A 1.5-V, 1.5-GHz CMOS Low Noise Amplifier

Derek K. Shaeffer, *Student Member, IEEE*, and Thomas H. Lee, *Member, IEEE*

Abstract—**A 1.5-GHz low noise amplifier (LNA), intended for use in a global positioning system (GPS) receiver, has been implemented in a standard 0.6-μm CMOS process. The amplifier provides a forward gain (S21) of 22 dB with a noise figure of only 3.5 dB while drawing 30 mW from a 1.5 V supply. In this paper, we present a detailed analysis of the LNA architecture, including a discussion on the effects of induced gate noise in MOS devices.**

Index Terms—**Amplifier noise, induced gate noise, low noise amplifier, microwave amplifier, MOSFET amplifier, noise figure, random noise, semiconductor device noise.**

I. Introduction

Radio frequency designs are increasingly taking advantage of technology advances in CMOS that make possible the integration of complete communications systems. As an example, global positioning system (GPS) receivers employ extensive digital signal processing to perform acquisition, tracking, and decoding functions. The use of CMOS technologies for implementation of the front end electronics in a GPS system is therefore attractive because of the promise of integrating the whole system on a single chip.

The first step in achieving this goal is to test the suitability of present-day CMOS for the task of low noise amplification at multigigahertz frequencies. Received GPS signal power levels at the antenna are around −130 dBm, and this low level degrades further in the presence of physical obstructions such as buildings and trees. Hence, a good amplifier is critical for enabling robust performance in obstructed environments.

One possible threat to low noise operation is the well-documented, but relatively unappreciated, excess thermal noise exhibited by submicron CMOS devices [1]–[4]. This noise is believed to arise from hot electron effects in the presence of high electric fields. Despite this excess noise, recent work has demonstrated the viability of CMOS low noise amplifiers (LNA's) at frequencies around 900 MHz [5]–[7]. As we will show, CMOS is also a suitable medium for implementing a GPS receiver, which must receive signals centered at 1.575 42 GHz.

To provide some background, Section II presents a review of recent LNA work in various technologies in the 900 MHz–2 GHz frequency range. A thorough mathematical treatment of the LNA architecture that we have chosen is presented in Section III. It is our hope that this treatment will be useful as a guide in future design efforts. In pursuing this goal, we will consider the effect of induced gate noise in CMOS, which is rarely cited but nonetheless of fundamental importance in establishing the limits of achievable noise performance. In Section IV, noise figure optimization techniques are discussed which permit selection of device geometries to maximize noise performance for a specified gain or power dissipation. In addition, numerical examples, employing the analytical techniques developed in this paper, illustrate some of the salient features of the LNA architecture. Implementation details are discussed in Section V, while Section VI presents experimental results.

Manuscript received August 20, 1996; revised November 24, 1996.

The authors are with the Center for Integrated Systems, Stanford University, Stanford, CA 94305 USA.

Publisher Item Identifier S 0018-9200(97)03419-7.

II. Recent LNA Research

Many authors have investigated LNA techniques in the 900 MHz–2 GHz frequency range. Table I summarizes the results of several recent studies dating from 1991–1996. This table has representative results from various process technologies and architectures. While the literature is full of examples of LNA work in GaAs and bipolar technologies, there are few examples of CMOS studies. The four references shown here are the only ones of which we are aware. In addition, despite a long history of LNA work in GaAs and bipolar technologies, these papers report a wide variety of noise figures, power dissipations, and gains. The remarkable spread in published results seems to suggest that a rational basis for the design of these amplifiers has not been elucidated. However, by examining these results from an architectural viewpoint, some order emerges.

In the design of low noise amplifiers, there are several common goals. These include minimizing the noise figure of the amplifier, providing gain with sufficient linearity—typically measured in terms of the third-order intercept point, IP3—and providing a stable 50 Ω input impedance to terminate an unknown length of transmission line which delivers signal from the antenna to the amplifier. A good input match is even more critical when a preselect filter precedes the LNA because such filters are often sensitive to the quality of their terminating impedances. The additional constraint of low power consumption which is imposed in *portable* systems further complicates the design process.

With these goals in mind, we will first focus on the requirement of providing a stable input impedance. The architectures in Table I can be divided into four distinct approaches, illustrated in simplified form in Fig. 1. Each of these architectures may be used in a single-ended form (as shown), or in a differential form. Note that differential forms will require the use of a balun or similar element to transform the single-ended signal from the antenna into a differential signal. Practical baluns introduce extra loss which adds directly to the noise figure of the system.

0018–9200/97$10.00 © 1997 IEEE

TABLE I
SUMMARY OF RECENT LNA RESULTS

Author [Ref.]	NF (dB)	Gain (dB)	IP3/-1dB[a] (dBm)	Power (mW)	f_0 (GHz)	Architecture	Technology	Year
Chang *et al.* [8]	6.0[b]	14	na / na	7	0.75	R-Term.	2μm CMOS	1993
Karanicolas *et al.* [5]	2.2	15.6	12.4 / na	20	0.9	L-Degen.	0.5μm CMOS	1996
Sheng *et al.* [7]	7.5	11.0	na / na	36	0.9	Shunt-Ser. FB	1μm CMOS	1996
Rofougaran *et al.* [6]	3.5	22	na / na	27	0.9	$1/g_m$-Term.	1μm CMOS	1996
Benton *et al.* [9]	2.7	28	na / 8.5	208	1.6	Shunt-Ser. FB	GaAs FET	1992
Cioffi [10]	2.2	17.4	na / na	10	1.6	L-Degen.	1μm GaAs FET	1992
	2.2	19.6	6 / -3	10	1.0	L-Degen.	1μm GaAs FET	
Nakatsugawa *et al.* [11]	2.0	12.2	5.1 / na	2	1.9	L-Degen.	0.3μm GaAs FET	1993
Heaney *et al.* [12]	1.5	14.5	11.2 / -1.1	12	1.9	L-Degen.	1μm GaAs FET	1993
Imai *et al.* [13]	2.5	11.5	9 / na	14	1.6	L-Degen.	0.3μm GaAs FET	1994
Sheng *et al.* [14]	5.7	7.8	23.9 / 11	115	1.0	Shunt-Ser. FB	GaAs HBT	1991
Meyer *et al.* [15]	2.2	16	6 / -4	40	0.9	L-Degen.	QUBiC BiCMOS	1994
Kobayashi *et al.* [16]	2.9	17.5	na / na	480	1.0	$1/g_m$-Term. & Shunt-Ser. FB	GaAs HBT	1994

[a]IP3 / -1dB compression point are output-referred.
[b]Neglects contribution of termination resistors. See text for discussion.

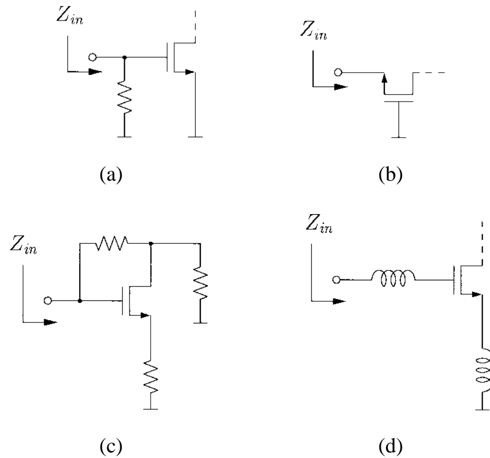

Fig. 1. Common LNA architectures. (a) Resistive termination, (b) $1/g_m$ termination, (c) shunt-series feedback, and (d) inductive degeneration.

The first technique uses *resistive termination* of the input port to provide a 50 Ω impedance. This approach is used in its differential form by Chang *et al.* [8], for example. Unfortunately, the use of real resistors in this fashion has a deleterious effect on the amplifier's noise figure. The noise contribution of the terminating resistors is neglected in that work because an antenna would be mounted directly on the amplifier, obviating the need for input matching. Hence, the reported noise figure of 6 dB corresponds to a hypothetical "terminationless" amplifier.

In general, however, the LNA is driven by a source that is located some distance away, and one must account for the influence of the terminating resistor. Specifically, we require that the amplifier possess a reasonably stable input impedance of approximately 50 Ω. To evaluate the efficacy of simple resistive input termination, suppose that a given LNA employing resistive termination has an available power gain of G_a and an available noise power at the output $P_{na,i}$ due to internal noise sources only; $P_{na,i}$ is, to first order, independent of the source impedance. Then, the noise factor is found to be[1]

$$F \triangleq \frac{\text{Total output noise}}{\text{Total output noise due to the source}} = 1 + \frac{P_{na,i} + kTBG_a}{kTBG_a} = 2 + \frac{P_{na,i}}{kTBG_a} \tag{1}$$

where B is the bandwidth over which the noise is measured. When the amplifier termination is removed, the noise figure expression becomes approximately

$$F = 1 + \frac{P_{na,i}}{4kTBG_a} \tag{2}$$

where we have assumed a high input impedance relative to the source. From (1) and (2), we may surmise that a "terminationless" amplifier with a 6 dB noise figure would likely possess an 11.5 dB noise figure with the addition of the terminating resistor. Two effects are responsible for this sharp degradation in noise figure. First, the added resistor contributes its own noise to the output which equals the contribution of the source resistance. This results in a factor of two difference in the first terms of (1) and (2). Second, the input is attenuated, leading to the factor of four difference in the second terms of (1) and (2). The large noise penalty resulting from these effects therefore makes this architecture unattractive for the more general situation where a good input termination is desired.

A second architectural approach, shown in Fig. 1(b), uses the source or emitter of a common-gate or common-base stage as the input termination. A simplified analysis of the $1/g_m$-termination architecture,

assuming matched conditions, yields the following lower bounds on noise factor for the cases of bipolar and CMOS amplifiers:

$$\text{Bipolar:} \quad F = \frac{3}{2} = 1.76 \text{ dB}$$

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

[1]Evaluated at $T = 290$ K.

where

$$\alpha \triangleq \frac{g_m}{g_{d0}}. \tag{3}$$

In the CMOS expressions, γ is the coefficient of channel thermal noise, g_m is the device transconductance, and g_{d0} is the zero-bias drain conductance. For long-channel devices, $\gamma = 2/3$ and $\alpha = 1$. The bipolar expression neglects the effect of base resistance in bipolar devices, while the value of 2.2 dB in the CMOS expression neglects both short-channel effects ($\alpha \leq 1$) and excess thermal noise due to hot electrons ($\gamma \geq 2/3$). Indeed, for short-channel MOS devices, γ can be much greater than one, and α can be much less than one. Accordingly, the minimum theoretically achievable noise figures tend to be around 3 dB or greater in practice.

