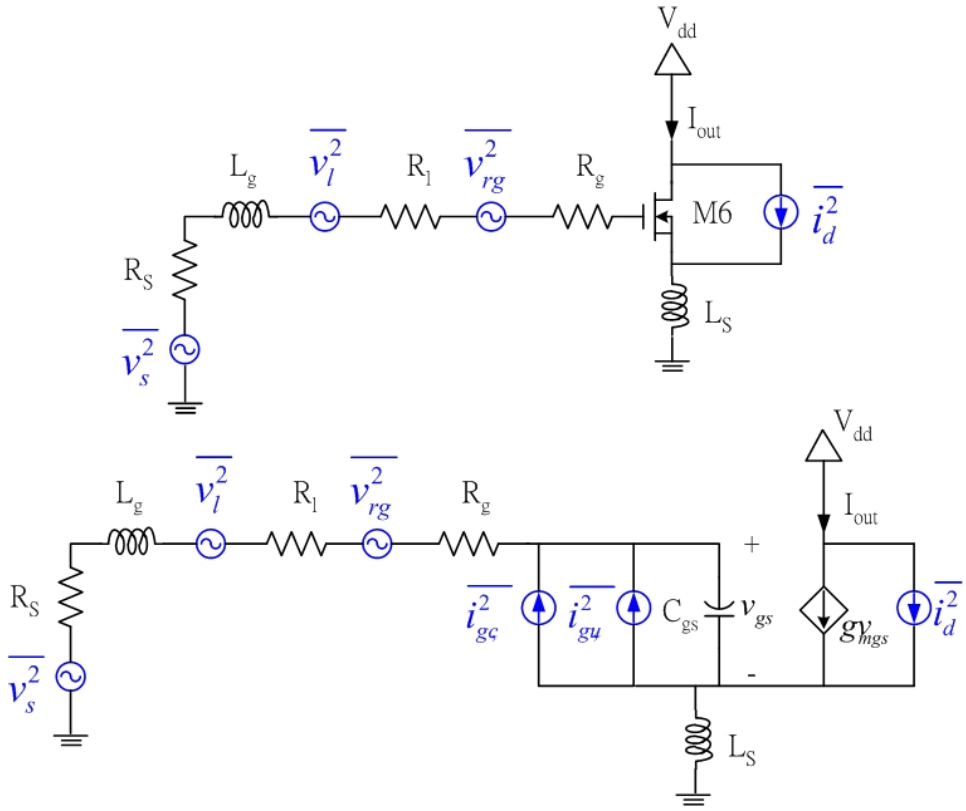

Fig. 2-2 Equivalent circuit of MOS for input stage noise calculation.

2.3 Topologies of low noise amplifier

In recent, there are four common LNA architectures as shown in Fig. 2-3, are the resistive termination, 1/gm termination, shunt-series feedback and inductive degeneration respectively. However, only the inductive degeneration architecture as shown in Fig. 2-3 (d) is the best choice in LNA design. Input matching is an important consideration for connection with external components. Focus on the requirement of providing a controlled input impedance typically 50Ω, we discuss the characteristics of each topology as follows:

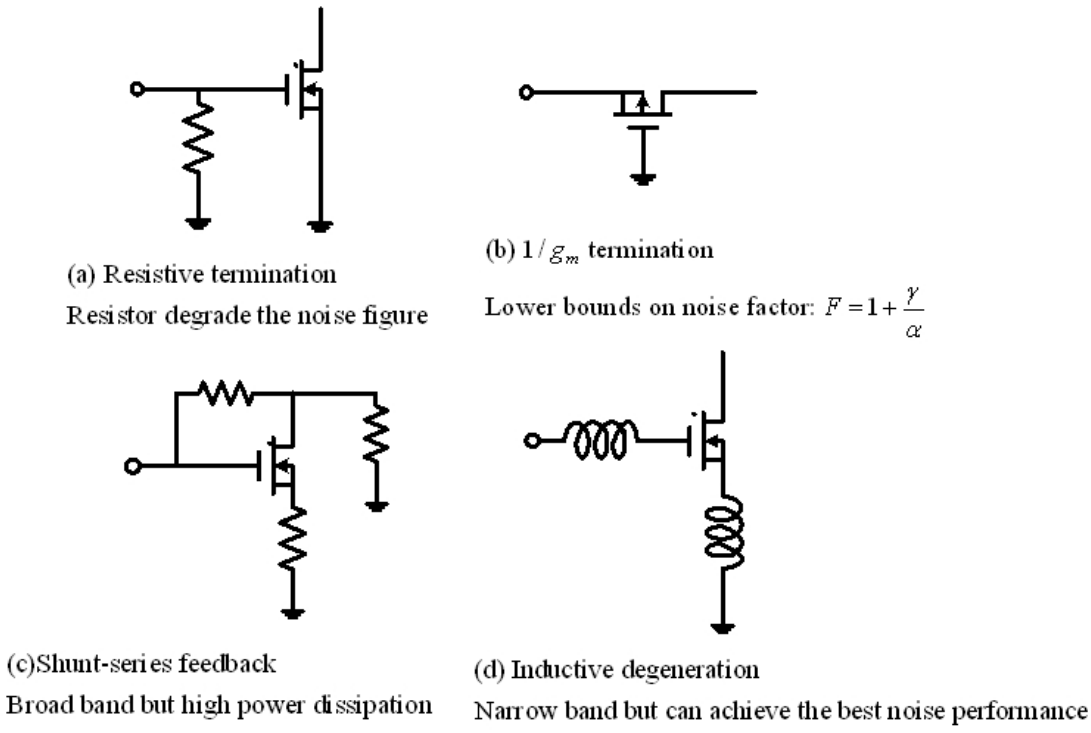

Fig. 2-3 Four topologies of low noise amplifiers.

