Numerical Investigation of Low Frequency Noise in MOSFETs with High-κ Gate Stacks

Yang Liu, Shuqing Cao, and Robert W. Dutton
Center for Integrated Systems, Stanford University, Stanford, CA, U.S.A. 94305-4075
E-mail: yangliu@gloworm.stanford.edu

***Abstract*—Two numerical models based on the impedance field method have been implemented to investigate the flicker noise in MOSFETs with high-κ gate stacks. The equivalent model uses approximate channel current noise source, while the physical model is based on the Langevin approach and accounts for the non-local carrier tunneling. The scaling impact on the flicker noise is investigated with the developed models. The validity of the models in the sub-threshold regime is examined. Comparison with experimental data indicates the importance of modeling the nonuniform trap energy distribution. The degradation of the flicker noise performance due to halo doping is also studied.**

Keywords-field effect transistor; noise; impedance field; HfO_2

I. INTRODUCTION

The timely deployment of alternative high-k gate dielectrics in MOS devices provides a promising yet challenging solution to the continuing aggressive scaling of CMOS technology. The high dielectric constants of those materials enable the reduction of the equivalent oxide thickness (EOT) while keeping the gate leakage current below tolerable limits. However, large trap densities are usually associated with high-k materials due to the poor interface quality. Since the flicker noise in MOS transistors is closely related to the surface trapping effect [1], degradation of the low frequency noise performance becomes a concern for high-k based devices [2]. In this work, we use the impedance field method (IFM) together with the local flicker noise source modeling to numerically investigate the low frequency noise property of devices with hafnium-based gate stacks.

The IFM approach was proposed in [3] and has been broadly used for multi-dimensional MOS thermal noise simulations. Simulations in this work are carried out using a general device simulator, PROPHET [4], where the IFM has been implemented to investigate thermal noise in highly scaled MOS transistors [5]. In this work, two approaches have been implemented to model the flicker noise in high-κ transistors. The equivalent model is based on the equivalent current density fluctuations, where the 1/f spectrum is described by the superposition of a set of approximate mono-polar generation-recombination (G-R) noise sources [6]. The physical model takes the Langevin approach. It solves the impedance field numerically for a system including the trap rate equation [7]. This method was recently implemented in a device simulator, FLOODS, to model carrier trapping induced low frequency noise in MOS transistors with SiO_2 gate stacks [8]. In order to account for the noise contributions induced by correlated mobility fluctuations, a unified model [9] has been adopted for both approaches in this work.

This work was partly supported by Stanford Center for Integrated Systems and MARCO MSD center.

The simulation results of the two models will be compared; and it will be shown that, unlike the physical model, the equivalent model becomes invalid in sub-threshold regime and requires explicit corrections according to Reimbold's theory [10]. Simulation results are correlated with experimental data, indicating the importance of accounting for the non-uniform energy distributions of the high-κ trap levels. The developed simulation capabilities are also used to reveal the impact of halo doping on flicker noise in highly scaled devices.

II. MODELING APPROACHES

A. Equivalent model for number fluctuations

A schematic plot of a typical high-k based n-type FET transistor is shown in Fig. 1(a). The trap density in the thin native SiO_2 interfacial layer is low. However, the high-κ layer above is usually found to be associated with high density of traps, where the inversion carriers can be trapped or de-trapped through tunneling. The band diagram showing this tunneling process is given in Fig. 1(b). Because the trapped electrons obey binomial statistics, only those traps with energy level near the channel quasi-Fermi level, E_F, have significant number fluctuations [1]. For a trap density energy distribution $N_T(E)$, the apparent trap density that contributes to the G-R noise is therefore given by $N_T=4k_BTN_T(E_F)$.

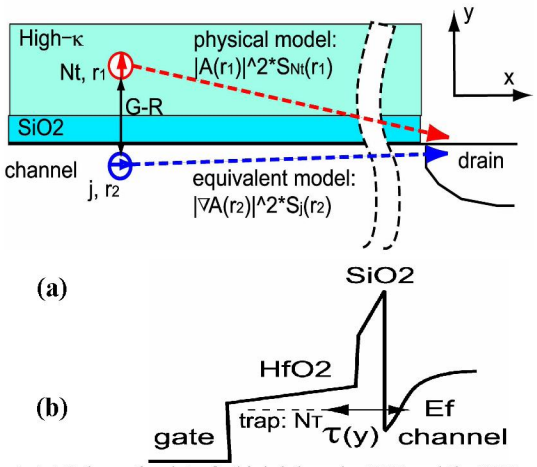

Fig. 1: (a) Schematic plot of a high-k based n-FET and the IFM-based equivalent and physical models.
(b) Band diagram along the gate stacks.