Fig. 1(c) illustrates yet another topology, which uses resistive shunt and series feedback to set the input and output impedances of the LNA. This approach is taken in [9] and [14] and as the second stage in [16]. It is evident from Table I that amplifiers using shunt-series feedback often have extraordinarily high power dissipation compared to others with similar noise performance. Intuitively, the higher power is partially due to the fact that shunt-series amplifiers of this type are naturally broadband, and hence techniques which reduce the power consumption through *LC* tuning are not applicable. For GPS applications, a broadband front end is not required, and it is desirable to make use of narrowband techniques to reduce power. In addition, the shunt-series architecture requires on-chip resistors of reasonable quality, which are generally not available in CMOS technologies. For these reasons, the shunt-series feedback approach is not pursued in this work.

The fourth architecture, and the one that we have used in this design, employs inductive source or emitter degeneration as shown in Fig. 1(d) to generate a real term in the input impedance. Tuning of the amplifier input becomes necessary, making this a narrow-band approach. However, this requirement is not a limitation for a GPS receiver.

Note that inductive source degeneration is the most prevalent method used for GaAs MESFET amplifiers. It has also been used in CMOS amplifiers recently at 900 MHz [5]. As we will see, the proliferation of this architecture is no accident; it offers the possibility of achieving the best noise performance of *any* architecture.

III. LNA Architectural Analysis

We will now pursue a careful analysis of the architecture in Fig. 1(d) to establish clearly the principle of operation and the limits on noise performance. A brief review of the standard CMOS noise model will facilitate the analysis.

A. Standard MOS Noise Model

The standard CMOS noise model is shown in Fig. 2. The dominant noise source in CMOS devices is channel thermal noise. This source of noise is commonly modeled as a shunt current source in the output circuit of the device. The channel noise is white with a power spectral density given by

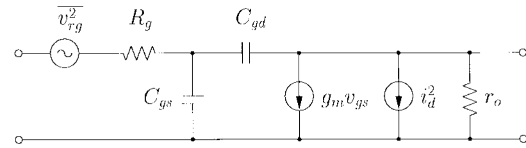

Fig. 2. The standard CMOS noise model.

$$\frac{\overline{i_d^2}}{\Delta f} = 4kT\gamma g_{d0} \tag{4}$$

where g_{d0} is the zero-bias drain conductance of the device, and γ is a bias-dependent factor that, for long-channel devices, satisfies the inequality

$$\frac{2}{3} \leq \gamma \leq 1. \tag{5}$$

The value of 2/3 holds when the device is saturated, and the value of one is valid when the drain-source voltage is zero. For short-channel devices, however, γ does not satisfy (5). In fact, γ *is much greater than 2/3 for short-channel devices operating in saturation* [1]–[4]. For 0.7-μm channel lengths, γ may be as high as two to three, depending on bias conditions [1].

This excess noise may be attributed to the presence of hot electrons in the channel. The high electric fields in submicron MOS devices cause the electron temperature, T_e, to exceed the lattice temperature. The excess noise due to carrier heating was anticipated by van der Ziel as early as 1970 [17].

An additional source of noise in MOS devices is the noise generated by the distributed gate resistance [18]. This noise source can be modeled by a series resistance in the gate circuit and an accompanying white noise generator. By interdigitating the device, the contribution of this source of noise can be reduced to insignificant levels. For noise purposes, the distributed gate resistance is given by [19]

$$R_g = \frac{R_\square W}{3n^2 L} \tag{6}$$

where $R_\square$ is the sheet resistance of the polysilicon, W is the total gate width of the device, L is the gate length, and n is the number of gate fingers used to lay out the device. The factor of 1/3 arises from a distributed analysis of the gate, assuming that each gate finger is contacted only at one end. By contacting at *both* ends, this term reduces to 1/12. In addition, this expression neglects the interconnect resistance used to connect the multiple gate fingers together. The interconnect can be routed in a metal layer that possesses significantly lower sheet resistance, and hence is easily rendered insignificant.

Though playing a role similar to that of base resistance in bipolar devices, the gate resistance is much less significant in CMOS because it can be minimized through interdigitation without the need for increased power consumption, unlike its bipolar counterpart. Its significance is further reduced in silicided CMOS processes which possess a greatly reduced sheet resistance, $R_\square$.

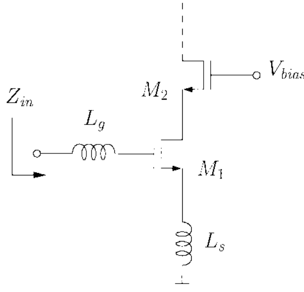

Fig. 3. Common-source input stage.

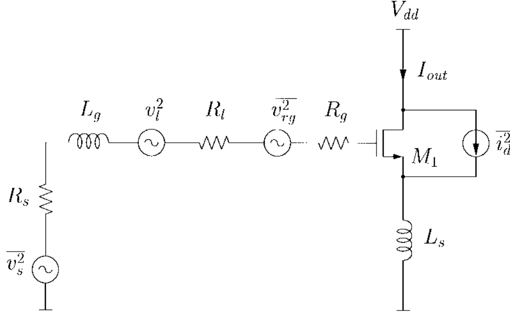

Fig. 4. Equivalent circuit for input stage noise calculations.

B. LNA Architecture

Having established the form of the CMOS noise model that we will use, we proceed to the analysis of the LNA architecture. Fig. 3 illustrates the input stage of the LNA. A simple analysis of the input impedance shows that

$$\begin{aligned} Z_{in} &= s(L_s + L_g) + \frac{1}{sC_{gs}} + \left(\frac{g_{m1}}{C_{gs}}\right)L_s \\ &\approx \omega_T L_s \qquad \text{(at resonance).} \end{aligned} \tag{7}$$

At the series resonance of the input circuit, the impedance is purely real and proportional to L_s. By choosing L_s appropriately, this real term can be made equal to 50 Ω. For example, if f_T is 10 GHz, a 50 Ω impedance requires only 800 pH for L_s. This small amount of inductance can easily be obtained with a single bondwire or on-chip spiral inductor. The gate inductance L_g is used to set the resonance frequency once L_s is chosen to satisfy the criterion of a 50-Ω input impedance.

The noise figure of the LNA can be computed by analyzing the circuit shown in Fig. 4. In this circuit, R_l represents the series resistance of the inductor L_g, R_g is the gate resistance of the NMOS device, and $\overline{i_d^2}$ represents the channel thermal noise of the device. Analysis based on this circuit neglects the contribution of subsequent stages to the amplifier noise figure. This simplification is justifiable provided that the first stage possesses sufficient gain and permits us to examine in detail the salient features of this architecture. Note that the overlap capacitance C_{gd} has also been neglected in the interest of simplicity. The use of a cascoded first stage helps to ensure that this approximation will not introduce serious errors.

Recall that the noise factor for an amplifier is defined as[2]

$$F \triangleq \frac{\text{Total output noise}}{\text{Total output noise due to the source}}. \tag{8}$$

To evaluate the output noise when the amplifier is driven by a 50-Ω source, we first evaluate the transconductance of the input stage. With the output current proportional to the voltage on C_{gs}, and noting that the input circuit takes the form of a series-resonant network

$$\begin{aligned} G_m = g_{m1}Q_{in} &= \frac{g_{m1}}{\omega_0 C_{gs}(R_s + \omega_T L_s)} \\ &= \frac{\omega_T}{\omega_0 R_s\left(1 + \frac{\omega_T L_s}{R_s}\right)} = \frac{\omega_T}{2\omega_0 R_s} \end{aligned} \tag{9}$$

where Q_{in} is the effective Q of the amplifier input circuit. In this expression, which is valid at the series resonance ω_0, R_l and R_g have been neglected relative to the source resistance, R_s. Perhaps surprisingly, the transconductance of this circuit at resonance is *independent* of g_{m1} (the device transconductance) as long as the resonant frequency is maintained constant. If the width of the device is adjusted, the transconductance of the *stage* will remain the same as long as L_g is adjusted to maintain a fixed resonant frequency. This result is intuitively satisfying, for as the gate width (and thus g_{m1}) is reduced, C_{gs} is also reduced, resulting in an increased Q_{in} such that the product of g_{m1} and Q_{in} remains fixed.

Using (9), the *output noise power density due to the 50-Ω source* is

$$S_{a,src}(\omega_0) = S_{src}(\omega_0)G_{m,eff}^2 = \frac{4kT\omega_T^2}{\omega_0^2 R_s\left(1 + \frac{\omega_T L_s}{R_s}\right)^2}. \tag{10}$$

In a similar fashion, the *output noise power density due to* R_l *and* R_g can be expressed as

$$S_{a,R_l,Rg}(\omega_0) = \frac{4kT(R_l + R_g)\omega_T^2}{\omega_0^2 R_s^2\left(1 + \frac{\omega_T L_s}{R_s}\right)^2}. \tag{11}$$

Equations (10) and (11) are also valid only at the series resonance of the circuit.

The dominant noise contributor internal to the LNA is the channel current noise of the first MOS device. Recalling the expression for the power spectral density of this source from (4), one can derive that the output noise power density arising from this source is

$$S_{a,i_d}(\omega_0) = \frac{\frac{\overline{i_d^2}}{\Delta f}}{\left(1 + \frac{\omega_T L_s}{R_s}\right)^2} = \frac{4kT\gamma g_{d0}}{\left(1 + \frac{\omega_T L_s}{R_s}\right)^2}. \tag{12}$$

The *total output noise power density* is the sum of (10)–(12). Assuming a 1 Hz bandwidth and substituting these into (8) yields

$$F = 1 + \frac{R_l}{R_s} + \frac{R_g}{R_s} + \gamma g_{d0} R_s\left(\frac{\omega_0}{\omega_T}\right)^2 \tag{13}$$

which is the noise factor of the LNA.

[2] Evaluated at $T = 290$ K.