- Resistive termination

The first type of LNA in Fig. 2-3 (a) is using resistive termination of the input port to provide a 50Ω impedance. The drawback of this topology is the use of real

resistors having a deleterious influence on the amplifier's noise finger. In other words, the added resistor termination contributes its own noise to the output directly. Hence it degrades the noise performance of the amplifier.

■ $\frac{1}{g_m}$ termination

The second topology of low noise amplifier in Fig. 2-3 (b) is a common-gate or common-base stage using the source or emitter as the input termination. With simplified analysis under matched condition, we can get the lower bound of noise figure shown below [3]:

Bipolar: $$F = 3/2 = 1.76\ dB \quad (2\text{-}1)$$

CMOS: $$F = 1 + \frac{\gamma}{\alpha} \geq \frac{5}{3} = 2.2\ dB \quad (2\text{-}2)$$

Where $\alpha \triangleq \frac{g_m}{g_{d0}}$, g_{d0} is the zero-bias drain conductance.

As mentioned before, for short channel device such as TSMC 0.25μm CMOS process, γ is significantly larger than 1 and α is less than 1. Hence the theoretical minimum achievable noise figure tends to be around 3dB or above.

■ Shunt-series feed back

Fig. 2-3 (c) illustrates the third topology of low noise amplifier, which uses resistive shunt-series feedback to match the input and output impedances of LNA. Higher power consumption is the major drawback of this LNA topology compared with other narrowband LNA topologies while achieve similar noise figure. Intuitively, the higher power dissipation is partially due to the fact that shunt-series feedback amplifiers of this type are naturally broadband, and hence the techniques which reduce the power consumption through L-C tuning are not applicable.

■ Inductive degeneration

The inductive degeneration architecture as shown in Fig. 2-3 (d) is the best choice in LNA design because it has two advantages: (1) the source inductor L_s provides a real part impedance $\mathrm{Re}[Z_{in}]$ for matching to generator impedance R_s usually is 50Ω at resonance condition. Refer to Fig. 2-4, the input impedance Z_{in} can be approximated to $Z_{in} = \mathrm{Re}[Z_{in}] + j \bullet \mathrm{Im}[Z_{in}] = \left(\frac{g_m}{C_{gs}}\right) L_s + \left[s(L_s + L_g) + \frac{1}{sC_{gs}} \right]$.

At resonance condition, then $Z_{in} \simeq \frac{g_m}{C_{gs}} L_s = \omega_T L_s = R_s$ $\left(here\ \omega_T = \frac{g_m}{C_{gs}} \right)$, ω_T is the NMOS cutoff frequency, If the effect of the gate-to-drain overlap capacitance is included, $Z_{in} \approx \frac{\omega_T L_s}{1 + 2\frac{C_{gd}}{C_{gs}}} = \omega_{T,eff} L_s$, Z_{in} shows a real part resistance in the input stage, but L_s does not result in thermal noise as a resistor, so this topology can mitigate the noise effects due to use resistive termination. In other words, the series feedback allows the use of almost noiseless reactive elements to produce a real input impedance at the carrier frequency ω_0 ($\omega_0 = \frac{1}{\sqrt{(L_g + L_S) C_{gs}}}$). Moreover, the source inductor L_s consumes no DC power since it is taken as short-circuited in DC analysis. (2) The source inductor L_s as feedback resistance at source terminal somewhat improves the linearity requirement while degrade the amplifier's power gain.

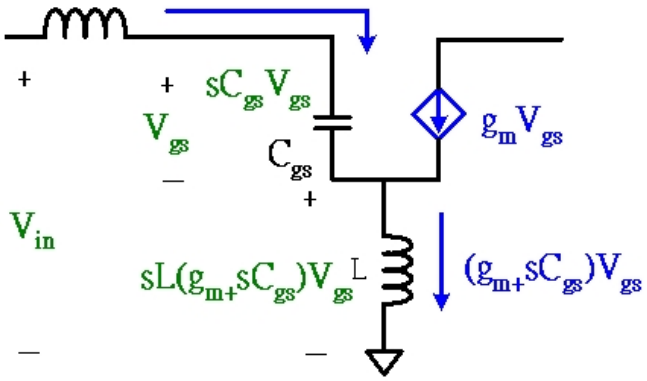

Fig. 2-4 Input impedance calculation of inductive degeneration.

■ Cascode amplifier (common source with common gate)

A novel high gain and low noise amplifier topology shown as Fig. 2-5 is a cascode amplifier with good isolation, high stability and good matching condition. In this topology, transistor M1 is used as common source with inductive source degeneration, and transistor M2 is used as a common gate circuit. This topology can provide good matching conditions for noise and gain simultaneously. It can also reduce the influence of Miller's effect due to the gate-to-drain overlap capacitance C_{gd1} and improves the linearity requirement due to use the inductor L_s as feedback component at source terminal. In addition, the cascode topology has two additional advantages of improving the reverse isolation and high stability. Today, the cascode amplifier with inductive source degeneration has been widely applied to narrowband LNA designs in RFIC/MMIC circuits because it can satisfy both noise and power gain requirement easily.

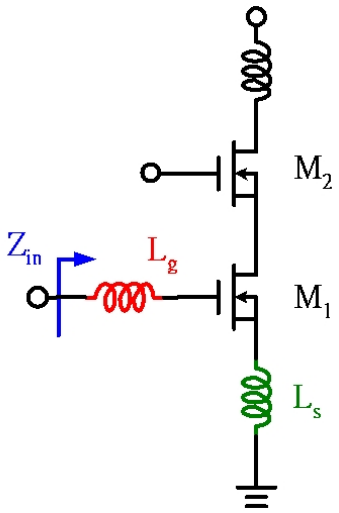

Fig. 2-5 Cascode amplifier with inductive source degeneration.