The scalar impedance field (IF), $A(\vec{r})$, is essentially a transfer function defined as the ratio of current perturbation at the terminal of interest (drain terminal in our case) to the

injected noise current at location r of the device. The vector IF is defined as the gradient of the scalar IF, $\nabla A(\vec{r})$. As illustrated in Fig. 1(a), the equivalent model assumes the trapped electron number fluctuation in the oxide is equal to that of the channel electrons: $\delta n = \delta n_t$. An equivalent microscopic current fluctuation at the corresponding channel position is then evaluated as $\delta \vec{j}_n = q\mu \vec{E} \cdot \delta n$. The total drain noise is then obtained by considering the propagation of the equivalent current noise sources along the channel:

$$S_{Id} = \int dv \cdot \left|\nabla A(\vec{r})\right|^2 S_j(f, \vec{r}),$$

where $S_j(f, \vec{r})$ is the power spectral density of the current noise source. The carrier number fluctuation is modeled by the superposition of a set of mono-polar G-R noise spectra with distributed time constants that produces 1/f behavior:

$$S_n \propto \int_0^{tox} N_T(y) \frac{\tau(y)}{1 + [2\pi f \tau(y)]^2} \cdot dy ,$$

where the distributed tunneling time constants are calculated using the WKB method based on the simulated potential profiles. The different affinities and tunneling masses in the SiO_2 and HfO_2 layers are considered in the calculations.

B. *Physical model for number fluctuations*

The calculation of the flicker noise based on the physical model is also illustrated in Fig. 1(a). Unlike the equivalent model, the microscopic noise source is modeled as to originate from the traps inside the oxide layers. The total drain noise is thus obtained as $S_{Id} = \int_{oxide} dv \cdot \left|A(\vec{r})\right|^2 S_{nt}(\vec{r})$, where $S_{nt}(\vec{r})$ is the power spectral density of the Langevin force associated with the trapping/de-trapping processes. In addition to the Poisson and carrier continuity equations, a rate equation for the trapped electron density is solved self-consistently:

$$dn_t / dt = G - R$$
$$= \left[(N_T - n_t)/\tau - n_t \exp(E_T - E_F / k_B T)/\tau\right]'$$

where the tunneling time, τ , is again calculated using the WKB method, and the trap energy level, E_T , is set equal to E_F for noise evaluations. This G-R term is also subtracted from the continuity equation for the channel carriers to ensure charge conservation. The trapped electrons also enter into the source term of the Poisson equation. It is noted that this approach is non-local from the simulation viewpoint. The grid points in the channel are coupled with those inside the oxides; and this is reflected by the corresponding non-zero off-diagonal entries in the Jacobian. The IFM is then extended to solve $A(\vec{r})$ taking into account this additional device equation.

The microscopic noise source is modeled as white G-R noise $S_{nt}(\vec{r}) = 2(G + R)$ according to [7]. In this physical model, the frequency dependence of the terminal noise is produced implicitly by the frequency dependence of the IF. In this model, two types of trap energy distributions are used: the uniform type and the exponential type. In the latter case, it is assumed that $N_T(E)$ decays exponentially away from the oxide band-edge E_c:

$$N_T(E) = \{1 + \exp[k_e(E_0 - E_c + E)]\} \cdot N_{T0} / 2 ,$$

where k_e, E_0, and N_{T0} are parameters to characterize the trap energy distribution profile.