This equation for noise factor reveals several important features of this LNA architecture. Note that the dominant term in (13) is the last term, which arises from channel thermal noise. Surprisingly, this term is *proportional* to g_{d0}. So, according to this expression, by reducing g_{d0} without modifying ω_T, we can simultaneously improve noise figure *and* reduce power dissipation. We can achieve this result by scaling the width of the device while maintaining constant bias voltages on its terminals and leaving the channel length unchanged. This scaling is consistent with the condition of constant ω_T, which depends only on the bias *voltages* on the device.

Recall, however, that this expression assumes that the amplifier is operated at the series resonance of its input circuit. So, a reduction in g_{d0} (and, hence in C_{gs}) must be compensated by an increase in L_g to maintain a constant resonant frequency. So, better noise performance and reduced power dissipation can be obtained by increasing the Q of the input circuit resonance.

By applying device scaling in this fashion to improve noise performance, the linearity of the amplifier will tend to degrade due to increased signal levels across C_{gs}. However, short-channel MOS devices operating in velocity saturation have a relatively constant transconductance with sufficient gate overdrive. This property is one advantage of implementing LNA's with MOS devices.

A second important feature in (13) is the inverse dependence on ω_T^2. Continued improvements in technology will therefore naturally lead to improved noise performance at a given frequency of operation.

Careful examination of (13) reveals a curious feature, however. Although finite inductor Q's will limit the amount of improvement practically available through device scaling, (13) does not predict a *fundamental* minimum for F. The implication is that a 0 dB noise figure may be achieved with zero power dissipation, and this simply cannot be true. Yet, the expression follows directly from the MOS noise model that we have assumed.

The conclusion can only be that our noise model is incomplete.

C. Extended MOS Noise Model

To understand the fundamental limits on noise performance of this architecture, we must turn our attention to induced gate current noise in MOS devices. Although absent from most (if not all) texts on CMOS circuit design, gate noise is given detailed treatment by van der Ziel [20].

Fig. 5 shows the cross section of a MOS device. If the device is biased so that the channel is inverted, fluctuations in the channel charge will induce a physical current in the gate due to capacitive coupling. This noise current can be (and has been) measured [21], but it is not included in the simple MOS noise model that we have used in the previous section.

A companion effect that occurs at very high frequencies arises due to the "distributed" nature of the MOS device. At frequencies approaching ω_T, the gate impedance of the device exhibits a significant phase shift from its purely capacitive value at lower frequencies. This shift can be accounted for by

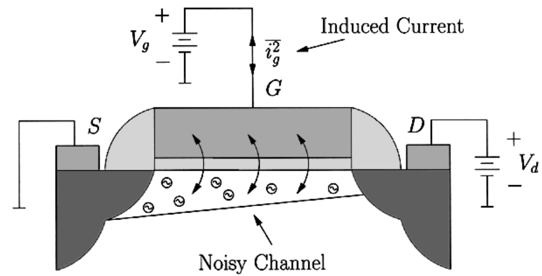

Fig. 5. Induced gate effects in MOS devices.

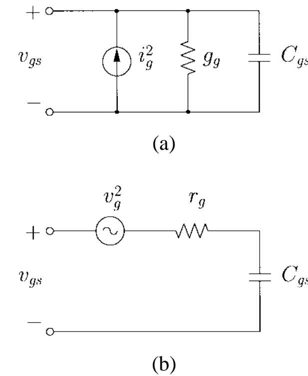

Fig. 6. Revised gate circuit model including induced effects. (a) Standard representation, as found in [20]. (b) The equivalent, but more intuitive, Thévenin representation.

including a real, noiseless conductance, g_g, in the gate circuit. Note that this conductance is distinct from the polysilicon resistance and is also distinct from the real term that occurs due to interaction of C_{gd} with g_m.[3]

A simple gate circuit model that includes both of these effects is shown in Fig. 6(a). A shunt noise current $\overline{i_g^2}$ and a shunt conductance g_g have been added. Mathematical expressions for these sources are [20][4]

$$\frac{\overline{i_g^2}}{\Delta f} = 4kT\delta g_g \tag{14}$$

$$g_g = \frac{\omega^2 C_{gs}^2}{5 g_{d0}} \tag{15}$$

where δ is the coefficient of gate noise, classically equal to 4/3 for long-channel devices. Equations (14) and (15) are valid when the device is operated in saturation.

Some observations on (14) and (15) are warranted. Note that the expression for the gate noise power spectral density takes a form similar to that of (4), which describes the drain noise power spectral density. However, in the gate noise expression, g_g is proportional to ω^2, and hence the gate noise is *not* a white noise source. Indeed, it is better described as a "blue" noise source due to its monotonically increasing power spectral density. It seems mysterious that the gate and drain noise terms have different types of power spectra, given

[3] A real conductance with a form similar to g_g is generated in cascoded amplifiers due to the feedback provided by C_{gd}. This effect is also significant at frequencies approaching ω_T.

[4] Our notation differs slightly from that found in [20], in which β is used in place of δ. The use of δ avoids confusion in cases where β represents $\mu_n C_{ox} W/L$, as is the practice in some texts on MOS devices.

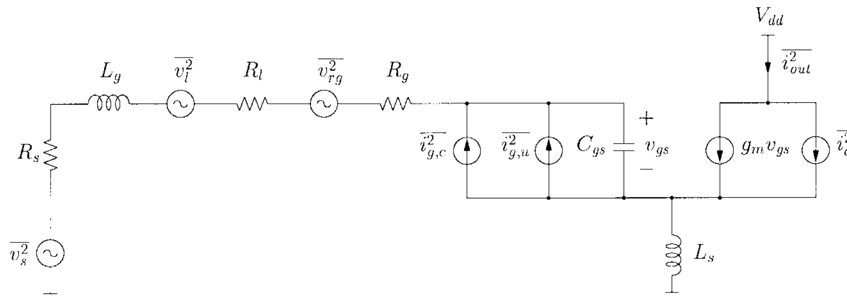

Fig. 7. Revised small-signal model for LNA noise calculations.

their common progenitor. The mystery is somewhat artificial, however, because the circuit of Fig. 6(a) can be cast into an equivalent, Thévenin representation as shown in Fig. 6(b) where

$$\frac{\overline{v_g^2}}{\Delta f} = 4kT\delta r_g \tag{16}$$

$$r_g = \frac{1}{5g_{d0}}. \tag{17}$$

We observe that v_g is now a *white* noise source proportional to a constant resistive term, r_g. This formulation of the gate circuit seems more intuitively appealing because the frequency dependence has been removed for *both* terms. Figs. 6(a) and (b) are interchangeable for frequencies where the Q of C_{gs} is sufficiently large, i.e.,

$$Q_{C_{gs}} = \frac{5g_{d0}}{\omega C_{gs}} \gg 1 \tag{18}$$

or, equivalently,

$$\omega \ll \frac{5g_{d0}}{C_{gs}} = \frac{5\omega_T}{\alpha} \tag{19}$$

where α was defined in (3) and is always less than one. This condition is automatically satisfied in all cases of practical interest.

In addition, we can expect the coefficient of gate noise δ to exhibit a dependence on electric field just as its counterpart, γ. To our knowledge, there are no published studies of the high-field behavior of δ.

The presence of gate noise complicates the analysis of F significantly. The gate noise is *partially correlated* with the drain noise, with a correlation coefficient given by [20]

$$c = \frac{\overline{i_g i_d^*}}{\sqrt{\overline{i_g^2}\,\overline{i_d^2}}} \approx 0.395j \tag{20}$$

where the value of $0.395j$ is exact for long-channel devices. The correlation can be treated by expressing the gate noise as the sum of two components, the first of which is fully correlated with the drain noise, and the second of which is uncorrelated with the drain noise. Hence, the gate noise is re-expressed as

$$\frac{\overline{i_g^2}}{\Delta f} = \underbrace{4kT\delta g_g(1-|c|^2)}_{\text{Uncorrelated}} + \underbrace{4kT\delta g_g|c|^2}_{\text{Correlated}}. \tag{21}$$

Because of the correlation, special attention must be paid to the reference polarity of the correlated component. The value of c is positive for the polarity shown in Fig. 6(a).

Having established this additional source of noise in MOS devices, we are now in a position to reevaluate the noise figure of the LNA. As we will see, the presence of gate noise establishes a lower bound on the achievable noise performance of the amplifier.

D. Extended LNA Noise Analysis

To evaluate the noise performance of the LNA in the presence of gate noise effects, we will employ the circuit of Fig. 7. In this circuit, we have neglected the effect of g_g under the assumption that the gate impedance is largely capacitive at the frequency of interest. Equation (19) specifies the condition under which this approximation holds. The gate noise has been subdivided into two parts. The first, $\overline{i_{g,c}^2}$, represents the portion of the total gate noise that is correlated with the drain noise. The second, $\overline{i_{g,u}^2}$, represents the portion that is uncorrelated with the drain noise.

With the revised small-signal model in mind, we can derive the noise factor of the LNA. A close examination of Fig. 7 allows us to anticipate the result of our analysis. As the Q of the input circuit is increased from zero, the noise figure will tend to improve in accordance with the earlier expression for F. However, the impedance at the gate of the device increases simultaneously, and hence the gate current noise will begin to dominate at some point. A minimum noise figure will thus be achieved for a particular input Q.

To analyze the circuit mathematically, we can draw on (10)–(12) from the previous section for the drain noise and resistive losses. However, the *amplitudes* of the correlated portion of the gate noise and the drain noise must be summed together before the *powers* of the various contributors are summed. Doing so yields a term representing the combined effect of the drain noise and the correlated portion of the gate noise

$$S_{a,i_d,i_{g,c}}(\omega_0) = \kappa S_{a,i_d}(\omega_0) = \frac{4kT\gamma\kappa g_{d0}}{\left(1+\frac{\omega_T L_s}{R_s}\right)^2} \tag{22}$$

where

$$\kappa = \frac{\delta\alpha^2}{5\gamma}|c|^2 + \left[1 + |c|Q_L\sqrt{\frac{\delta\alpha^2}{5\gamma}}\right]^2 \tag{23}$$

$$Q_L = \frac{\omega_0(L_s+L_g)}{R_s} = \frac{1}{\omega_0 R_s C_{gs}}. \tag{24}$$

Note that if $\delta \to 0$, then $\kappa \to 1$ and (22) then reduces to (12).

