Hot-carrier reliability of submicron NMOSFETs and integrated NMOS low noise amplifiers

Sasan Naseh, M. Jamal Deen *, Chih-Hung Chen

Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ont., Canada L8S 4K1

Received 7 March 2005
Available online 26 July 2005

Abstract

The effects of hot-carrier stress (HCS) on the performance of NMOSFETs and a fully integrated low noise amplifier (LNA) made of NMOSFETs in a 0.18 μm CMOS technology are studied. The main effects of HCS on single NMOSFETs are an increase in threshold voltage and a decrease in channel carrier mobility, which lead to a drop in the biasing current of the transistors. In the small-signal model of the transistor, hot-carrier effects appear as a decrease in the transconductance and an increase of the output conductance. No clear change was observed in the parasitic gate-source and gate-drain capacitances in the devices under test due to hot carriers. The main effects of hot carriers in the LNA were a drop of the power gain and an increase of its noise figure. The input and output matching, $S_{11}$ and $S_{22}$, slightly increased after hot-carrier stress. The third-order input-referred intercept point (IIP3) of the LNA improved after stress. This is believed to be due to the improvement of the linearity of the current–voltage (I–V) characteristics of the transistors in the LNA at the particular operating point where they were biased.

© 2005 Elsevier Ltd. All rights reserved.

1. Introduction

With the continual scale-down of the channel length of MOSFETs and on-chip passive devices such as inductors with good performance parameters, CMOS technology has become a viable choice for the implementation of radio frequency (RF) circuit blocks such as low noise amplifiers (LNA), oscillators and mixers [1–8]. By reducing the channel length of MOSFETs, however, the issues related to short channel effects such as drain induced barrier lowering (DIBL) [9], punch-through and hot carriers become important [10]. In NMOSFETs, applying the biasing voltages to the device's terminals naturally leads to the appearance of electric fields inside the device. In the saturation mode of operation, in a part of the channel near drain known as the pinch-off region, the inversion layer almost completely disappears and the electric field in this region is significantly higher than that in the rest of the channel [11]. When the carriers enter this region, they gain high energy and therefore are called hot carriers. As hot-carriers flow toward the drain, some of them collide with the silicon atoms and generate new electron—hole pairs. This process is called avalanche, similar to the breakdown of a reverse biased pn junction. The new electrons and holes generated during the avalanche process are swept by the drain and substrate terminals, respectively. The current flowing through the substrate terminal is called $I_{sub}$. During the avalanche process, hot carriers can also impinge on the atomic bonds between silicon and silicon dioxide.

* Corresponding author. Tel.: +1 905 525 9149x27137; fax: +1 905 525 4407.
E-mail address: jamal@mcmaster.ca (M.J. Deen).
at the interface and break them. The un-bound atomic bonds at the interface are known as interface traps [12–15]. An increase in the density of interface traps can result in an increase of the threshold voltage [11,13] and a reduction of the mobility of the electrons in the channel [16,17], which lead to a lower drain current and transconductance. The rate of damage to the interface caused by hot carriers is proportional to both the biasing current of the transistor \( I_{ds} \) and to the intensity of the electric field in the pinch-off region. At the same time, \( I_{sub} \) is also dependent on \( I_{ds} \) and the electric field in the channel in the following form [18]:

\[
I_{sub} \sim I_{ds} e^{-\phi_i/qE_m},
\]

in which \( \phi_i \) is the minimum energy that a hot electron must have in order to cause impact ionization with reported values in literature between 1.12 eV and 3 eV [18], \( q \) is the electron charge, \( \lambda \) is the mean-free path of hot electrons which is related to \( \phi_i \) according to \( \lambda = \phi_i/(1.6 \times 10^6) \) with \( \lambda \) in meter and \( \phi_i \) in electronvolt [18], and \( E_m \) is the maximum electric field in channel, which occurs near drain end of the channel. Since both \( I_{sub} \) and the rate of generation of hot carrier damage to the interface follow the variations of \( I_{ds} \) and the channel electric field, then \( I_{sub} \) can be used as a measure of the rate of hot carrier damage generation in the device.