■ Fully differential amplifier

The drawback of single-ended LNA topology is its sensitivity to the parasitic ground inductance and resistance. See Fig. 2-5, small amount of additional parasitic resistance of source inductor can have a large effect on amplifier performance. It forms a parasitic feedback loop that destabilizes the amplifier. To overcome this problem, we can apply more additional ground pins or use sophisticated packages. But such measures may be either of limited effectiveness and somewhat costly.

Instead of using above techniques, adopting symmetrical differential structure is another choice. It can achieve a virtual ground at the symmetrical point to eliminate the effects due to parasitic ground inductance and resistance. Fig. 2-6 is an example circuit with differential structure. Another important advantage of the differential structure is the ability to reject common mode noise. This consideration is particularly important in mix-mode circuits since the supply and substrate voltages may be noisy. To carry out good common-mode rejection at high frequency, we must let the circuit layout as symmetrical as possible.

An important property that we must keep in mind, the noise figure of differential

type amplifier is higher than its single-ended type amplifier while consuming equal power. In other words, the differential amplifier consumes twice power compared to its single-ended amplifier while achieving the same noise figure. Although the differential structure has above disadvantage, but trades-off the improved linearity. Hence the dynamic range increases even though the noise figure remains constant while assuming twice the power consumption.

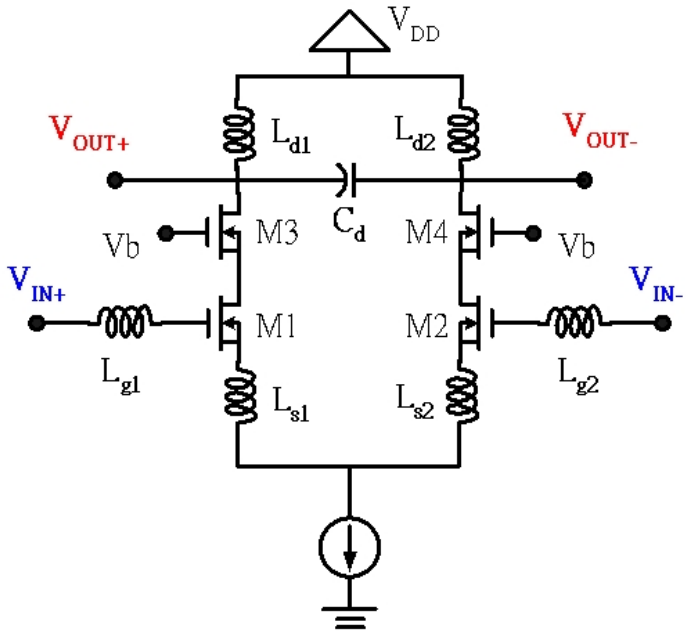

Fig. 2-6 Fully differential amplifier with inductive source degeneration.

2.4 Design considerations of low noise amplifier

The LNA usually has the specifications:

(a) Noise figure (NF)

(b) Stability

(c) Forward gain (S_{21})

(d) Input reflection coefficient (S_{11})

(e) Output reflection coefficient (S_{22})

(f) Power consumption (P_D)

(g) The linear output power IP3 (Input-referred 3rd-order intercept)

(h) P_{-1dB} (1dB compression point)

■ Noise figure (NF)

Noise figure has been defined in a number of different ways, the most commonly accepted definition is

$$\text{Noise factor (F)} = \frac{SNR_{(input)}}{SNR_{(output)}} = \frac{\textit{Total noise power}}{\textit{Total noise power from the source}}$$

$$\text{Noise figure (NF)} = 10\log(F) \geqq 0 \quad \text{, in dB unit.} \tag{2-3}$$

Physically, the noise figure is a measure of how much the SNR degrades as the signal passes through a system. It is desired that noise figure as small as possible since we hope the system adds few noise to the input signal. For instance, if a system has no noise, then noise figure is equal to unity the ideal case we wanted.

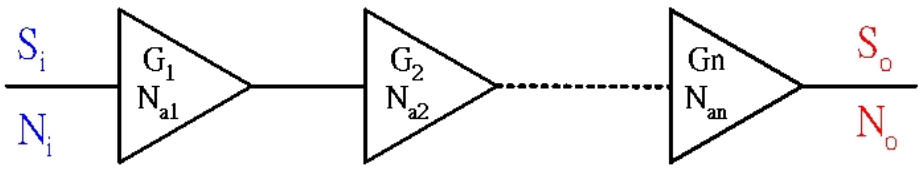

Fig. 2-7 Noise figure calculation of a cascade stage.

For a cascade stage shown as Fig. 2-7, the overall noise figure of full system can be obtained by the following calculation:

The total output signal $S_o = S_i G_1 G_2 G_3 .....G_n$

The total output noise $N_o = N_i G_1 G_2 G_3 .....G_n + N_{a1} G_2 G_3 .....G_n + ... + N_{an}$

Then noise factor $F = \dfrac{S_i / N_i}{S_o / N_o} = \dfrac{S_i / N_i}{\dfrac{S_i G_1 G_2 G_3 .....G_n}{N_i G_1 G_2 G_3 .....G_n + N_{a1} G_2 G_3 .....G_n + ... + N_{an}}}$

$$= \frac{S_i N_i G_1 G_2 G_3 .....G_n + S_i N_{a1} G_2 G_3 .....G_n + ... + S_i N_{an}}{S_i N_i G_1 G_2 G_3 .....G_n}$$

$$= 1 + \frac{N_{a1}}{N_i G_1} + \frac{N_{a2}}{N_i G_1 G_2} + ... + \frac{N_{an}}{N_i G_1 G_2 G_3 .....G_n}$$