The last noise term is the contribution of the uncorrelated portion of the gate noise. This contributor has the following power spectral density:

$$S_{a,i_{g,u}}(\omega_0) = \xi S_{a,i_d}(\omega_0) = \frac{4kT\gamma\xi g_{d0}}{\left(1+\frac{\omega_T L_s}{R_s}\right)^2} \tag{25}$$

where

$$\xi = \frac{\delta\alpha^2}{5\gamma}(1-|c|^2)(1+Q_L^2). \tag{26}$$

We observe that all of the noise terms contributed by the first device M_1 are proportional to $S_{a,i_d}(\omega_0)$, the contribution of the drain noise. Hence, it is convenient to define the contribution of M_1 as a whole as

$$S_{a,M_1}(\omega_0) = \chi S_{a,i_d}(\omega_0) = \frac{4kT\gamma\chi g_{d0}}{\left(1+\frac{\omega_T L_s}{R_s}\right)^2} \tag{27}$$

where, after some slight simplification

$$\chi = \kappa + \xi = 1 + 2|c|Q_L\sqrt{\frac{\delta\alpha^2}{5\gamma}} + \frac{\delta\alpha^2}{5\gamma}(1+Q_L^2). \tag{28}$$

With (27) and (28), it is clear that the effect of induced gate noise is to modify the noise contribution of the device in proportion to χ. It follows directly that

$$F = 1 + \frac{R_l}{R_s} + \frac{R_g}{R_s} + \gamma\chi g_{d0} R_s\left(\frac{\omega_0}{\omega_T}\right)^2 \tag{29}$$

where χ is defined as in (28). By factoring out Q_L from the expression for χ, and noting that

$$g_{d0}Q_L = \frac{g_m}{\alpha}\frac{1}{\omega_0 R_s C_{gs}} = \frac{\omega_T}{\alpha\omega_0 R_s} \tag{30}$$

we can re-express F as

$$F = 1 + \frac{R_l}{R_s} + \frac{R_g}{R_s} + \frac{\gamma}{\alpha}\frac{\chi}{Q_L}\left(\frac{\omega_0}{\omega_T}\right). \tag{31}$$

To understand the implications of this new expression for F, we observe that χ includes terms which are constant, proportional to Q_L, and proportional to Q_L^2. It follows that (31) will contain terms which are *proportional* to Q_L as well as *inversely proportional* to Q_L. Therefore, a minimum F exists for a particular Q_L, as argued earlier.

IV. LNA Design Considerations

The analysis of the previous section can now be drawn upon in designing the LNA. Of primary interest is insight into picking the appropriate device width and bias point to optimize noise performance given specific objectives for gain and power dissipation.

To select the width of M_1, we turn to (28) and (31). Note that all of the terms are well defined in these expressions, except for γ and δ. Because γ and δ both depend on drain bias in an unspecified fashion, it is difficult to account properly for their contributions. To surmount this difficulty, we adopt the assumption that although each may be a function of bias, the *ratio* can be expected to show less variation because γ and δ will likely have similar dependence on bias, given their common progenitor. The reader is cautioned, however, that this assumption is somewhat arbitrary; it is necessary because the detailed high-field behavior of γ and δ is presently unknown. Modifications may be required once further research yields information about these coefficients. It is our hope that, having made this assumption, the analysis which follows will be easily adapted to account for the high-field natures of γ and δ.

In preparation for optimizing the noise performance of the LNA, it will be useful to formulate the quantities α, ω_T, and Q_L in terms of the gate overdrive voltage of M_1.

A. Definition of Terms

To quantify these terms, a simple second-order model of the MOSFET transconductance can be employed which accounts for high-field effects in short-channel devices. Assume that I_d has the form [22]

$$I_d = WC_{ox}\nu_{sat}\frac{V_{od}^2}{V_{od}+L\varepsilon_{sat}} \tag{32}$$

with

$$V_{od} = V_{gs} - V_T \tag{33}$$

where C_{ox} is the gate oxide capacitance per unit area, ν_{sat} is the saturation velocity, and ε_{sat} is the velocity saturation field strength. We can differentiate this expression to determine the transconductance, yielding

$$g_m = \frac{\partial I_d}{\partial V_{gs}} = \mu_{eff}C_{ox}\frac{W}{L}V_{od}\underbrace{\left[\frac{1+\rho/2}{(1+\rho)^2}\right]}_{\alpha} \tag{34}$$

with the definition that

$$\rho = \frac{V_{od}}{L\varepsilon_{sat}} \tag{35}$$

where μ_{eff} is the field-limited electron mobility. The term in square braces is α itself.

Having established an expression for I_d, we can formulate the power consumption of the amplifier as follows:

$$P_D = V_{dd}I_d = V_{dd}WC_{ox}\nu_{sat}\frac{V_{od}^2}{V_{od}+L\varepsilon_{sat}}. \tag{36}$$

Note that the power dissipation is proportional to the device width, W. Another quantity which depends directly on W is Q_L, which has been specified in (24). Combining this equation with (36), and noting that $C_{gs} = \frac{2}{3}WLC_{ox}$, we can relate Q_L to P_D with

$$Q_L = \frac{P_0}{P_D}\frac{\rho^2}{1+\rho} \tag{37}$$

where

$$P_0 = \frac{3}{2}\frac{V_{dd}\nu_{sat}\varepsilon_{sat}}{\omega_0 R_s}. \tag{38}$$

Note that for the purposes of our analysis, P_0 is a constant determined solely by physical technological parameters (ν_{sat} and ε_{sat}) and design target specifications (V_{dd}, ω_0, and R_s).

Another factor required in the design process is ω_T. This can also be evaluated with the help of (34) to be

$$\omega_T \approx \frac{g_m}{C_{gs}} = \frac{g_m}{\frac{2}{3}WLC_{ox}} = \frac{3}{2}\frac{\alpha\mu_{eff}V_{od}}{L^2} = \frac{3\alpha\rho\nu_{sat}}{L}. \quad (39)$$

This expression is approximate because we are neglecting C_{gd}, the gate-drain overlap capacitance. Note that proportionality to α limits the ω_T that can be achieved with a given device.

B. Noise Figure Optimization Techniques

With the relevant quantities now defined, we can proceed to optimize the noise performance of the amplifier. In low noise amplifier design, determination of the minimum noise figure is a common and well-understood procedure. Typically, a small-signal model of the amplifier is assumed *a priori*, an expression for F is formed, and differentiation leads to the unique conditions for optimized noise performance. The reader is referred to [23] for an excellent treatment of the general approach. There is a significant distinction, however, between that type of optimization and the one which we seek to perform here. In this analysis, we seek the conditions that guarantee optimized noise performance for a specified fixed design parameter, such as gain or power consumption, under the condition of perfect input matching. Accordingly, we fix the necessary design criteria and determine the appropriate small-signal model *a posteriori* through the optimization procedure. Because the architecture permits selection of Q_L and L_s independently, a solution exists for which the optimum noise performance coincides with the best input match.

There are two approaches to this optimization problem which deserve special attention. The first assumes a fixed transconductance, G_m, for the amplifier. The second assumes a fixed power consumption. To illustrate the second approach, the expression for F in (31) can be recast to make its dependence on power dissipation (P_D) explicit. It is, however, nontrivial to make the dependence on G_m explicit. Fortunately, the condition for constant G_m is equivalent to the condition of constant ω_T, as is clear from (9). To maintain a fixed ω_T, we need only fix the value of ρ. Hence, we will reformulate F in terms of P_D and ρ to facilitate both optimizations.

We can draw on (34), (37), and (39) and substitute into (31) expressions for α, Q_L, and ω_T in terms of the relative gate overdrive, ρ. The result is that

$$F = 1 + \frac{\gamma\omega_0 L}{3\nu_{sat}}P(\rho, P_D) \quad (40)$$

in which we have neglected the contributions of the gate resistance and inductor losses to the noise factor. In this new expression, $P(\rho, P_D)$ is a ratio of two sixth-order polynomials of ρ given by

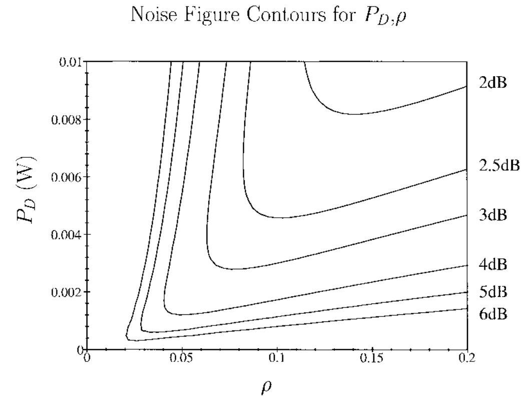

Fig. 8. Contours of constant noise figure relating ρ and P_D, for $L = 0.35\,\mu$m, $R_s = 50\Omega$, $\omega_0 = 10$ Grps, $V_{dd} = 1.5$ V, $\gamma = 2.5$ [1], $\delta = 5.0$, $|c| = 0.395$ [20], $\nu_{sat} = 1 \times 10^5$ m/s, and $\varepsilon_{sat} = 4.7 \times 10^6$ V/m [22].

$$P(\rho, P_D) = \frac{\frac{P_D}{P_0}P_1(\rho) + P_2(\rho) + \frac{P_0}{P_D}P_3(\rho)}{\rho^3\left(1+\frac{\rho}{2}\right)^2(1+\rho)} \quad (41)$$

with

$$P_1(\rho) = (1+\rho)^6 + \frac{\delta}{5\gamma}(1+\rho)^2\left(1+\frac{\rho}{2}\right)^2$$

$$P_2(\rho) = 2|c|\sqrt{\frac{\delta}{5\gamma}}(1+\rho)^3\left(1+\frac{\rho}{2}\right)\rho^2$$

$$P_3(\rho) = \frac{\delta}{5\gamma}\left(1+\frac{\rho}{2}\right)^2\rho^4.$$

Because F is a function of two variables, one can define contours of constant noise figure in ρ and P_D. An example is shown in Fig. 8. To generate this plot, we have adopted the assumption that the ratio of γ to δ is unchanged by hot-electron effects. These contours give a useful indication of the design tradeoffs between noise figure, power dissipation, and gate overdrive.