Hot-carrier generation is a strong function of the biasing voltages and the device's channel length. In Fig. 1, the variation of \( I_{sub} \) versus gate voltage \( V_{gs} \) with drain voltage \( V_{ds} \) as a parameter is shown. It is seen that the maximum \( I_{sub} \) occurs at \( V_{gs} \approx V_{ds}/2 \) similar to that in [11]. It is also seen that \( I_{sub} \) is almost an exponential function of the drain voltage; with an increase of 0.1 V in \( V_{ds} \) resulting in an almost twofold increase in \( I_{sub} \).

To observe the effect of channel length on the hot-carrier generation, the substrate current in two devices with different channel length can be compared. However, to have a fair comparison, the width of the two devices should be such that their performance (i.e., their drain currents) are the same. Choosing devices with square shape gates guarantees that the two devices will have the same performance to a first-order approximation. Based on this argument, the measured currents in the devices with gate's length \( L \) and width \( W \) not equal, should be normalized to the ratio of their width-to-length \( W/L \) and then be compared. In Fig. 2(a) and (b), the normalized drain and substrate current, respectively, measured for two devices with different channel lengths are shown. It is seen that although the normalized drain current in the two devices are almost the same, the normalized substrate current in the device with shorter channel length is much higher than that of the device with longer channel length. This observation clearly shows that by reducing the channel length, hot-carrier generation strongly increases.

In addition to the generation of interface traps explained above, as hot electrons in the channel flow from
the source towards the drain terminal, due to scattering, some of them may be redirected towards the gate oxide and penetrate into the gate oxide. A fraction of these electrons, which are called lucky electrons, permanently stay in the gate oxide and are known as trapped charges \[19\]. Trapped charges in the gate oxide increase the threshold voltage and therefore, affect the device’s \( I-V \) characteristics. However, it has been reported that drain avalanche is the dominant mechanism of damage caused by hot carriers in NMOSFETs \[20–23\].

In MOS integrated circuits, the degradation of the device characteristics due to interface traps and trapped charges can affect the circuit’s performance. Therefore, it is important to study the effects of hot-carrier damages on the performance of the circuits made of NMOSFETs. Recently, there have been some studies of hot-carrier effects on the RF performance of single NMOSFETs and RF circuits fabricated in CMOS technology reported \[24–35\]. For the case of CMOS LNAs, there have been some investigations carried out on the effects of hot carriers on their performance; however, in most of these works, few experimental results have been provided. In this paper, we expand on our previously published result \[35\] by including additional measurement results of single NMOSFETs which can be used for lifetime estimation, a more detailed analysis of the effect of stress on the noise performance of single devices and some details of the extraction of the parameters of the LNA.

In the next section, the devices and the circuit of the LNA which were used for hot-carrier reliability experiments, and the experimental procedures are described. The result of hot-carrier measurements on the DC and small-signal parameters of single NMOSFETs are presented and a method for lifetime estimation is proposed in Section 3. In Section 4, the effects of hot-carrier stress on the gain, matching, linearity and noise figure of a fully integrated LNA made of NMOSFETs will be explained, and finally, a summary of important results is provided in Section 5.

2. Device and circuit description and measurement procedure

The single transistors used for experiments were NMOSFETs fabricated in a commercial 0.18 \( \mu \)m CMOS technology. The transistor’s channel width were 120 \( \mu \)m and 125 \( \mu \)m and the gate oxide thickness \( t_{\text{ox}} = 40 \) Å. The schematic of the LNA is shown in Fig. 3. The circuit is fully integrated with metal–insulator–metal capacitors and spiral inductors. The resistor \( R_b \) is implemented with N-well diffusion in the p-type substrate. The layouts of the transistors and the LNA were designed with pads in the ground–signal–ground configuration at input and output so that scattering (\( S \)) parameters could be measured with 3-point probes.

Fig. 3. Schematic of the LNA used in this work.