Rearrange above result,

$$\Rightarrow F = 1 + \frac{(F_1 - 1) N_i G_1}{N_i G_1} + \frac{(F_2 - 1) N_i G_2}{N_i G_1 G_2} + ... + \frac{(F_n - 1) N_i G_n}{N_i G_1 G_2 G_3 .....G_n}$$

$$= 1 + (F_1 - 1) + \frac{(F_2 - 1)}{G_1} + \frac{(F_3 - 1)}{G_1 G_2} + ... + \frac{(F_n - 1)}{G_1 G_2 G_3 .....G_{n-1}}$$

$$= F_1 + \frac{(F_2 - 1)}{G_1} + \frac{(F_3 - 1)}{G_1 G_2} + ... + \frac{(F_n - 1)}{G_1 G_2 G_3 .....G_{n-1}} \qquad (2\text{-}4)$$

where F_n is the noise factor of n^{th} stage evaluated with respect to the driving impedance of the preceding stage and G_n is the available power gain of the n^{th} block.

The answer why low noise amplifier is the first stage of a wireless receiver can be known from Eq. (2-4) since the noise factor of 1st stage is added directly to the total noise value without degrading by gain. The noises of following stages can be reduced by the gain of preceding stage, hence it becomes the more minor factors than 1st stage. Signals from antenna suffer from strong noise interference, so low noise amplifier needs to mitigate noise as much as possible. According to Eq. (2-4), the total noise figure of a system can be degraded while the gain of first stage LNA is large.

■ Gain

In design, several equations of power gain are defined and expressed as shown in Fig. 2-8 with the forms [11] [12]:

$$G_T = \frac{power\ delivered\ to\ the\ load}{power\ available\ from\ source} = \frac{P_L}{P_{AVS}} = \frac{1-|\Gamma_S|^2}{|1-\Gamma_{IN}\Gamma_S|^2}|S_{21}|^2\frac{1-|\Gamma_L|^2}{|1-\Gamma_{OUT}\Gamma_L|^2} \quad (2\text{-}5)$$

$$G_P = \frac{power\ delivered\ to\ the\ load}{power\ input\ to\ the\ network} = \frac{P_L}{P_{IN}} = \frac{1}{|1-\Gamma_{IN}|^2}|S_{21}|^2\frac{1-|\Gamma_L|^2}{|1-S_{22}\Gamma_L|^2} \quad (2\text{-}6)$$

$$G_A = \frac{power\ available\ from\ the\ network}{power\ available\ from\ the\ source} = \frac{P_{AVN}}{P_{AVS}} = \frac{1-|\Gamma_S|^2}{|1-S_{11}\Gamma_S|^2}|S_{21}|^2\frac{1}{|1-\Gamma_{OUT}|^2} \quad (2\text{-}7)$$

Where $\Gamma_{IN} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1-S_{22}\Gamma_L}$ and $\Gamma_{OUT} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1-S_{22}\Gamma_S}$

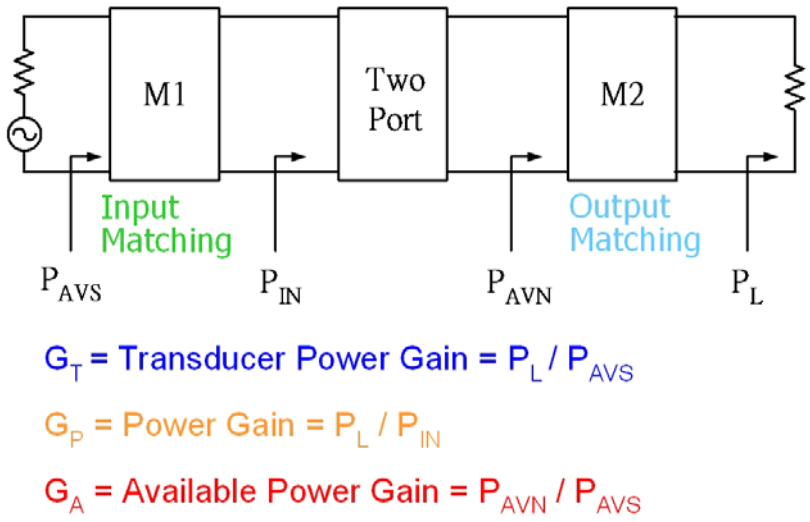

Fig. 2-8 Definitions of power gain.

■ Linearity (IIP3, OIP3, 1dB compression point)

Dynamic range is one of the important requirements in currently wireless communication systems. The most useful measures for linearity are third-order intercept (IP3) and 1dB compression point (P_{1dB}). We first consider the characteristics of gain compression. Gain compression is a useful index of distortion generation. It provides a good indication of the signal amplitude that an amplifier or a mixer will tolerate before really bad distortion is generated. It is specified in terms of an input

power level at which the small signal gain drops off by 1dB.

Amplifiers are nonlinear devices, and the output signal of an amplifier can be expressed as

$$v_o = \alpha_1 v(t) + \alpha_2 v^2(t) + \alpha_3 v^3(t) + ... \quad (2\text{-}8)$$

If an input signal $v(t) = A\cos(\omega_1 t)$ is applied to the amplifier, the output signal becomes

$$v_o(t) = \alpha_1 A\cos(\omega_1 t) + \alpha_2 A^2 \cos^2(\omega_1 t) + \alpha_3 A^3 \cos^3(\omega_1 t) + ...$$

The value of coefficient α_3 is usually negative. We concern the fist term and third term:

$$v_o(t) = \alpha_1 - |\alpha_3| A^3 \cos^3(\omega_1 t) \quad (2\text{-}9)$$

We can predict the 1dB compression input voltage while we know the value of coefficient α_3.