C. *Correlated mobility fluctuations*

In order to explain the flicker noise behavior in MOS transistor devices, the carrier mobility fluctuation (Δμ theory) is proposed as an important physical mechanism along with the carrier number fluctuation (ΔN theory). The two theories usually predict different gate voltage dependence of flicker noise. In reality, correlations between these two mechanisms are often observed. Unified models (Δμ–ΔN theory) are therefore developed to conform to such observations. In this work, we adopt a unified model developed in [9], where the mobility fluctuation is modeled as the consequence of additional Coulomb scattering of channel carriers by the trapped charges. For both the equivalent and physical models, the local noise source is multiplied by a correction term as follows [9]:

$$\eta = \left(1 + \mu \sqrt{n_{2D}} / \mu_{c0}\right)^2 ,$$

where n_{2D} is the 2-D inversion carrier density and μ_{c0} is a fitting Coulomb scattering parameter. Throughout this work, a value of $1.5 \times 10^8 \, cm/Vs$ has been used for μ_{c0} [2].

III. SIMULATION RESULTS AND DISCUSSIONS

The impact of device scaling on the flicker noise behavior is examined using the equivalent model. NMOS devices with HfO_2 gate dielectric (EOT=1.3nm) and varying gate length from 30nm to 1.2 μm are simulated. Fig. 2(a) shows simulated drain noise as a function of frequency for devices with varying channel length. It can be seen that 1/f type noise is reproduced for frequencies up to 100 KHz. $1/f^2$ behavior is observed for higher frequencies, which can be explained by the fact that SiO_2 trap density is much smaller than that of HfO_2. The flicker noise and thermal noise meet at a corner frequency f_c, which increases with reducing channel length. To further illustrate this trend, we plotted flicker noise at 1Hz and thermal noise as functions of channel length in Fig. 2(b). It can be seen that, as the channel length decreases, the magnitude of flicker noise increases much faster than that of the thermal noise.

As described in the previous section, the flicker noise source in the physical model originates from the oxide traps. Fig. 3 shows the simulated profiles of flicker noise contributions inside the oxides at different frequencies. For this and the following simulations, the device EOT is set to 2.3nm (physical thickness 5.5nm) with 1nm interfacial layer to match

with the actual device parameters given in [2]. A moderate dielectric constant of $13.5\varepsilon_0$ is assumed for the HfO_2 layer, and an affinity difference of 1.2eV is used between the SiO_2 and HfO_2 layers. It is shown in Fig. 3 that, for a given frequency, the major noise contribution comes from traps at a certain depth into the oxides. As the frequency decreases, the peak of the distributed noise moves deeper into the oxides, because those traps correspond to longer tunneling time. The height of the noise peak increases by about 4 orders of magnitudes as the frequency decreases from 10KHz to 1Hz, leading to the 1/f type noise behavior. Such an observation agrees with that firstly reported in the work of [8].

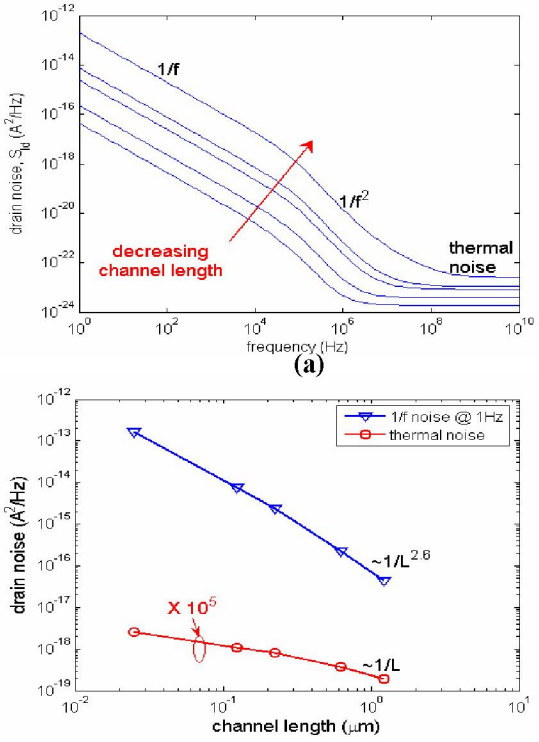

Fig. 2: (a) Simulated drain noise spectrum based on the equivalent model. Devices with varying gate length of 30, 130, 230, 630 and 1230nm are simulated, respectively. Vd=0.1V and Vg-V_T=0.2V. (b) Flicker noise at 1Hz and thermal noise as functions of channel length.