The LNA circuit is made of two stages, with each stage being a common-source amplifier. The inductor \( L_s \) at the source of the transistor in the first stage is for the purpose of providing input matching by the inductive source degeneration technique \[3\].

The hot-carrier stress were applied to the single NMOSFETs and the transistors in the LNA for different periods of time for a total time of 500 s, and after each stressing period, the parameters of the transistors and the LNA were measured. To apply stress, the devices were biased at \( V_g = 0.7 \) V and \( V_d = 3.3 \) V and 3.5 V. A HP4145B Semiconductor Parameter Analyser was used for providing the DC biasing for the single NMOSFETs and the LNA. The \( S \)-parameters were measured with a HP8510B network analyser. The dummy pad structures were also measured and used for de-embedding the parasitic effects of the pads \[36\]. The noise figure (NF) of the LNA was measured with HP8970B noise figure meter.

The noise parameters of the single transistors were measured with an ATN noise parameter measurement system. To measure the nonlinearity parameters of the LNA, the third-order input-referred intercept point, IIP3, and the 1-dB compression point, \( P_{1\text{dB}} \), were measured. For measurement of IIP3, two signal generators HP8662A and Agilent 83752A were used to generate input signals at 1.09 GHz and 1.1 GHz. The output spectrum of the LNA was measured using a HP8594E spectrum analyser. The same setup was used for measurement of 1-dB compression point. The control of the instruments and data acquisition were done by computer.

3. Hot-carrier effects on single NMOSFETs

As explained in the introduction, damages in the form of interface traps created by hot carriers in the NMOSFETs, cause an increase of its threshold voltage and also a drop in the electron’s mobility in the channel. The threshold voltage in a n-channel MOSFET can be written in terms of the device’s parameters as
where the parameters are defined as, \( \varphi_0 = V_T \ln(N_A/n_i) \); \( n_i \), the intrinsic electron concentration; \( V_T = kT/q \), the thermal voltage; \( N_A \), the bulk dopant concentration density; \( q \), the electron charge; \( \varepsilon_s \), the silicon permittivity; \( \varphi_{ms} \), the gate–substrate work function difference; \( Q_{th} \), the trapped charge density in the oxide; \( Q_{tr} \), the interface trap charge density; and \( V_{\text{thr}} \), the source–substrate voltage.

From (2), it is seen that increase of interface trap charge density \( Q_{th} \) leads to an increase of the threshold voltage. In Fig. 4, the normalized variation of threshold voltage with stress time is shown for a single value of gate voltage \( V_g = 0.7 \) V and several values of drain voltage \( V_{ds} \) for the stress condition. From the graphs in Fig. 4, it is observed that the variation of the threshold voltage with stressing time \( t \) follows a power law in the form of

\[
\Delta V_{\text{th}}/V_{\text{th0}} \approx K_1 t^{n_1},
\]

in which the coefficient \( K_1 \) is a function of the drain voltage during stress, and \( n_1 \approx 0.67 \) for the value of \( V_g = 0.7 \) V used in the stress measurements. Since the change in the threshold voltage is caused by increase of interface traps during the avalanche process, it might be concluded that the generation of the interface traps for the stress conditions considered here has the same trend as shown by (3).

It was also mentioned that another effect of the damage due to hot carriers is a drop of mobility of electrons in the inversion layer after stress [16,17]. This is because with the increase of interface traps density with stress, the charge trapped at the interface traps also increases and affects the mobility of the channel carriers through Coulomb scattering. A drop in mobility and an increase of threshold voltage lead to a lower drain current in the device. In Fig. 5, a typical measured \( I_d-V_{ds} \) of a single NMOSFET is shown for several gate voltages at different stress conditions. The normalized variation of drain current with stress time are shown in Fig. 6 for several different levels of hot-carrier stress. It can be seen that similar to the threshold voltage, the normalized variation of the drain current with stress time follows a power law:

\[
\Delta I_{ds}/I_{ds0} \approx k_2 t^{n_2},
\]

where the parameters are defined as, \( \varphi_0 = V_T \ln(N_A/n_i) \); \( n_i \), the intrinsic electron concentration; \( V_T = kT/q \), the thermal voltage; \( N_A \), the bulk dopant concentration density; \( q \), the electron charge; \( \varepsilon_s \), the silicon permittivity; \( \varphi_{ms} \), the gate–substrate work function difference; \( Q_{th} \), the trapped charge density in the oxide; \( Q_{tr} \), the interface trap charge density; and \( V_{\text{thr}} \), the source–substrate voltage.