Second, we consider the characteristics of intermodulation distortion (IMD). Intermodulation distortion (IMD) is a measure of large-signal capability. It occurs when two or more signals are presented at the input to the amplifier. These two signals can interact with the nonlinearities to generate unwanted IMD products (distortion).

When two input signals $v(t) = A\cos(\omega_1 t) + B\cos(\omega_2 t)$ is applied to the amplifier, the output signal become

$$v_o = \alpha_1 + \alpha_2 \left[A\cos(\omega_1 t) + B\cos(\omega_2 t)\right]^2 + \alpha_3 \left[A\cos(\omega_1 t) + B\cos(\omega_2 t)\right]^3 + ...$$

Rearrange it, then

$$v_o(t) = \alpha_1 + 2\alpha_2 AB\cos(\omega_1 t)\cos(\omega_2 t) + \alpha_2 \left[A^2 \cos^2(\omega_1 t) + B^2 \cos^2(\omega_2 t)\right]$$

$$+\alpha_3\left[A^3\cos^3(\omega_1 t)+B^3\cos^3(\omega_2 t)\right]+3\alpha_3 A^2 B\cos^2(\omega_1 t)\cos(\omega_2 t)$$

$$+3\alpha_3 AB^2\cos(\omega_1 t)\cos^2(\omega_2 t)+\ldots\ldots \quad (2\text{-}10)$$

We rewrite the 5th term since we mainly concern with the third-order IMD:

$$3\alpha_3 A^2 B\cos^2(\omega_1 t)\cos(\omega_2 t)=3\alpha_3 A^2 B\bullet\cos(\omega_1 t)\bullet\frac{\cos(\omega_1+\omega_2)t+\cos(\omega_1-\omega_2)t}{2}$$

$$=\frac{3\alpha_3 A^2 B}{2}\left[\cos(\omega_1 t)\cos(\omega_1+\omega_2)t+\cos(\omega_1 t)\cos(\omega_1-\omega_2)t\right]$$

$$=\frac{3\alpha_3 A^2 B}{2}\left\{\frac{1}{2}\left[\cos(2\omega_1+\omega_2)t+\cos(\omega_2 t)\right]+\frac{1}{2}\left[\cos(2\omega_1-\omega_2)t+\cos(\omega_2 t)\right]\right\}$$

$$=\frac{3\alpha_3 A^2 B}{2}\left\{\cos(\omega_2 t)+\frac{1}{2}\left[\cos(2\omega_1+\omega_2)t+\cos(2\omega_1-\omega_2)t\right]\right\} \quad (2\text{-}11)$$

= (Cross-modulation) + (Third-order IMD)

Cross-modulation term: $\frac{3\alpha_3 A^2 B}{2}\cos(\omega_2 t)$ (2-12)

Third-order IMD term: $\frac{3\alpha_3 A^2 B}{2}\bullet\frac{1}{2}\left[\cos(2\omega_1+\omega_2)t+\cos(2\omega_1-\omega_2)t\right]$ (2-13)

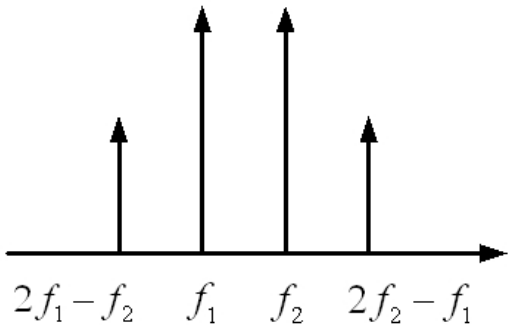

Fig. 2-9 Third-order intermodulation product.

According to Eq. (2-13) and Fig. 2-9, the third order products at $2f_1-f_2$ and $2f_2-f_1$ are obtained. These two components may fall within the signal bandwidth and thus cause interference to a desired signal. IMD power has a slope of 3 on a dB plot.

In addition, the cross-modulation effect can also be seen from Eq. (2-12). The amplitude of on signal (say ω_1) influences the amplitude of the desired signal at ω_2 through the coefficient $\frac{3\alpha_3 A^2 B}{2}$. This cross-modulation can have annoying or error generating effects at the output. Other higher odd-order intermodulation products, such as 5th and 7th, are also of interest. Fig. 2-10 is a diagram used to illustrate IP3 and 1dB compression point.

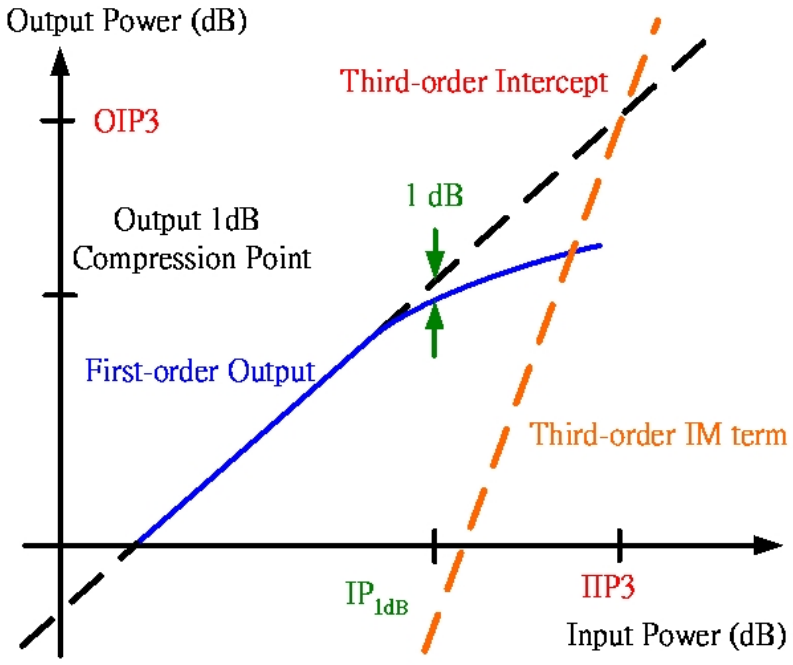

Fig. 2-10 Definitions of 1dB gain compression and IIP3 [1] [2].