The form of (40) suggests that optimization of F proceeds by minimizing $P(\rho, P_D)$ with respect to one of its arguments, keeping the other one fixed. The complexity of this polynomial will force us to make some simplifying assumptions when optimizing for a fixed power dissipation. Fortunately, the optimization for a fixed G_m can proceed directly from (41) without further simplifications.

1) Fixed G_m Optimization: To fix the value of the transconductance, G_m, we need only assign a constant value to ρ. The appropriate value for ρ is easily determined by substituting (39) into the expression for G_m as found in (9). The result, which relates G_m to ρ, is

$$G_m = \frac{3\nu_{sat}}{2\omega_0 R_s L}\frac{\rho\left(1+\frac{\rho}{2}\right)}{(1+\rho)^2}. \quad (42)$$

Once ρ is determined, we can minimize the noise factor by

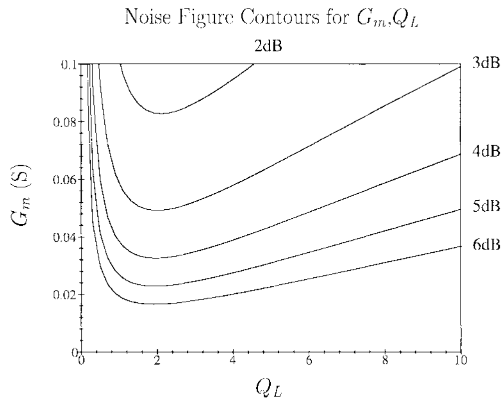

Fig. 9. Contours of constant noise figure relating Q_L and G_m. The same assumptions as in Fig. 8 apply.

taking

$$\frac{\partial P(\rho, P_D)}{\partial P_D} = 0 \tag{43}$$

which, after some algebraic manipulations, results in

$$P_{D,opt,G_m} = P_0\sqrt{\frac{P_3(\rho)}{P_1(\rho)}} = P_0\frac{\rho^2}{1+\rho}\left[1+\frac{5\gamma}{\delta\alpha^2}\right]^{-1/2}. \tag{44}$$

This expression gives the power dissipation which yields the best noise performance for a given G_m under the assumption of a matched input impedance. By comparing (44) to (37), we see immediately that this optimum occurs when

$$Q_L = Q_{L,opt,G_m} = \sqrt{1+\frac{5\gamma}{\delta\alpha^2}} \geq 1.87. \tag{45}$$

Hence, the best noise performance for a given transconductance is achieved at some specific input Q. Note that the value 1.87 is valid only for long-channel devices. For short-channel lengths, where $\alpha < 1$, we can expect the optimum Q_L to be somewhat larger. Note that if we substitute Q_{L,opt,G_m} into (28), the sum of the second two terms (which are attributed to the presence of gate noise) exceeds unity, thus indicating that the gate current contributes *more* noise than the drain current. The Q_L for which the contribution of the two sources is equal may be easily found from (28) to be less than Q_{L,opt,G_m}.

By substituting (45) into (31), we determine that the minimum noise factor (neglecting inductor and gate losses) is

$$F_{min,G_m} = 1+\sqrt{\frac{4}{5}\delta\gamma}\left(\frac{\omega_0}{\omega_T}\right)\left\{|c|+\sqrt{1+\frac{\delta\alpha^2}{5\gamma}}\right\} \tag{46}$$
$$\geq 1+1.33\left(\frac{\omega_0}{\omega_T}\right).$$

The value of 1.33 is only valid for long-channel devices; it may be three to four times larger in the presence of high electric fields.

The constant noise figure contours plotted in Fig. 9 illustrate the behavior of G_m for arbitrary Q_L.

2) Fixed P_D Optimization: An alternate method of optimization fixes the power dissipation and adjusts ρ to find the minimum noise factor. The expression for $P(\rho, P_D)$ is too complex in ρ to yield a closed form solution for the optimum point. However, we can adopt a simplifying assumption and check its validity by graphical comparison. If we assume that $\rho \ll 1$, then $P(\rho, P_D)$ can be simplified to

$$P(\rho, P_D) \approx \frac{\frac{P_D}{P_0}\left(1+\frac{\delta}{5\gamma}\right)+2|c|\sqrt{\frac{\delta}{5\gamma}}\rho^2+\frac{P_0}{P_D}\frac{\delta}{5\gamma}\rho^4}{\rho^3}. \tag{47}$$

This expression is minimized for a fixed P_D when

$$\frac{\partial P(\rho, P_D)}{\partial \rho} = 0. \tag{48}$$

The solution of this equation, under the assumption that $\rho \ll 1$ is

$$\rho^2_{opt,P_D} = \frac{P_D}{P_0}|c|\sqrt{\frac{5\gamma}{\delta}}\left[1+\sqrt{1+\frac{3}{|c|^2}\left(1+\frac{\delta}{5\gamma}\right)}\right]. \tag{49}$$

By comparing (49) to (37), it is clear that this value for ρ is equivalent to an optimum Q_L of

$$Q_{L,opt,P_D} = |c|\sqrt{\frac{5\gamma}{\delta}}\left[1+\sqrt{1+\frac{3}{|c|^2}\left(1+\frac{\delta}{5\gamma}\right)}\right] \approx 3.9. \tag{50}$$

So, it is clear that the optimum Q_L for a fixed power dissipation is *larger* than the optimum Q_L for a fixed G_m. We can now evaluate (28) and use the result in (31) to show that

$$F_{min,P_D} \approx 1+2.4\frac{\gamma}{\alpha}\left(\frac{\omega_0}{\omega_T}\right) \geq 1+1.62\left(\frac{\omega_0}{\omega_T}\right) \tag{51}$$

where the value of 1.62 is valid only in the long-channel limit; the value will be somewhat larger for short-channel devices in velocity saturation.

To examine the validity of our simplifying assumption that $\rho \ll 1$, the noise figure is plotted in Fig. 10 for the two cases defined in (41) and (47). Evidently, the approximation of $\rho \ll 1$ is reasonable near the optimum point, though the curves diverge somewhat as Q_L increases. Note that the simplified expression slightly underestimates the necessary Q_L for best performance. Nonetheless, it predicts F_{min,P_D} remarkably well.

Finally, Fig. 11 shows contours of constant noise figure relating P_D and Q_L. These contour plots are useful for selecting device geometries for a particular power dissipation and desired noise figure.

C. Discussion of F_{min}

Although we have derived expressions for F_{min} in this section under two different optimization procedures, the question arises as to whether the analysis has produced, indeed, the minimum F which can be achieved *for any architecture*.

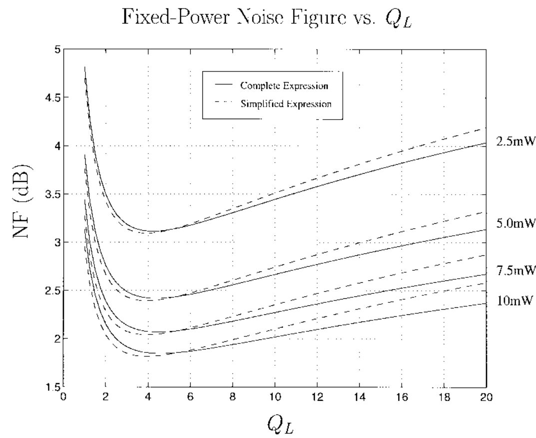

Fig. 10. Theoretical predictions of noise figure F for several power dissipations. The same assumptions as in Fig. 8 apply.

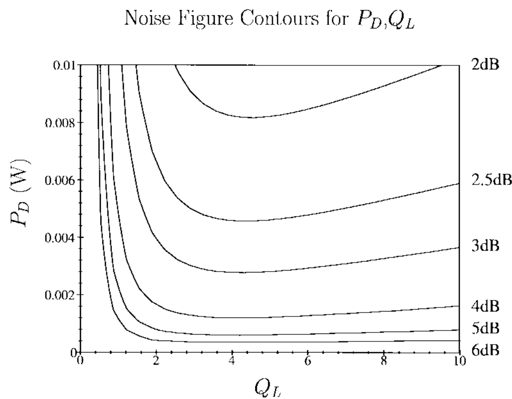

Fig. 11. Contours of constant noise figure relating Q_L and P_D. The same assumptions as in Fig. 8 apply.

The difficulty in answering this question is that our optimization procedures identify the best MOS device for a fixed R_s under particular design constraints (e.g., power consumption or gain). In contrast, traditional noise figure minimization techniques seek to determine the optimum Z_s for a *given* MOS device at a specified power level, and though this latter approach achieves the best performance for a particular device, the performance of the amplifier may be suboptimal for other figures-of-merit (such as input reflection coefficient). Indeed, the traditional techniques do not aid the selection of the appropriate device geometries at all. In this respect, the techniques that we have presented are more useful for integrated circuit design, where the device geometry is controlled by the designer.

But the question remains: how do these expressions for F_{min} relate to the minimum-F of the MOS device itself? As proven by Haus *et al.* [23], the minimum noise factor for a MOS device is achieved with a particular source conductance when the source susceptance cancels the noise correlation susceptance of the device. Such a condition is commonly referred to as a *conjugate noise match.* A MOS device with partially correlated gate noise has a correlation susceptance given by

$$B_{cor} = \omega C_{gs}\left[1 + |c|\sqrt{\frac{\delta\alpha^2}{5\gamma}}\right] \le 1.25\omega C_{gs}. \quad (52)$$

Hence, the optimum source susceptance is an inductance which resonates with the gate capacitance at a frequency slightly *higher* than ω_0. This is sufficient to specify the imaginary part of Y_s. A simple transformation can be used to put the source admittance into a series impedance form which is equivalent at a particular frequency. This transformation preserves the value of inductance for moderate values of Q, thus ensuring that the series resonance will occur at nearly the same frequency as its parallel counterpart. This series equivalent corresponds to the architecture of the LNA.

Because the analysis presented in this paper assumes a series resonance at the frequency of operation, we may conclude that it does not quite yield F_{min} for a particular device. However, the difference in the optimum series resonance frequency and ω_0 is only about 15%. So, we can expect the proposed architecture to possess near-optimum noise performance.

