Inversion Layer Quantization Impact on the Interpretation of 1/f Noise in Deep Submicron CMOS Transistors

A. Mercha, E. Simoen, G. Richardson and C. Claeys*
IMEC, Kapeldreef 75, B-3001 Leuven, Belgium
*also at E.E. Dept, KU Leuven, Belgium
Mercha@imec.be

Abstract

This paper discusses the impact of inversion layer quantization on the interpretation of the 1/f noise characteristics in deep submicron CMOS transistors. In order to describe the strong gate voltage dependence of the input-referred noise spectral density, a model will be developed which consistently takes account of inversion layer quantization. The only adjustable parameter is the density of near interface oxide traps, which sets the level of the flat-band voltage noise spectral density in weak inversion. It is shown that a good agreement with measured data obtained on 0.13 µm CMOS transistors is found both for n- and p-channel devices. Finally, the implications with respect to practical noise modeling will be discussed.

1. Introduction

It is well established nowadays that in order to properly model the dc characteristics of deep submicron transistors, one has to account for the inversion layer quantization [1]-[3]. These quantization effects originate from the band splitting induced by the high vertical electric field, at sufficiently high gate bias V_G and substrate doping density. The major consequences are an increase of the band gap at the interface, whereby the first allowed energy level is shifted by an amount ΔE above the conduction band minimum, for an n-channel device. This has an impact on the threshold voltage V_T. In addition, the electrical gate oxide thickness becomes higher than the physical one (t_oxx), due to the peaked distribution of the charge density in the 2-dimensional electron gas (2-DEG). This results in a gate voltage dependent effective oxide thickness t_oxeff and gate capacitance C_oxeff. As a consequence, also the 2-DEG charge density and the drain current I_D are modified by quantization.

Although some groups have recognized the importance of quantum effects (QE) on the low-frequency (LF) noise of submicron MOSFETs [4]-[6], no rigorous study has been undertaken so far. It has been pointed out that if QE are ignored, the density of oxide traps N_t derived from the 1/f noise, based on a number-fluctuations model can be seriously over- or underestimated, depending on the nature of the traps [4]. It has also been shown that inversion layer quantization affects the time constants of Random Telegraph Signals (RTSs) [5]-[6].

It is the aim of this paper to assess the impact of QE on the 1/f noise in deep submicron transistors. Theory will be compared with results from a 0.13 µm CMOS technology. The main conclusions coming out of this study is that if the inversion layer quantization is neglected, inaccurate oxide trap densities will be derived from the input-referred voltage noise spectral density S_V at sufficiently high gate overdrives. A similar conclusion will hold for the extraction of the scattering coefficient derived in the frame of a correlated number fluctuations model [7]-[9] or for the Hooge parameter, derived from the normalized drain current spectral density S_{ID}/I_D^2 [7].

2. Experimental

The devices studied have been processed in a 0.13 µm CMOS technology and correspond with a reoxidized nitrided gate oxide (RNO) of 2 nm thickness. LDDs and HALO implants were utilized to control the short-channel effects. Noise was measured on transistor arrays with a common gate and source contact and a fixed width W=10 µm. The polysilicon gate length varied from L=0.1 to 10 µm. The devices were biased in linear operation for an absolute drain voltage |V_DS|=0.05 V. The spectra were recorded using a BTA9812A and Noise Pro software to drive a HP35665A Dynamic Signal Analyzer.

Inversion layer quantization and polysilicon gate depletion were modelled using the approach derived in Ref. 2.

The simulation results were calibrated to Capacitance-Voltage (C-V) measurements at 100 kHz obtained on MOS structures from the same wafer. A good agreement was obtained for the 1D simulation assuming a 5.10^{17} cm^{-3} uniformly doped p- and n-well and a 10^{20} cm^{-3} doped polysilicon gate.