From (2), it is seen that increase of interface trap charge density \( Q_{th} \) leads to an increase of the threshold voltage. In Fig. 4, the normalized variation of threshold voltage with stress time is shown for a single value of gate voltage \( V_g = 0.7 \) V and several values of drain voltage \( V_{ds} \) for the stress condition. From the graphs in Fig. 4, it is observed that the variation of the threshold voltage with stressing time \( t \) follows a power law in the form of

\[
\Delta V_{\text{th}}/V_{\text{th0}} \approx K_1 t^{n_1},
\]

in which the coefficient \( K_1 \) is a function of the drain voltage during stress, and \( n_1 \approx 0.67 \) for the value of \( V_g = 0.7 \) V used in the stress measurements. Since the change in the threshold voltage is caused by increase of interface traps during the avalanche process, it might be concluded that the generation of the interface traps for the stress conditions considered here has the same trend as shown by (3).

It was also mentioned that another effect of the damage due to hot carriers is a drop of mobility of electrons in the inversion layer after stress [16,17]. This is because with the increase of interface traps density with stress, the charge trapped at the interface traps also increases and affects the mobility of the channel carriers through Coulomb scattering. A drop in mobility and an increase of threshold voltage lead to a lower drain current in the device. In Fig. 5, a typical measured \( I_d-V_{ds} \) of a single NMOSFET is shown for several gate voltages at different stress conditions. The normalized variation of drain current with stress time are shown in Fig. 6 for several different levels of hot-carrier stress. It can be seen that similar to the threshold voltage, the normalized variation of the drain current with stress time follows a power law:

\[
\Delta I_{ds}/I_{ds0} \approx k_2 t^{n_2},
\]
and from the graphs in Fig. 6, it is found that for the value of $V_g = 0.7$ V used in the stress measurements, $n_2 \approx 0.55$ and $k_2$ is a function of the drain voltage during the stress.

As a result of the drop in the biasing current of the transistor and the channel carrier mobility, the transconductance $g_m$ also decreases. In Fig. 7, the normalized variation of $g_m$ with stress time are shown for a particular operating point and different stress conditions. Here, it is shown that the normalized variation of $g_m$ also follows a power law $k_3 t^{n_3}$ with the value of exponent $n_3 \approx 0.57$ and $k_3$ dependent on the drain voltage during the stress. It is possible that the degradation of mobility due to stress is the dominant factor in the decrease of both the drain current and the transconductance, which also could be the reason why $n_2$ and $n_3$ are close, as the dominant degradation mechanism of both is the same.

The slope of the $I_g$–$V_{ds}$ graphs in the saturation region is known as the output conductance $g_{ds}$ of the transistor and $g_{ds}$ can have an important effect on the operation of CMOS analog integrated circuits. From Fig. 5, it can be seen that this slope changes with hot-carrier stress. Variations of $g_{ds}$ with stress time are shown in Fig. 8 for a particular biasing point and several stress condition. In contrast to threshold voltage and transconductance, the variations of $g_{ds}$ with stress time when plotted in a log–log scale shows some saturation, especially at higher stress conditions.