Similar to the noise figure of cascade stage discussed before, the IIP3 of cascade stage can be expressed as [1]:

$$\frac{1}{A_{IP3,total}^2} \approx \frac{1}{A_{IP3,1}^2} + \frac{\alpha_1^2}{A_{IP3,2}^2} + \frac{\alpha_1^2 \beta_1^2}{A_{IP3,3}^2} + ...$$

$$\text{or} \quad \frac{1}{IIP3} \approx \frac{1}{IIP3_1} + \frac{G_{a1}}{IIP3_2} + \frac{G_{a2}}{IIP3_3} + ... \qquad (2\text{-}14)$$

According to above equation, the nonlinearity of last stage is most critical since the IIP3 of each stage is scaled down by the gain preceding that stage.

2.5 Modified noise analysis of MOSFET

Thomas H. Lee and Derek K. Shaeffer suggest an extended noise model of MOSFET [3] to estimate the noise figure of LNA. Following, we give a short illustration about this noise analysis method.

Recall Fig. 2-2, Fig. 2-5 and Eq. (2-3), the input impedance and noise figure of a cascode amplifier are respectively represented by

$$Z_{in} = \left(\frac{g_{m1}}{C_{gs}}\right)L_s + \left[s\left(L_s + L_g\right) + \frac{1}{sC_{gs}}\right]$$

$$NF = \frac{Total\ noise\ power}{Total\ noise\ power\ from\ the\ source}$$

To estimate the output noise, the first work is the evaluation of the transconductance of the input stage.

$$G_m = g_{m1}Q_{in} = \frac{g_{m1}}{\omega_0 C_{gs}\left(R_s + \omega_T L_s\right)} = \frac{\omega_T}{\omega_0 L_s\left(1 + \frac{\omega_T L_s}{R_s}\right)} = \frac{\omega_T}{2\omega_0 L_s} \tag{2-15}$$

Here, ω_0 is the resonant frequency, Q_{in} is the effective Q of the input stage.

Refer to Fig. 2-2 and modified noise model illustrated in Chapter 2.2, the output noise power density due to the source is

$$S_{a,src}\left(\omega_0\right) = S_{src}\left(\omega_0\right)G_{m,eff}^2 = \frac{4kT\omega_T^2}{\omega_0^2 R_s\left(1 + \frac{\omega_T L_s}{R_s}\right)^2} \tag{2-16}$$

Similar to above, the output noise power density due to R_l and R_g can be expressed as

$$S_{a,R_l,R_g}(\omega_0)=\frac{4kT\left(R_l+R_g\right)\omega_T^2}{\omega_0^2R_s^2\left(1+\frac{\omega_TL_s}{R_s}\right)^2} \tag{2-17}$$

The output noise power density due to channel current noise of MOSFET is

$$S_{a,i_d}(\omega_0)=\frac{\frac{i_d^2}{\Delta f}}{\left(1+\frac{\omega_TL_s}{R_s}\right)^2}=\frac{4kT\gamma g_{d0}}{\left(1+\frac{\omega_TL_s}{R_s}\right)^2} \tag{2-18}$$

As illustrated in Chapter 2.2, the induced gate noise is the combinations of correlated and uncorrelated gate induced noise. The combined power density of drain noise with correlated gate induced noise is

$$S_{a,i_d,i_{g,c}}(\omega_0)=\kappa S_{a,i_d}(\omega_0)=\frac{4kT\gamma\kappa g_{d0}}{\left(1+\frac{\omega_TL_s}{R_s}\right)^2} \tag{2-19}$$

where $\kappa=\frac{\delta\alpha^2}{5\gamma}|c|^2+\left[1+|c|\sqrt{\frac{\delta\alpha^2}{5\gamma}}\right]^2$ (2-20)

The uncorrelated gate induced noise power density is

$$S_{a,i_{g,u}}(\omega_0)=\xi S_{a,i_d}(\omega_0)=\frac{4kT\gamma\xi g_{d0}}{\left(1+\frac{\omega_TL_s}{R_s}\right)^2} \tag{2-21}$$

where $\xi=\frac{\delta\alpha^2}{5\gamma}\left(1-|c|^2\right)\left(1+Q_S^2\right)$ (2-22)

$$Q_S=\frac{\omega_0\left(L_S+L_g\right)}{R_S}=\frac{1}{\omega_0R_SC_{gs}} \tag{2-23}$$

Hence the total current noise of M1 is

$$S_{a,M_1}(\omega_0)=(\kappa+\xi)S_{a,i_d}(\omega_0)=\chi S_{a,i_d}(\omega_0)=\frac{4kT\gamma\chi g_{d0}}{\left(1+\frac{\omega_T L_s}{R_s}\right)^2} \quad (2\text{-}24)$$

where $\chi=\kappa+\xi=1+2|c|\sqrt{\frac{\delta\alpha^2}{5\gamma}}+\frac{\delta\alpha^2}{5\gamma}\left(1+Q_S^2\right)$

(2-25)