3. Results

Figure 1 summarizes the 1/f noise data at f=10 Hz obtained for the n- and p-channel transistors. The curves show the extremes (minimum and maximum) and the average values. Error bars indicate the variance. There are a number of striking observations in Fig. 1. First, the noise of the p-channel transistors is larger than the value for n-MOSFETs, which seems to be typical for scaled
technologies [10]. Next, there exists a strong gate voltage
dependence of $S_{V_G}$, irrespective of the channel length $L$,
which was also observed before [9]-[11]. Note that in
Fig. 1, the normalized input spectral density $L x S_{V_G}$ is
represented, to filter out the gate area dependence. From
this, one can conclude that to a first approximation,
channel-length dependent mechanisms like the series
resistance or the lateral doping profile cannot explain the
observed dependence, for the gate voltage range studied.

Fig. 1. Normalized input-referred noise $L x S_{V_G}$ for the
0.13 $\mu$m n- and p-channel transistors studied. $W=10$
$\mu$m and $V_{DS} - 0.05$ V, $f=10$ Hz.

Before tackling the origin of the $S_{V_G}$-$V_G$ dependence
in more detail, it is worthwhile to identify the most likely
origin of the 1/f noise. For that, one can follow the
procedure proposed by Ghibaudo et al. [7], which relies
on the representation of the normalized current noise
spectral density versus the drain current. The data corresponding
with Fig. 1 are represented in Figs 2 and 3.

For the n-channel devices, one can clearly see the
leveling off of $S_{D} / I_0^2$ in weak inversion, strongly
pointing to a number fluctuations origin. This can be
further verified by comparing with the $g_m / I_0^2$ ratio, with
g$_m$ the transconductance. For the p-channel devices, the
results in Fig. 3 are a bit more dubious, as only the 0.3
$\mu$m transistor shows a tendency to level off in weak
inversion. It has been pointed out recently that especially
for p-channel devices the unified modeling approach may
not be the correct one [9].

Fig. 3. Normalised noise power density for the 0.13
$\mu$m p-MOSFETs studied, at $f=10$ Hz, $V_{DS}=0.05$ V.

4. Theory

In case that charge trapping is at the origin of the 1/f
noise, one usually follows an approach based on the flat-
band noise power spectral density $S_{V_{FB}}$ to extract the
effective density of oxide traps from the weak inversion
part of Fig. 1 [7]. $S_{V_{FB}}$ is given by:

$$S_{V_{FB}} = \frac{q^2 kT \lambda \tilde{N}_{\alpha}}{W L C_{ox} C_{ox} f}$$  \hspace{1cm} (1)

In Eq. (1), $k$ is Boltzmann’s constant, $T$ the absolute
temperature, $q$ the electronic charge and $\lambda$ is a tunneling
parameter of the order 0.1 nm, which represents the
penetration depth of the electron (or hole) wave function
in the oxide. In principle, the only adjustable parameter
in Eq. (1) is the effective oxide trap density $N_{\alpha}$ which is
proportional to the NOIA parameter in the BSIM3v3 model [9].

Classically, $N_{\alpha}$ is derived from the input-referred
noise in weak inversion assuming $C_{ox}$ corresponds to the
physical gate oxide thickness. It is clear, however, that
this approach is questionable when applied to the data of
Figs 1-3. For the gate voltage range of the measurements,
the effective oxide thickness is expected to be larger than
$\tilde{t}_{ox}$. This follows also from the calculated result of Fig. 4,
representing the ratio $C_{ox}^{eff} / C_{ox}$ versus gate bias.
Therefore, neglecting inversion layer quantization will
result in a too high $N_{\alpha}$ by a factor of 2. The
overestimation is more pronounced for p- than for n-
channel devices. It turns out in fact that the average (or
effective) $N_{\alpha}$ extracted for the p-channel devices is
slightly smaller ($5x10^{16}$ cm$^{-3}$ eV$^{-1}$) than for the n-
MOSFETs ($1x10^{17}$ cm$^{-3}$ eV$^{-1}$). The difference stems from
the smaller perpendicular effective mass for holes, compared with electrons [3]. Note that the derived oxide trap densities are reasonable values, compared with literature data [11].