Observe that, in our analysis, the constraint that leads to optimum noise performance is in terms of an optimum Q_L. This optimum does not constrain the value of L_s itself, but rather the *sum* of L_s and L_g. This degree of freedom permits the optimum noise performance to be obtained while simultaneously permitting selection of L_s for a good input match. So, in return for a slight noise degradation, the quality of the input match is assured, which is a desirable design goal. Evidently, this architecture exhibits the well-known tradeoff between input reflection coefficient and noise figure. Simply put, this tradeoff exists because $B_{cor} \neq \omega C_{gs}$.

On the other hand, the optimum source *resistance* is that which balances the contributions of drain and gate noise generators. In the traditional approach, R_s is varied to locate the optimum. In our approach, the device characteristics are varied, which changes the relative powers of the two noise generators. The minimum is achieved for the optimum balance of these two generators, and hence the result is the same as the traditional analysis. Indeed, if we fix all of the terms in (31) and evaluate the optimum source resistance R_s, the optimum results in the same expression for Q_L as in the fixed-G_m analysis. That is, the optimum *device* determined by the fixed-G_m analysis of the previous section necessarily possesses the quality that the specified R_s is optimal for that particular device at that particular power level.

Given that the fixed-G_m analysis optimally matches the device to the source, one is tempted to reject the second optimization method (fixed-P_D) which results in a different "optimum" for Q_L. Clearly, this procedure does *not* match the device to R_s in the sense of Haus *et al.* However, as we will show, the second method is likely to be preferred in most cases.

Consider Fig. 12, which illustrates a thought experiment entailing several tradeoffs inherent in a constant-power optimization. In this figure, the solid arcs represent fixed-P_D optimizations, while the dashed arcs represent optimizations

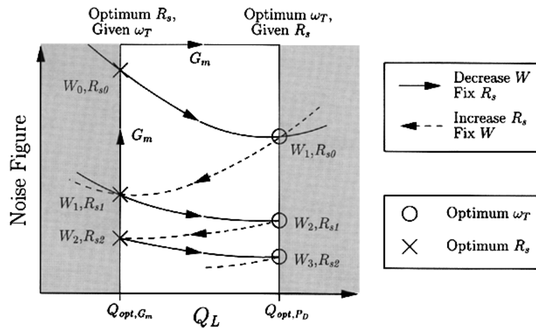

Fig. 12. Noise figure optimization experiment illustrating the significance of Q_{opt,G_m} and Q_{opt,P_d}. Note that the curves shown represent constant-P_D.

where R_s is modified to noise-match the given device. Suppose that we begin with a device which has been optimized using the fixed-G_m analysis for a particular R_{s0}, resulting in a device width W_0. Although R_s is optimally matched to this particular device, superior noise performance can be obtained on the same power dissipation by decreasing the device width to W_1, following the fixed-P_D arc. The noise performance improves in this procedure despite the nonoptimal source resistance because ω_T improves as the scaling is performed. This offsets the loss in noise match until Q_{opt,P_D} is reached. At this point, the gate noise dominates the output noise of the device. So, degrading the noise match in favor of the gate noise permits operation at an elevated ω_T; the net result is improved noise performance. Also note that the gain, G_m, actually *improves* in this procedure.

Of course, once the new width W_1 is determined, an increased source resistance can be found which is noise-matched to this new device. This procedure takes the design back along the dashed arc, yielding improved noise performance until Q_{opt,G_m} is reached. However, there is a significant penalty in G_m which is incurred by this increase in R_s (recall that G_m is inversely proportional to R_s). Nonetheless, this procedure could be repeated (at the expense of G_m) as long as it is reasonable to increase R_s and decrease W, maintaining Q_L to lie within the white region of Fig. 12.

The question is: at what point (and at which Q_L) should the ultimate design be placed? Assuming that a maximum realistic R_s can be specified, it seems reasonable always to design the LNA to operate at Q_{opt,P_D} because this design point will always possess a larger G_m than its lower-Q counterpart. The result is that the best LNA design operates at a Q_L which is *different* from the value corresponding to the conjugate noise match. A noise mismatch is tolerated in return for a higher ω_T at the same power dissipation.

We conclude that the optimization procedures given here, though not yielding F_{min} precisely as outlined in [23], permit selection of the best device for *two* constraints simultaneously: perfect input match and a specific gain; or perfect input match and a specific power dissipation. Of these, the second set of constraints yields the best combination of noise, power, and gain. There is only *one* device in a given technology that optimizes noise performance while satisfying either set of two of these specifications for a particular R_s.

Finally, it is clear that the minimum noise factor improves as ω_T increases with advances in technology. This fact, taken in conjunction with the experimental results of this study, signifies that CMOS LNA's will soon achieve noise performance at GPS frequencies that are largely parasitic-limited, making CMOS an attractive alternative to more costly silicon bipolar and GaAs technologies.

D. A Note on MOS Noise Simulation Models

The preceding analysis facilitates the design of CMOS low-noise amplifiers using this topology. It is important to note, however, that existing MOS noise models—as implemented in circuit simulators such as HSPICE—do not adequately account for hot-electron effects or induced gate effects. The options available for level 13, 28, and 39 MOS models (BSIM-I, Modified BSIM-I, and BSIM-II, respectively) do not account for even the most elementary of short-channel noise effects, much less the more advanced considerations of the previous section. This is particularly disturbing, given that the optimal LNA design will undoubtedly be limited by the gate noise of the device.

Some strides have been made recently with the adoption of the BSIM-III model. This model makes use of an alternative formulation for channel thermal noise in which the noise power is treated as proportional to the total inversion layer charge [24]. This is the same model proposed by Wang *et al.* [4]. Short-channel effects can be included in the formulation of the inversion layer charge, and hence in the noise power. However, even this model discounts the possibility that elevated carrier temperature is an important factor. The assumption of a uniform carrier temperature along the entire channel length may explain the departure of the model's predictions from measured data for relatively short-channel devices [25].

V. LNA Implementation

To probe further the ability of CMOS to deliver low noise amplification at 1.575 42 GHz, we have implemented an LNA in a 0.6-μm CMOS technology provided through the MOSIS service (0.35 μm L_{eff}). The only information about the technology available at the time of design referred to interlayer dielectric thicknesses, sheet resistances, and diffusion capacitances. Thankfully, the value of t_{ox} was also available, making possible a crude extrapolation from 0.8-μm models to provide some basis for simulation. The success of the implementation demonstrates that knowledge of device capacitances is the most important factor in the design of tuned amplifiers.

The width of the input device was initially chosen without regard to the induced gate noise term because the detailed nature of gate noise was unknown to the authors at design time. It will prove useful to know the optimum width for this technology so that we can determine whether our performance is limited by the induced gate noise or by the drain current noise. From Fig. 10, the optimum Q_L for a power dissipation of 7.5 mW (which corresponds to the measured P_D of the first stage of our LNA) is about 4.5, with a corresponding F_{min,P_D} of 2.1

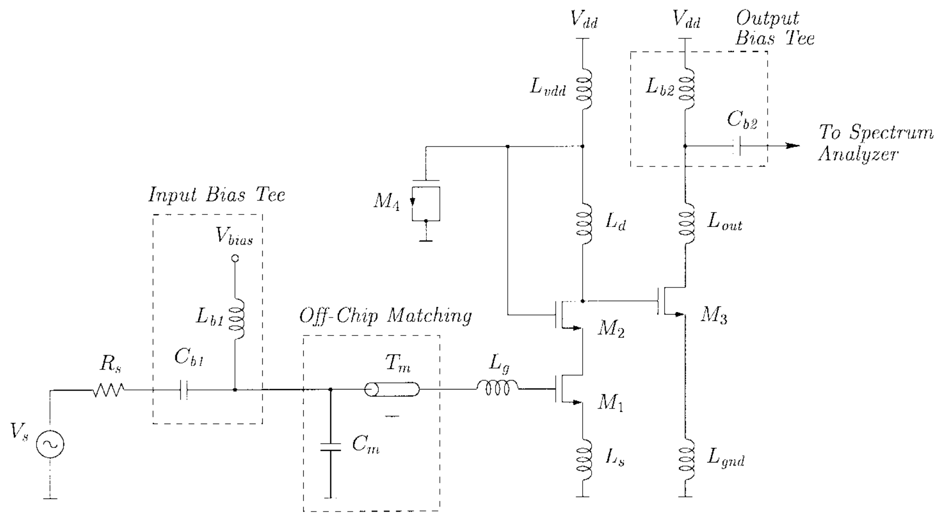

Fig. 13. Complete schematic of the LNA, including off-chip elements.

dB. We can immediately determine the optimum width to be

$$W_{M_1,opt,P_D} = \left[\frac{2}{3}\omega_0 L C_{ox} R_s Q_{L,opt,P_D}\right]^{-1} \approx 496\,\mu\text{m} \quad (53)$$

where $\omega_0 = 10$ Grps, $L = 0.35\mu$m, $C_{ox} = 3.84\,\text{mF/m}^2$, and $R_s = 50\Omega$. Because the product of L and C_{ox} remains roughly constant as device geometries scale, the optimum width as given by (53) is also relatively insensitive to scaling. The actual width of M_1, as implemented, is about 403 μm, which corresponds to a Q_L of 5.5, still very close to the minimum noise figure point for 7.5 mW of power dissipation. Because our Q_L is greater than the optimum, we expect that our measured performance will be limited by the gate noise. Note, however, that the predicted F neglects any contribution to the noise factor by parasitic losses, particularly those due to on-chip spiral inductors, which influence the noise figure of the LNA. Accordingly, the amplifier will possess a noise figure which is greater than 2.1 dB.

The complete schematic of the LNA is shown in Fig. 13. The amplifier is a two-stage, cascoded architecture. The drain of M_2 is tuned by a 7-nH on-chip spiral inductor, L_d. This inductor resonates with the total capacitance at the drain of M_2, including C_{gs} of M_3. Transistor M_3 serves as an open-drain output driver providing 4.6 dB of gain, and the amplifier uses the test instrument itself as the load. Note that M_3 has a gate width of about 200 μm, or half of M_1.

Four of the inductors shown (L_s, L_{gnd}, L_{vdd}, and L_{out}) are formed by bondwire inductances. Of these four, L_s is the only one whose specific value is significant in the operation of the amplifier, since it sets the input impedance of the LNA. L_{gnd} and L_{out} are unwanted parasitics, so their values are minimized by proper die bonding. L_{vdd} aids in supply filtering with M_4, which acts as a supply bypass capacitor. Because a large value of inductance is beneficial for this use, L_{vdd} is formed from a relatively long bondwire.