The total noise figure of M1 is the sum of (2-16), (2-17) and (2-24), then

$$F=\frac{S_{a,source}(\omega_0)+S_{a,R_l,R_g}(\omega_0)+S_{a,M_1}(\omega_0)}{S_{a,source}(\omega_0)}=1+\frac{R_l}{R_s}+\frac{R_g}{R_s}+\gamma\chi g_{d0}R_S\left(\frac{\omega_0}{\omega_T}\right)^2 \quad (2\text{-}26)$$

Noting that $g_{d0}Q_S=\frac{g_m}{\alpha}\frac{1}{\omega_0 R_S C_{gs}}=\left(\frac{g_m}{C_{gs}}\right)\bullet\left(\frac{1}{\alpha\omega_0 R_S}\right)=\frac{\omega_T}{\alpha\omega_0 R_S}$, then

$g_{d0}R_S=\frac{\omega_T}{\alpha\omega_0 Q_S}$.

Substitute into (2-26), we can repress (2-26) as

$$F=1+\frac{R_l}{R_s}+\frac{R_g}{R_s}+\frac{\gamma}{\alpha}\frac{\chi}{Q_s}\left(\frac{\omega_0}{\omega_T}\right) \quad (2\text{-}27)$$

The coefficient χ according to (2-25) includes terms which are constant and terms which are proportional to Q_S^2. From (2-27), we see the noise figure will contain terms which are proportional to Q_S and terms which are inversely proportional to Q_S. Therefore, a minimum F exists for a particular Q_S.

2.6 Design steps of low noise amplifier

LNA design is full of trade-offs between optimum gain, lowest noise figure, input and output port matching, linearity and power consumption. According to the selected specification such as IEEE 802.11 Families or Bluetooth, we can specify the performance needed such as gain, noise figure, power consumption, and linearity. A

popular and useful design procedure presented by Thomas H. Lee and Derek K. Shaeffer [2] [3] will be used to optimize the performance of low noise amplifier. The design steps of basic cascode LNA with inductive source degeneration are illustrated as following:

(a) Specify the performance needed such as gain, noise figure, linearity and power consumption.

(b) Construct circuit architecture and choose MOS size. Determining the device size is an important thing for obtaining low noise and input matching condition under specified power consumption condition. The first step, we need to set up the power consumption of LNA. Because the noise performance of a cascode amplifier with inductive source degeneration is mainly determined by the transistor size of input stage M1 (see Fig. 2-5), the key point is to determine the size of M1 under specified power consumption condition. Now, we give a short illustration about how to determine the device size here.

■ Noise figure optimization technique – fixed P_D optimization [3]

Here, we review the noise figure optimization technique under fixed P_D condition that presented by Thomas H. Lee and Derek K. Shaeffer [3].

In Eq. (2-27), the term $\left(\frac{R_l}{R_s}+\frac{R_g}{R_s}\right)$ can be neglected since $\frac{R_l}{R_s},\frac{R_g}{R_s}\ll 1$, then

$$F=1+\frac{\gamma}{\alpha}\frac{\chi}{Q_s}\left(\frac{\omega_0}{\omega_T}\right) \tag{2-28}$$

The drain current can be expressed as $I_d=WC_{ox}v_{sat}\frac{V_{od}^2}{V_{od}+L\varepsilon_{sat}}$ (2-29)

Then differentiate (2-29) to get the transconductance

$$g_m = \frac{\partial I_d}{\partial V_{gs}} = \mu_{eff} C_{ox} \frac{W}{L} V_{od} \left[\frac{1+\frac{\rho}{2}}{(1+\rho)^2} \right] \text{ , here } \rho \triangleq \frac{V_{od}}{L\varepsilon_{sat}} \tag{2-30}$$

Hence, we can formulate the power consumption as follows,

$$P_D = V_{dd} I_d = V_{dd} W C_{ox} v_{sat} \frac{V_{od}^2}{V_{od} + L\varepsilon_{sat}} \tag{2-31}$$

Note that the P_D is proportional to W. We also see that Q_S having a dependence on W from Eq. (2-23), then we can link P_D and Q_S together. Hence,

$$Q_S = \frac{P_0}{P_D} \frac{\rho^2}{1+\rho} \text{ , here } P_0 = \frac{3}{2} \frac{V_{dd} v_{sat} \varepsilon_{sat}}{\omega_0 R_s} \tag{2-32}$$

Besides, $\omega_T \approx \frac{g_m}{C_{gs}} = \frac{3\alpha\rho v_{sat}}{L}$ (2-33)

Combine Eq. (2-30), Eq. (2-32), Eq. (2-33) and substitute into (2-28), after some manipulations, we can get the following expression.

$$F = 1 + \frac{\gamma\omega_0 L}{3v_{sat}} P(\rho, P_D) \tag{2-34}$$

where $P(\rho, P_D) = \frac{\frac{P_D}{P_0} P_1(\rho) + \frac{P_0}{P_D} P_2(\rho)}{\rho^3 \left(1+\frac{\rho}{2}\right)^2 (1+\rho)}$ (2-35)

$$P_1(\rho) = (1+\rho)^6 + 2|c| \sqrt{\frac{\delta}{5\gamma}} \left(1+\frac{\rho}{2}\right)(1+\rho)^4 + \frac{\delta}{5\gamma}(1+\rho)^2 \left(1+\frac{\rho}{2}\right)^2 \tag{2-36}$$

$$P_2(\rho) = \frac{\delta}{5\gamma} \left(1+\frac{\rho}{2}\right)^2 \rho^4 \tag{2-37}$$

Normally $\rho \ll 1$, hence

$$P(\rho, P_D) \approx \frac{\frac{P_D}{P_0}\left(1 + 2|c|\sqrt{\frac{\delta}{5\gamma}} + \frac{\delta}{5\gamma}\right) + \frac{P_0}{P_D}\frac{\delta}{5\gamma}\rho^4}{\rho^3} \tag{2-38}$$