Another interesting parameter is the hole or electron capture cross section for interface traps $\sigma_{int}$. In the classical approach, it is assumed constant, while in case of inversion layer quantization, the capture cross section increases strongly with the gate bias, which is derived from the ratio $\sigma_{int}/\sigma_{id}$ [1]. Following the analysis of Ref. 1, we obtain the result of Fig. 5, as a function of $V_{G}$.

In the same graph, the input-referred noise for $n$- and $p$-MOSFETs is replotted. Interestingly, the $\sigma_{id}$ exhibits a similar quasi-exponential dependence on $V_{G}$ as the measured noise. The physical reason for the strong vertical field dependence is related to the exponential increase of the cross section with $\Delta E$, the shift of the first allowed band in the inversion layer [1].

Combining the foregoing elements, a straightforward expression for the gate voltage dependence of $S_{VG}$ can be proposed, which accounts for the QE, namely:

$$S_{VG} = \frac{q^2 kT\lambda N_a \sigma_{2D}}{WLC_{ox}} \frac{\sigma_{id}}{f}$$  \hspace{1cm} (2)

The inversion layer quantization is included in $C_{ox}$, which mainly relies on the shift of the charge centroid with $V_{G}$ and in the cross section ratio, which depends firstly on the increase of the band gap energy at the interface, with increasing $V_{G}$. The latter dependence is the most pronounced and will dominate at high(er) $V_{G}$, while the $C_{ox}$ effect is most important in weak inversion, where it sets the $N_{ox}$. Note that in Eq. (2) the only adjustable parameter is the oxide trap density. All other parameters can be self-consistently calculated from known or measured quantities. Furthermore, these measured quantities are independent of the noise experiment. Equation (2) is represented in Fig. 6, showing a good fit with the experimental noise data up to a gate bias of about 1 V. At higher gate biases, other effects contribute to the gate voltage dependence of the noise, like correlated mobility fluctuations (scattering) or series resistance.

Irrespective of the correctness of the proposed approach, it is clear that one has to account for the QE when modeling the 1/f noise in deep submicron MOSFETs. This is not only valid when dealing with the correlated number fluctuations approach, but also for the empirical Hooge law, which can be represented by:

$$\frac{S_{ID}}{I_{D}^2} = \frac{\alpha_{H}}{Nf}$$  \hspace{1cm} (3)

In Eq. (3), $N$ is the total number of carriers in the channel and $\alpha_{H}$ is the empirical Hooge parameter. $N$ can be derived from the inversion charge density $Q_{av}$ which is proportional to $C_{ox}(\lambda_{p}/\lambda_{n})$. Again, if no proper account is made for $C_{ox}$ and $\Delta E$, the extracted $\alpha_{H}$ will be erroneous. Comparing the classical with the quantum approach in Fig. 7 leads to the conclusion that

![Fig. 4. Calculated gate capacitance ratio for the n- and p-channel devices studied.](image)

![Fig. 5. Cross section ratio as a function of the gate bias for holes and electrons. In the same graph, the normalized input referred spectral density is plotted for comparison.](image)

![Fig. 6. Comparison of the model of Eq. (2) with the measured noise data for 0.13 μm n- and p-channel devices.](image)

5. Discussion
particularly in strong inversion, $N_{\text{class}}$ will be too high, leading to a higher $\alpha_t$ value. So in order to compare the $\alpha_t$ of different technology generations, the proper normalization needs to be done in Eq. (3), based on $N_{0t}$ for deep submicron devices.

In summary, it is clear that the introduction of QE in the noise modeling will open doors to a better (unified?) understanding of the underlying fundamental mechanisms.

6. Conclusion

Concluding, a new approach has been formulated to interpret the $1/f$ noise in deep submicron MOSFETs, taking account of inversion layer quantization. Although there is certainly still room for improvement, one can clearly state that neglecting QE leads to erroneous data for both $N_{0t}$ or $\alpha_t$. One should, therefore, take care in comparing noise parameters obtained on older technology generations, not suffering from QE, with the ones from deep submicron CMOS.

7. References