We compare the simulated normalized drain noise at 1Hz, S_{Id}/I_d^2, from both the equivalent and the physical models (Fig. 4). The device channel length is 1μm and the overdrive voltage varies from sub-threshold to strong inversion condition. The two models agree with each other in strong inversion regime. However, significant difference is observed at sub-threshold regime. As pointed out in the work of Reimbold [10], the relation $\delta n = \delta n_t$ no longer holds under sub-threshold condition. Instead, the correction due to the capacitances related to the charges at the gate and depletion regions needs to be considered. Therefore, S_{Id}/I_d^2 approaches to a saturated value in sub-threshold regime [10]. Without explicit inclusion of such a correction, the equivalent model significantly overestimates the drain noise near and under the threshold. On the other hand, the physical model is based on the direct evaluation of the impedance field of the complete system, with the electrostatics taken into account self-consistently. Hence, the saturation of the normalized drain noise is correctly reproduced in agreement with Reimbold's theory.

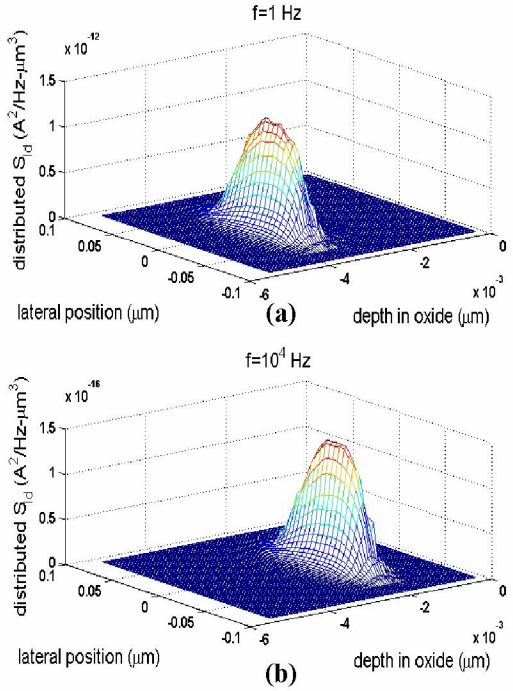

Fig. 3: Simulated profiles of distributed drain noise contributions at inside the gate oxides using the physical model. Two frequencies are simulated: (a) 1Hz; (b) 10KHz. The device gate length is 0.18μm. Vd=50mV and Vg-V_T=0.3V.

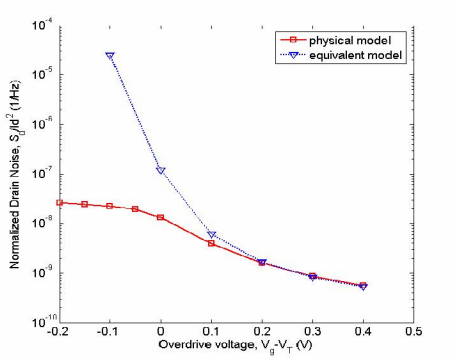

Fig. 4: Normalized drain noise as a function of overdrive voltage from simulations using both the equivalent and physical models. The device gate length is 1μm. Vd=50mV.

Simulated S_{Id}/I_d^2 based on the physical model is compared with measured data [2] at varying overdrive voltages. Simulation results are shown for both the uniform and exponential trap energy distribution models (Fig. 5). For the uniform model, a uniform trap spectral density of $3x10^{18}$/cm^{3}eV in the HfO_2 is assumed. For the exponential

model, values of N_{T0}, k_e and E_0 are assigned to $4x10^{17}/cm^3eV$, 6/eV and 3eV, respectively. The uniform model produces much stronger overdrive voltage dependence of the normalized drain current noise than the exponential model does. This is because when the gate bias increases, the channel quasi-Fermi level can access trap levels closer to the conduction band-edge, which possess higher apparent trap density according to the exponential model. The measured data show relatively weak bias dependence and are matched by the exponential trap energy distribution model.