Under fixed-power consumption condition, we can differentiate above equation.

$$\frac{\partial P(\rho, P_D)}{\partial \rho} = 0$$

$$\Rightarrow \quad \rho^2_{opt,P_D} = \frac{P_D}{P_0}\sqrt{3}\left\{1+2|c|\sqrt{\frac{5\gamma}{\delta\alpha^2}}+\frac{5\gamma}{\delta\alpha^2}\right\}^{\frac{1}{2}} \tag{2-39}$$

Substitute into Eq. (2-32), hence

$$Q_{s,opt,P_D} = \sqrt{3}\left\{1+2|c|\sqrt{\frac{5\gamma}{\delta\alpha^2}}+\frac{5\gamma}{\delta\alpha^2}\right\}^{\frac{1}{2}} \tag{2-30}$$

After getting Q_{s,opt,P_D}, we can evaluate χ according to Eq. (2-25). Substitute χ into (2-28), hence we get

$$F_{\min,P_D} = 1+\sqrt{\frac{16}{15}\delta\gamma}\left(\frac{\omega_0}{\omega_T}\right)\left\{1+2|c|\sqrt{\frac{\delta\alpha^2}{5\gamma}}+\frac{\delta\alpha^2}{5\gamma}\right\}^{\frac{1}{2}} \tag{2-31}$$

Finally, the optimum width of MOSFET can be obtain by

$$W_{opt,P_D} = \frac{3}{2}\frac{1}{\omega L C_{ox} R_s Q_{s,opt,P_D}} \tag{2-32}$$

Hence, as long as we know the process parameters, we can determine the optimum Q value and optimum width of MOSFET according to Eq. (2-30) and Eq. (2-32) under the assumed power consumption condition.

Worthy of mentioned, the optimum conditions are not extremely sensitive to the width since a variation of 20% one way or the other will only degrade the noise figure by one or two tenths of a decibel.

For example, under some given process parameters, we can perform Fig. 2-11 and Fig. 2-12 used to select the best width of MOSFET and forecast the minimum noise figure we can achieve.

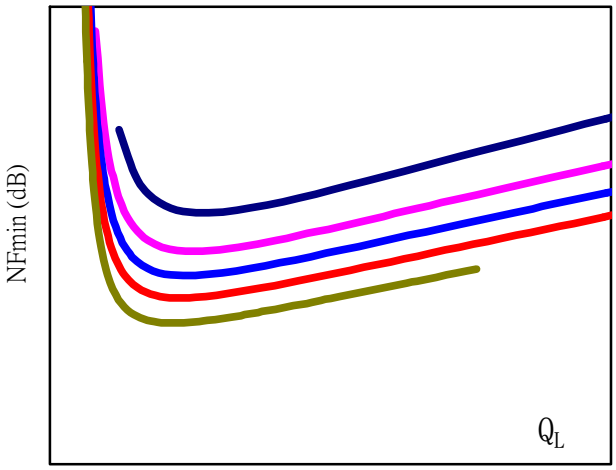

Fig. 2-11 Determine the optimum Q value (schematic).

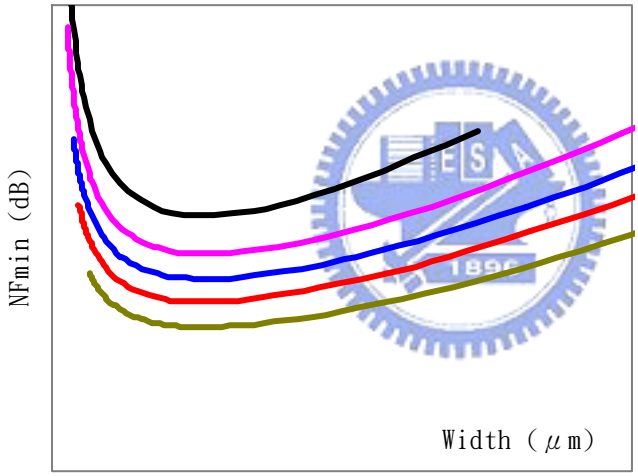

Fig. 2-12 Determine the optimum width of MOSFET (schematic).

■ Note

Noting that, above method is correct only if the MOS devices are in the strong inversion region of operation. Actually, for typical current value (≦10mA) and a scaled down technology (below 0.5μm), the devices work in the moderate inversion region at the optimum bias point. Francesco Gatta [13] provides a modified method to determine the optimum width:

$$W_{opt,P_D} = \frac{K(I)}{\omega_0 L_{eff}^{0.6} C_{ox}^{1.2} R_S^{0.8}} \quad (2\text{-}33)$$

Where $K(I)$ is a constant, dependent on biasing current I and different for PMOS and NMOS devices. The key point is that the optimum width is inversely proportional to the frequency and almost inversely proportional to the source resistance. Moreover, for the same frequency and source resistance, going from 0.35μm to 0.18μm technology, the optimum gate width decreases by 25%.

(c) Second, set the bias to meet the assumed power consumption condition, and determine the S parameters and characteristics of the transistors.

(d) Check stability in a wide frequency ranges (0~5 GHz). Do some compensation if it is not stable. For instance, in a cascode amplifier, we can increase the power gain by wider the size of the cascode M2 shown in Fig. 2-5, but this action will decrease the stability, so we must trade-off between them.

(e) To compromise the performance between gain, noise figure, linearity, isolation, and power consumption in matching networks. All performances must fit the specifications of wireless communication standard we chosen.

(f) Complete the circuit, and tuning for optimization.