Following the same method as in [37], it is possible to estimate the lifetime of a NMOSFET at different levels of stress using the data of the variations of parameters $V_{th}$ and $g_m$ with stress time presented above. As explained in the introduction, the substrate current $I_{sub}$ can be monitored as a measure of the level of hot-carrier stress. The device lifetime $\tau$ can be defined as the time it takes for a device parameter, e.g., $V_{th}$, to degrade a certain amount, e.g., 10%. In Fig. 9, an example of lifetime prediction is shown. The device’s lifetime is plotted in terms of the density of the substrate current per gate width $I_{sub}/W$. By using the linear density of substrate current as opposed to substrate current, the effect of the device’s width is taken into account. With the help of this graph, the device’s lifetime when it is biased at normal operating condition, e.g., $I_{sub}/W = 1.67 \times 10^{-8}$ A/µm, can be estimated. Two examples are shown in Fig. 9 in which the criteria for the amount of degradation of the monitored parameter $V_{th}$, are 3% and 10%. According to the graphs in Fig. 9, the device’s lifetime is a function of the substrate current density in the form of [37]

$$\tau \approx A(I_{sub}/W)^{-m}, \tag{5}$$

for $m = 2$, $W_{total} = 120$ µm, $L = 0.18$ µm, gate voltage during stress was 0.7 V and the drain voltages are specified in the figure.
in which \( A \) is a function of the change used as index of lifetime measurement (e.g., 3% or 10%) and \( m \approx 2 \) for the stress condition considered in these experiments.

Using \( g_m \) as the monitored parameter for lifetime estimation leads to a similar result, as shown in Fig. 10.

Another important effect in the device caused by hot-carrier stress is the degradation of its noise performance. Due to the thermal noise in the channel, small fluctuations occur in the drain current even when the potentials at different terminals of the transistor are kept constant. These small fluctuations in the drain current can be presented in the form of a current source between the drain and the source terminal. A simple small-signal model of the device is shown in Fig. 11 in which the noise power current source \( I_d \) connected between the drain and the source terminals represents the channel thermal noise. The conductance \( g_{ds} \) represents the variation of the drain current with drain voltage, which is the slope of the \( I_d-V_{ds} \) graphs in the saturation region, shown in Fig. 5.

The power spectral density of the channel thermal noise current was extracted using the technique described in [40] after different periods of stress and the results are shown in Fig. 12. There is a decrease of the channel current noise after stress which can be explained by considering the relation [10,38,48]:

\[
\bar{I}_d = 4kT \frac{\mu}{L_{cl}^2} (-Q_1),
\]

in which \( k \) is the Boltzmann constant, \( T \) is the absolute temperature, \( \mu \) is the channel electron mobility, \( L_{cl} \) is the electrical channel length and \( |Q_1| \) is the magnitude of the charge density in the inversion layer in the channel. Since the threshold voltage increases after stress, then the inversion charge density \( Q_1 \) decreases after stress for a constant \( V_g \). Also, as explained before, the channel’s carrier mobility \( \mu \) decreases after stress and therefore, \( \bar{I}_d \) decreases after stress for a fixed biasing point.

Although \( \bar{I}_d \) decreases, the overall noise performance of the device deteriorates. This can be seen if the current noise power \( \bar{I}_d^2 \) is referred to the device’s input gate terminal by dividing it by \( g_{ds}^2 \), and is presented in the form of a voltage noise power source \( \bar{I}_d^2 \). As shown in Fig. 12, the power spectral density of this voltage increases with stress time.

To summarize, the important effects on the small-signal model parameters of the device are a decrease of transconductance \( g_m \), an increase of output conductance \( g_{ds} \), and a deterioration of the noise performance of the device. The parasitic gate-source \( C_{gs} \) and gate-drain \( C_{gd} \) capacitances of single transistors were extracted from the measured S-parameters following the method in [38] after different stress times and at different operating points similar to the operating point of the transistors in the LNA. Unlike what has been reported before [30,33,39], there was no significant change observed in the values of the parasitic capacitances \( C_{gs} \) and \( C_{gd} \) in the transistors tested.