Due to the lack of simulation models before fabrication, a flexible topology was chosen which would permit postfabrication adjustment of the bias points of M_1 and M_2. The input matching is accomplished with the aid of an off-chip network. Off-chip tuning was required because the necessary value of L_g was prohibitively large for on-chip fabrication. However, a 4-nH inductor was integrated on-chip in series with the gate of M_1. This inductor, together with the input bondwire inductance, reduces the matching burden of the off-chip network. Unfortunately, it also introduces additional resistive losses which degrade the noise performance of the LNA.

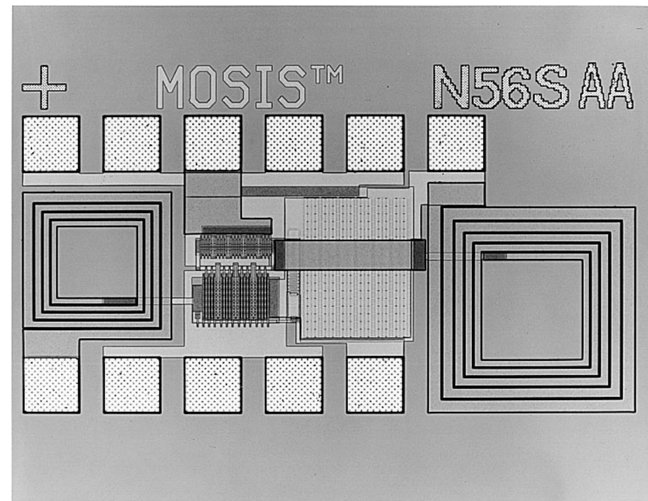

Fig. 14. Die photo of the LNA.

A die photo of the LNA is shown in Fig. 14. The two spiral inductors are clearly visible. The input pad is on the lower left corner of the die. The spiral on the left is a 4-nH inductor which forms a portion of L_g. The spiral on the right is a 7-nH inductor that tunes the output of the first stage. The spirals are fabricated in metal-three, which permits Q's of about three to be achieved. This value of Q is typical of on-chip spiral inductors that have been reported in the literature [26]. To improve the Q slightly, the inductors are tapered so that the outer spirals use wider metal lines than the inner

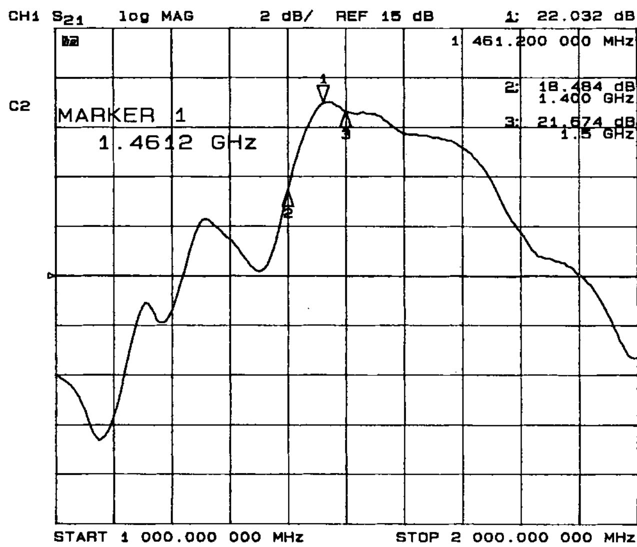

Fig. 15. Measured S21 of the LNA.

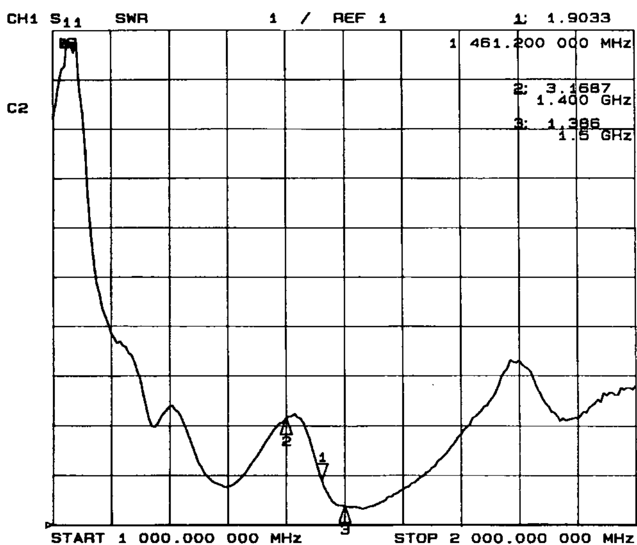

Fig. 16. Measured S11 of the LNA.

spirals. The goal of this tapering is to distribute the loss to yield a roughly constant loss per turn. A magnetic field solver, FastHenry, was used during the design of the LNA to predict the values of inductance and the winding loss associated with various geometries. From these simulations, we determined that tapering provides a slight, but welcome, increase in Q (approximately 20%). Several inner turns were also removed in a naive attempt to increase Q further.

VI. Experimental Results

To test the LNA, the die was mounted in a high-frequency package and bonded. The measured gain (S21) of the amplifier appears in Fig. 15. The gain has a peak value of 22 dB at 1.46 GHz and remains above 20 dB to almost 1.6 GHz. The bandpass nature of the amplifier is evident from the plot. The input reflection coefficient (S11) is also plotted in Fig. 16. The input VSWR at 1.5 GHz is quite good (about 1.4) with the addition of off-chip tuning elements.

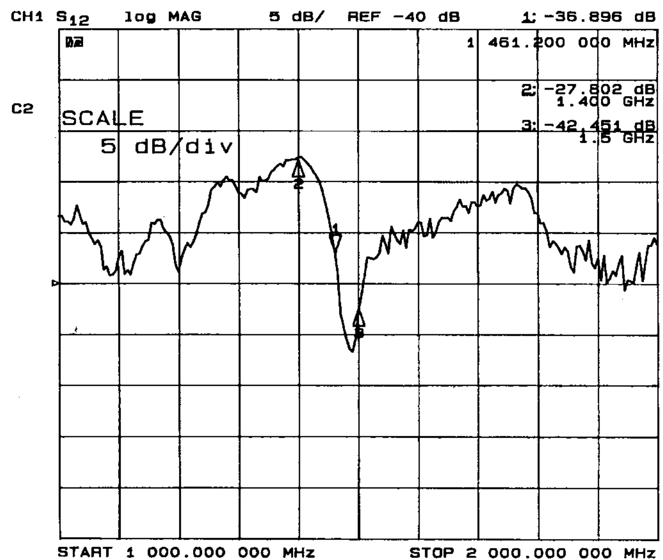

Fig. 17. Measured S12 of the LNA.

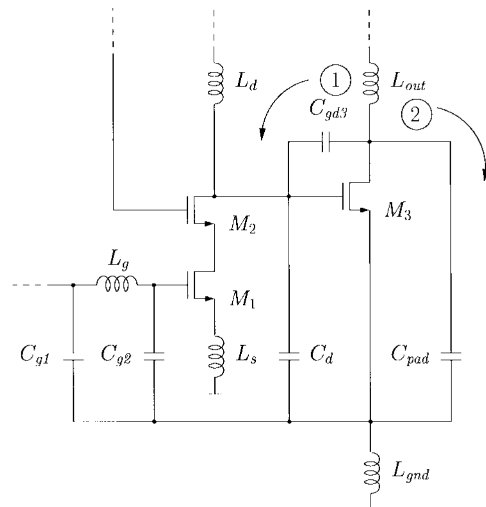

Fig. 18. Detailed LNA schematic showing parasitic reverse paths.

It is interesting that both plots exhibit some anomalies at about 1.4 GHz. On the S21 curve, the gain begins to dip sharply, whereas the S11 plot shows a bump in the reflection coefficient. This point is indicated by marker 2 on both plots. An examination of the reverse gain of the amplifier (S12) in Fig. 17 provides a plausible explanation for these anomalies. Marker 2 is positioned at the same frequency as in the two previous plots. Note that it coincides with a pronounced peak in the reverse gain. Indeed, the approximate loop gain magnitude of the LNA at marker 2 is −6 dB. This value is insufficient to cause oscillation of the amplifier, but is nonetheless substantial. Accordingly, we are compelled to attribute the formerly mentioned anomalies to this reverse isolation problem.

Another feature of the S12 characteristic is a sharp null at 1.5 GHz. This null is a clue to the source of our troubles. In Fig. 18, a partial schematic of the LNA is shown along with various significant parasitic capacitances. The substrate of the die was connected to the lowest inductance signal

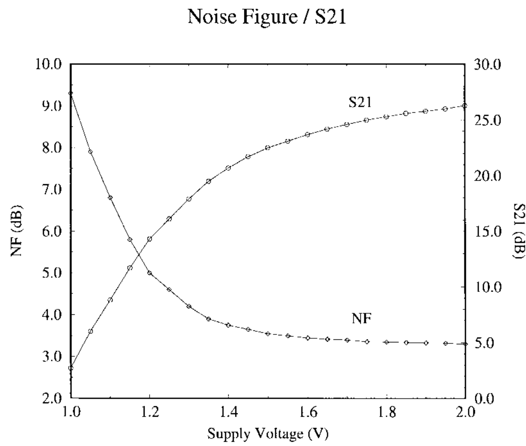

Fig. 19. Noise figure and forward gain of the LNA.

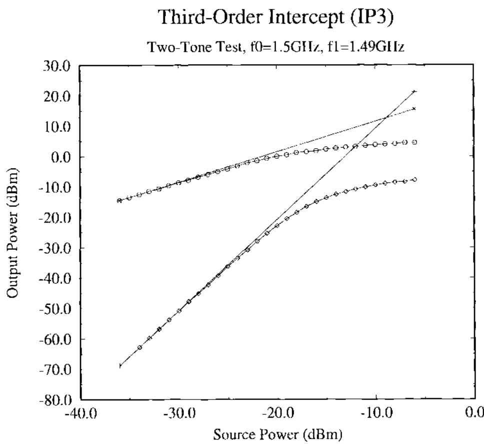

Fig. 20. Results of two-tone IP3 measurement.

ground, L_{gnd}. As shown in the diagram, this choice degrades the reverse isolation by allowing signal currents in the output driver to couple back to the input through the large parasitic capacitance of the gate inductance and its bond pad. There are actually two significant paths for this to occur, opening the possibility of cancellation at a particular frequency. Indeed, a significant phase shift along path 1 in the diagram occurs near the resonance of L_d and C_d. A null in the reverse gain could thus occur near this frequency. This problem could be mitigated by terminating the substrate differently or by moving to a differential structure.