In Table 1, a summary of important changes in a NMOSFET caused by hot-carrier stress is presented.
4. Hot-carrier effects on the performance of NMOS-LNA

To characterize a linear two-port network at microwave frequencies, usually scattering parameters of the network are measured [41,42]. Scattering parameters or in short, $S$-parameters, are defined in terms of the incident and reflected power waves at each port of the network [43]. A parameter $S_{ij}$ is defined as the ratio of the power flowing out of port $i$ to the power incident to the port $j$. The power gain of an LNA is represented with the parameter $S_{21}$. In Fig. 13, the measured and simulated parameter $S_{21}$ of the LNA are shown versus frequency for the circuit before and after stress. For the simulation of the LNA, a SPICE netlist was extracted from its layout and the inductors were replaced with their $π$-equivalent models [44]. For the simulation of the circuit after stress, three parameters: mobility, threshold voltage and channel length modulation in the BSIM3 model of the transistors were adjusted as shown in Table 2, in order to take into account the effects of hot-carrier degradation.

The variation of $|S_{21}|$ at 1.1 GHz with stress time presented in Fig. 14 shows a drop of about 4 dB in the gain of the LNA after 500 s stress. By considering that the main changes in the small-signal model of the transistors are the degradation of $g_m$ and $g_{ds}$, it can be said that the decrease of the gain of the LNA is mostly due to the drop of $g_m$, and to some extent due to the increase of $g_{ds}$ after hot-carrier stress.

The parameters $S_{11}$ and $S_{22}$ can be considered as a measure of the input and output matching of the LNA, respectively. In Fig. 15, variations of these two parameters with stress time are shown. For both parameters, their variations at 1.1 GHz where the maximum gain occurs are shown. Also, variations of $S_{11}$ and $S_{22}$ at 1.8 GHz and 1.45 GHz, respectively, where the best matching (i.e., lowest values of $S_{11}$ and $S_{22}$) occurs are

![Figure 13](image1.png)

![Figure 14](image2.png)

![Figure 15](image3.png)

Table 1
Summary of important changes in the single NMOSFET

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Direction of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{th}$</td>
<td>Increases</td>
</tr>
<tr>
<td>$I_{ds}$</td>
<td>Decreases</td>
</tr>
<tr>
<td>$g_m$</td>
<td>Decreases</td>
</tr>
<tr>
<td>$g_{ds}$</td>
<td>Increases</td>
</tr>
<tr>
<td>$C_{gd}$</td>
<td>No change</td>
</tr>
<tr>
<td>$C_{gs}$</td>
<td>No change</td>
</tr>
</tbody>
</table>

Table 2
BSIM3v3 parameters which changed to model the hot carrier stress

<table>
<thead>
<tr>
<th>Parameter’s name</th>
<th>Original value</th>
<th>Modified value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold voltage, $V_{TH0}$ (V)</td>
<td>0.448</td>
<td>0.528</td>
</tr>
<tr>
<td>Mobility, $U_0$ ($m^2/s$)</td>
<td>0.0439</td>
<td>0.0379</td>
</tr>
<tr>
<td>Channel length modulation parameter, $P_{CLM}$</td>
<td>1.1</td>
<td>8.1</td>
</tr>
</tbody>
</table>

Fig. 13. (a) Measured and (b) simulated $S_{21}$ as a function of frequency, of the LNA before and after hot carrier stress.
shown. Some slight increase in the values of these two parameters were caused due to stress. At the input of the LNA, it is speculated that the gate–source and gate–drain capacitances are the dominant elements and therefore variations of transconductance $g_m$ and output conductance $g_{ds}$ of the input transistor do not affect the input impedance significantly. Similarly, at the LNA’s output, it can be said that the inductor at the drain of $M_2$, which is in parallel with the output conductance of $M_2$, is the dominant element in determining the LNA’s output impedance and therefore variation of $g_m$ and $g_{ds}$ of $M_2$ do not have a major effect on the output matching.

The parameter $S_{12}$ is a measure of transmission of a signal from output to the input port of a two-port network. For the LNA in this work, the measured parameter $S_{12}$ was very small (<0.01), both before and after stress and was not affected significantly by stress. This small value of $S_{12}$ indicates that the LNA’s input is highly isolated from its output. This high level of isolation can be qualitatively explained by considering that the major path of signal from the LNA’s output to its input is through the gate–drain parasitic capacitance of the transistors, which has a rather high-impedance and therefore, signal transmission from output to input is not significant.

The effects of stress on the nonlinearity of the LNA are studied by considering the third-order input-referred intercept point IIP3 and the 1-dB compression point $P_{1\text{dB}}$. To measure the parameter IIP3, two signals at frequencies $f_1 = 1.09$ GHz and $f_2 = 1.1$ GHz with equal amplitude were combined with a power-combiner and the combined signal was applied to the LNA’s input. At the output of the LNA, in addition to the amplified signals at $f_1$ and $f_2$, the cross modulation of the original signals also appear due to the nonlinearity of the LNA. Cross-modulated components appear at frequencies $mf_1 \pm nf_2$ where $m$ and $n$ are integers. In Fig. 16, the result of a typical measurement of the spectrum at the LNA’s output is shown. By increasing the power of the input signals, the signals at the LNA’s output also increase. The signals at LNA’s output at frequencies $(2f_1 - f_2)$ and $(2f_2 - f_1)$ grow almost 3 times faster than the signals at $f_1$ and $f_2$ [45]. This is illustrated in Fig. 17. The extrapolated value of the input power at which the signals at $f_1$ and $(2f_1 - f_2)$ become equal is called the IIP3 and is considered as a measure of the nonlinearity of the LNA [45].

Variations of IIP3 versus stress time is shown in Fig. 18, which shows an improvement of about 1.3 dB after stress. To explain this, the nonlinear elements in the LNA circuit are investigated. One such element is the current source between the drain and source of each transistor, which is a function of the gate and drain voltages $I_d(V_g, V_d)$. Since in the saturation region, the drain current is a strong function of the gate voltage $V_g$. 

![Fig. 15. Variation of the $S_{11}$ and $S_{22}$ of the LNA with stress time.](image1)

![Fig. 16. A typical measured spectrum at the LNA’s output during IIP3 measurement. The intermodulation signals are at $(2f_2 - f_1) = 1.11$ GHz and $(2f_1 - f_2) = 1.08$ GHz.](image2)

![Fig. 17. Determining IIP3 from the measured output component as a function of the input signal power.](image3)
whereas it does not change much with drain voltage $V_d$, its nonlinearity with respect to $V_g$ is considered. For this purpose, a new parameter called the third-order input-referred intercept voltage $V_{ip3}$ is defined which is a measure of the linearity of the $I_d-V_g$ characteristics of the device [46]. To define this parameter, a sinusoidal voltage with the amplitude $v_p$ is added to the biasing DC voltage at gate and the potentials at other terminals are kept constant. In response to this input, a periodic current will flow through the device which has a fundamental component with the same frequency as the input sinusoidal voltage, plus higher order harmonics. By increasing the amplitude $v_p$, the higher order harmonics in the drain current will also increase. The value of $v_p$ at which the fundamental and the third harmonic in $I_d$ become equal is defined as $V_{ip3}$ and it can be shown to be [46]

$$V_{ip3} = \sqrt{\frac{24 g_m}{g_{m3}}}$$

(7)

in which $g_m$ is the transistor’s transconductance and $g_{m3}$ is the third derivative of drain current with respect to gate voltage. This parameter before and after stress is shown in Fig. 19. It can be seen that the value of gate biasing voltage at which drain current has its highest linearity with respect to $V_g$ (i.e., highest $V_{ip3}$) has shifted to higher values after stress. Therefore, it can be said that depending on the biasing, the device can become more linear or more nonlinear after stress.

In addition to the drain current, the depletion capacitance of the drain–substrate and the source–substrate pn junctions are also nonlinear elements. Since these capacitances are far from the pinch-off region of the device, it can be assumed that hot-carrier damage does not have a significant effect on these capacitances.

Overall, the improvement in IIP3 of the LNA after stress might be considered to be due to the improvement of the linearity behaviour of the transistors in the circuit at the particular operating point they were biased.