The noise figure and gain of the LNA are plotted in Fig. 19. From this plot, we can see that at $V_{dd} = 1.5$ V, the LNA exhibits a 3.5 dB noise figure with 22 dB of forward gain. The power dissipation is 30 mW total. Of this power, only 7.5 mW is attributed to the first amplifier stage. The other 22.5 mW is used to drive 50 Ω with the open-drain output driver. This added power could be nearly eliminated if the LNA were to drive an on-chip mixer rather than an off-chip transmission line.

Although the measured noise figure exceeds the theoretical minimum of 2.1 dB, it is a simple matter to account for the difference. In particular, our theoretical predictions must be modified to include the loss of the 4-nH spiral inductor, which contributes significantly to the noise figure, and to account for the actual impedance level at the LNA input, as determined by $\omega_T L_s$. In the final amplifier, $\omega_T L_s$ was less than 50 Ω. In fact, the real portion of the input impedance, before matching, was about 35 Ω. If we assume that the 4-nH inductor possesses a Q of about three, then it would contribute about 0.38 to F in a 35 Ω environment. In addition, the theoretical minimum increases to about 2.5 dB when R_s is 35 Ω. These two effects therefore elevate the predicted noise figure from 2.1 dB to 3.3 dB. The remaining 0.2 dB may be attributed to the second stage of the amplifier.

A two-tone IP3 measurement was performed on the LNA and the results are shown in Fig. 20. The two tones were applied with equal power levels at 1.49 GHz and 1.5 GHz. The measurement indicates a −9.3 dBm input-referred third-order intercept point (+12.7 dBm output-referred). The linearity is primarily limited by M_3, due to the gain which precedes it.

The measured performance of the LNA is summarized in Table II.

TABLE II
LNA PERFORMANCE SUMMARY

Frequency	1.5 GHz
Noise Figure	3.5 dB
S21	22 dB
IP3 (Output)	12.7 dBm
1 dB Compression (Output)	0 dBm
Supply Voltage	1.5 V
Power Dissipation	30 mW
(First Stage)	7.5 mW
Technology	0.6-μm CMOS
Die Area	0.12 mm2

VII. CONCLUSIONS

We have demonstrated a low noise amplifier in a 0.6-μm CMOS process which is suitable as a first amplifier in a GPS receiver. Based on this result, we firmly believe that CMOS is a serious contender for the technology of choice in future wireless receiver designs. As CMOS progresses to smaller and smaller channel lengths, driven by the digital VLSI industry, the performance of circuits such as this one will continue to improve. Based on the results of this study, we expect noise figures of about 1.8 dB on 5 mW of power dissipation with the 0.35 μm ($\approx$0.25 μm L_{eff}) generation of CMOS.

Theoretical analysis of the amplifier architecture has demonstrated the fundamental role of induced gate noise, which is essential in defining the minimum noise figure. That in many practical cases this source of noise may *dominate* the output noise of the amplifier underscores the critical need for improved MOS noise models. Given the intense interest in RF CMOS, it is likely that improved models will be developed in the near future.

ACKNOWLEDGMENT

The authors would like to thank A. Jerng and R. Farjad-Rad for their assistance during the design of the LNA and K. Yang and A. C.-L. Lu for vital help in testing the LNA.

In addition, they are indebted to H. Swain, formerly of the Hewlett-Packard Company, for many helpful and enlightening discussions on noise in FET devices and to the anonymous reviewers of this article, whose insightful comments helped to strengthen the final manuscript.

REFERENCES

[1] A. A. Abidi, "High-frequency noise measurements on FET's with small dimensions," *IEEE Trans. Electron Devices*, vol. ED-33, pp. 1801–1805, Nov. 1986.
[2] R. P. Jindal, "Hot-electron effects on channel thermal noise in fine-line NMOS field-effect transistors," *IEEE Trans. Electron Devices*, vol. ED-33, pp. 1395–1397, Sept. 1986.
[3] S. Tedja, J. Van der Spiegel, and H. H. Williams, "Analytical and experimental studies of thermal noise in MOSFET's," *IEEE Trans. Electron Devices*, vol. 41, pp. 2069–2075, Nov. 1994.
[4] B. Wang, J. R. Hellums, and C. G. Sodini, "MOSFET thermal noise modeling for analog integrated circuits," *IEEE J. Solid-State Circuits*, vol. 29, pp. 833–835, July 1994.
[5] A. N. Karanicolas, "A 2.7V 900MHz CMOS LNA and mixer," in *ISSCC Dig. Tech. Papers*, 1996, vol. 39, pp. 50–51.
[6] A. Rofougaran *et al.*, "A 1 GHz CMOS RF front-end IC for a direct-conversion wireless receiver," *IEEE J. Solid-State Circuits*, vol. 31, pp. 880–889, July 1996.
[7] S. Sheng *et al.*, "A low-power CMOS chipset for spread-spectrum communications," in *ISSCC Dig. Tech. Papers*, 1996, vol. 39, pp. 346–347.
[8] J. Y.-C. Chang, A. A. Abidi, and M. Gaitan, "Large suspended inductors on silicon and their use in a 2-μm CMOS RF amplifier," *IEEE Electron Device Lett.*, vol. 14, pp. 246–248, May 1993.
[9] R. Benton *et al.*, "GaAs MMICs for an integrated GPS front-end," in *GaAs-IC Symp. Dig. Tech. Papers*, 1992, pp. 123–126.
[10] K. R. Cioffi, "Monolithic L-band amplifiers operating at milliwatt and sub-milliwatt DC power consumptions," in *IEEE Microwave and Millimeter-Wave Monolithic Circuits Symp.*, 1992, pp. 9–12.
[11] M. Nakatsugawa, Y. Yamaguchi, and M. Muraguchi, "An L-band ultra low power consumption monolithic low noise amplifier," in *GaAs-IC Symp. Dig. Tech. Papers*, 1993, pp. 45–48.
[12] E. Heaney *et al.*, "Ultra low power low noise amplifiers for wireless communications," in *GaAs-IC Symp. Dig. Tech. Papers*, 1993, pp. 49–51.
[13] Y. Imai, M. Tokumitsu, and A. Minakawa, "Design and performance of low-current GaAs MMIC's for L-band front-end applications," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 209–215, Feb. 1991.
[14] N. H. Sheng *et al.*, "A 30 GHz bandwidth AlGaAs-GaAs HBT direct-coupled feedback amplifier," *IEEE Microwave Guided Wave Lett.*, vol. 1, pp. 208–210, Aug. 1991.
[15] R. G. Meyer and W. D. Mack, "A 1-GHz BiCMOS RF front-end IC," *IEEE J. Solid-State Circuits*, vol. 29, pp. 350–355, Mar. 1994.
[16] K. W. Kobayashi and A. K. Oki, "A low-noise baseband 5-GHz direct-coupled HBT amplifier with common-base active input match," *IEEE Microwave Guided Wave Lett.*, vol. 4, pp. 373–375, Nov. 1994.
[17] A. van der Ziel, "Noise in solid-state devices and lasers," *Proc. IEEE*, vol. 58, pp. 1178–1206, Aug. 1970.
[18] R. P. Jindal, "Noise associated with distributed resistance of MOSFET gate structures in integrated circuits," *IEEE Trans. Electron Devices*, vol. ED-31, pp. 1505–1509, Oct. 1984.
[19] B. Razavi, R.-H. Yan, and K. F. Lee, "Impact of distributed gate resistance on the performance of MOS devices," *IEEE Trans. Circuits Syst. I*, vol. 41, pp. 750–754, Nov. 1994.
[20] A. van der Ziel, *Noise in Solid State Devices and Circuits*. New York: Wiley, 1986.
[21] ——, "Gate noise in field effect transistors at moderately high frequencies," *Proc. IEEE*, pp. 461–467, Mar. 1963.
[22] N. G. Einspruch, Ed., *VLSI Electronics: Microstructure Science*. New York: Academic, 1989, vol. 18, ch. 1, pp. 1–37.
[23] H. A. Haus *et al.*, "Representation of noise in linear twoports," *Proc. IRE*, vol. 48, pp. 69–74, Jan. 1960.
[24] P. K. Ko, C. Hu *et al.*, *BSIM3v3 Manual*, Dept. Electrical Eng. Comp. Sci., Univ. California, Berkeley, 1995.
[25] B. Wang, "Wide band noise in MOSFETs," M.S. thesis, Mass. Inst. Technol., Oct. 1992.
[26] K. B. Ashby *et al.*, "High Q inductors for wireless applications in a complementary silicon bipolar process," *IEEE J. Solid-State Circuits*, vol. 31, pp. 4–9, Jan. 1996.

Derek K. Shaeffer (S'90) received the B.S. degree from the University of Southern California, Los Angeles, in 1993 and the M.S. degree from Stanford University, Stanford, CA, in 1995 where he is currently engaged in research toward the Ph.D. degree.

Since 1992 he has worked for Tektronix, Inc., Beaverton, OR, where he cut his teeth designing A/D converter and communications circuits in CMOS and bipolar technologies. His current research interests are in CMOS and bipolar implementations of low noise, high linearity wireless communications receivers.

Thomas H. Lee (S'87–M'87) received the S.B., S.M., and Sc.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1983, 1985, and 1990, respectively.

He worked for Analog Devices Semiconductor in Wilmington, MA, until 1992, where he designed high-speed clock-recovery PLL's that exhibit zero jitter peaking. He then worked for Rambus Incorporated in Mountain View, CA, where he designed the phase- and delay-locked loops for 500 MB/s DRAM's. In 1994, he joined the faculty of Stanford University, Stanford, CA, as an Assistant Professor, where he is primarily engaged in research into microwave applications for silicon IC technology, with a focus on CMOS IC's for wireless communications.

Dr. Lee has twice received the "Best Paper" award at ISSCC.