To measure $P_{1\text{dB}}$, a signal at 1.09 GHz was applied to the LNA’s input and its value was increased gradually. As long as input is small, the LNA acts as a linear circuit and its output is proportional to the input. As the input becomes larger, the linear relation becomes increasingly less valid and the output will no longer be proportional to the input. This is illustrated in Fig. 20. The input power at which output power is 1 dB below its linearly extrapolated value, is called the 1-dB compression point, $P_{1\text{dB}}$ of the LNA. In Fig. 18, the variation of $P_{1\text{dB}}$ of the LNA is also shown. It does not show a clear change with stress. It should be mentioned that IIP3 is a measure of the nonlinearity of the LNA at low and moderate input signal levels, while $P_{1\text{dB}}$ is usually the point at which the input signals are large and the LNA output is limited by the supply voltage or when the transistors move into the cut-off region [47]. Therefore,
it is plausible that these two parameters show different sensitivities with respect to hot-carrier stress.

The last parameter considered for the LNA is its noise figure (NF). In Fig. 21(a), the measured noise figure of the LNA before and after stress are shown. An increase of about 1.5 dB occurs in the noise figure at a frequency of 1.1 GHz due to stress. Although, as explained before, the noise current source in the small-signal model shown in Fig. 11 decreases, which tends to improve the noise performance, but this is overshadowed by the drop of the transconductance of the transistors which reduces the LNA's gain and increases its noise figure. The noise figure of the LNA was simulated using the same circuit as that were used for the simulation of the power gain, except that the transistor’s channel noise were slightly modified by adding an extra current noise with a value of about $2.76 \times 10^{-23}$ A$/\sqrt{\text{Hz}}$ both before and after stress. The reason was to bring the channel noise to a level in agreement with the experimental result [48]. There is a good agreement between the simulation and measurement results in Fig. 21, which confirms that the drop of gain of the LNA is the main reason for the degradation of its noise figure.

Finally, in Table 3, a summary of measured changes in a NMOS-LNA caused by hot-carrier stress is presented.

Table 3  
Summary of the effect of hot carrier stress on the important parameters of the LNA

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Direction of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gain, $S_{21}$</td>
<td>Decrease</td>
</tr>
<tr>
<td>Matching</td>
<td>Slight degradation</td>
</tr>
<tr>
<td>IIP3</td>
<td>Increase</td>
</tr>
<tr>
<td>$P_{1\text{db}}$</td>
<td>No change</td>
</tr>
<tr>
<td>NF</td>
<td>Increase</td>
</tr>
</tbody>
</table>

5. Conclusion

The effects of hot-carrier stress on the RF performance of single NMOSFETs with 0.18 μm channel length and on a fully integrated low noise amplifier fabricated in a 0.18 μm CMOS technology were investigated.

In single NMOSFETs, the main effects of hot carriers are an increase of the threshold voltage and a drop of carriers’ mobility in the channel. As a result, for any constant biasing voltages to the drain and gate terminals, the drain current decreases.

In the small-signal model of the transistor, the effects of hot-carriers result in a decrease of the transconductance $g_m$, an increase of the output conductance $g_{ds}$ and a deterioration of the device’s noise performance.

The degradation in the noise performance itself is mostly a result of the decrease of the transconductance.

The NMOSFET’s lifetime can be defined in terms of the length of time required for a certain amount of change in one of the device’s parameter. The substrate current per gate width can be used as a measure of level of hot-carrier stress. By measuring this parameter at several levels of hot-carrier stress, it is possible to estimate the device’s lifetime at normal operating conditions.

The main effects of hot-carrier stress on the performance of a LNA were a decrease of its power gain and an increase in its noise figure. A slight degradation in the input and output matching of the LNA due to hot carriers were also observed. The main cause for these changes is the decrease of the transconductance of the transistors, and to some extent, an increase of their output conductance.

Depending on the biasing point at which the transistors are operated, hot carriers can lead to a higher linearity of the LNA as it was observed that the third-order input-referred intercept point of the LNA IIP3 increased after stress.
Acknowledgements

The authors wish to acknowledge Canadian Microelectronics Corporation for supplying the devices and circuits studied here. This work was funded by NSERC and the Canada Research Chair Program.

References