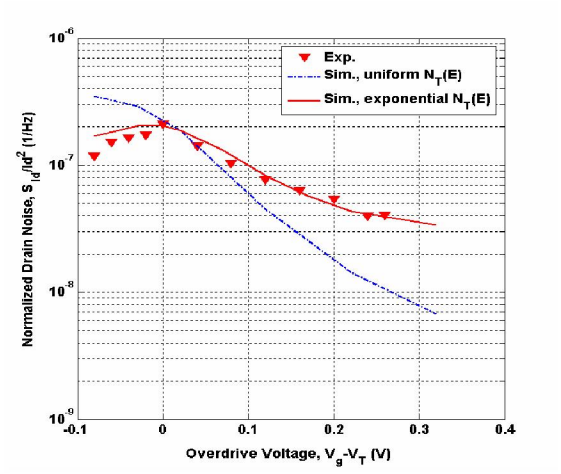

Fig. 5: Measured [2] and simulated normalized drain noise as a function of overdrive voltage. The simulations are based on the physical model with both the uniform and exponential trap energy distributions. The device gate length is 0.18μm and Vd=50mV.

In highly scaled MOS transistors, halo doping profiles are used to suppress the short channel effect. It has been experimentally observed that the presence of halo doping significantly degrades the flicker noise of short channel devices, mainly due to the nonuniform threshold voltage distribution along the channel [11]. In that work, an analytical model considering channel segments with different V_T's has also been developed. However, the sub-threshold correction according to Reimbold's theory has not been included, despite its importance considering the elevated V_T in the halo doping regions. In this work, a device with source/drain halo doping profiles ($1.8x10^{18}/cm^3$) is also simulated with the physical model. As shown in Fig. 6(a), the distributed noise contribution profile exhibits two evident peaks in the oxides, corresponding to the source/drain halo doping positions in the lateral direction. It indicates that the same amount of electron fluctuations cause greater current fluctuations in the halo regions than in the rest of the channel. This can be explained by the reduced inversion carrier density in the halo regions. In Fig. 6(b), the normalized flicker noise as a function of the overdrive voltage is simulated for devices with and without halo doping, respectively. The device with halo doping constantly exhibits higher noise level over the entire range of overdrive voltage.

IV. SUMMARY

Two IFM-based numerical models have been implemented to investigate the flicker noise performance in FETs with high-κ gate stacks. The equivalent model is based on approximate microscopic channel current noise source and becomes less accurate in sub-threshold regime. On the other hand, the physical model is applicable in all operation regimes. Comparison with experimental data indicates that it is important to consider the nonuniform trap energy distributions in the modeling. The negative impact of halo doping on the flicker noise is demonstrated with the aid of simulations.

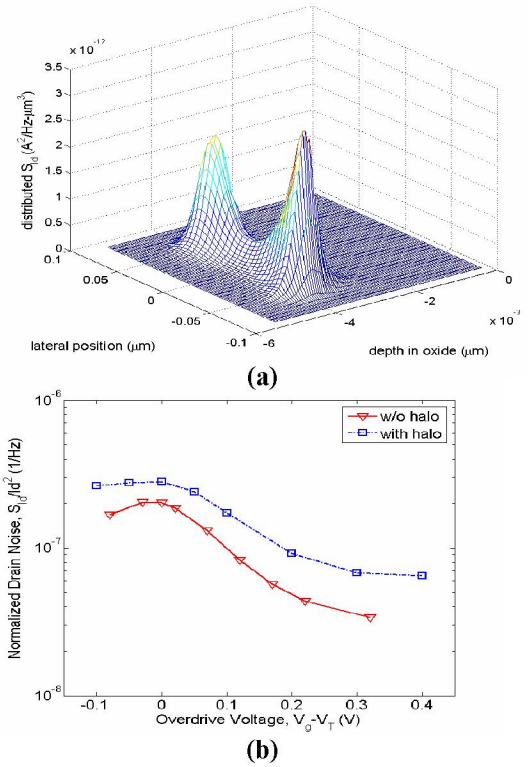

Fig. 6: (a) Simulated profile of noise contribution at 1Hz for a device with halo doping. (b) Normalized drain noise as a function of overdrive voltage for two devices with and without halo doping, respectively. The simulations are based on the physical model. The device gate length is 0.18μm and Vd=50mV.

REFERENCES

[1] A. van der Ziel, Noise in Solid State Dev. and Circuits, Wiley, NY, 1986

[2] B. Min et al., IEEE Trans. Elec. Dev. Vol. 51, pp.1679-1687, 2004

[3] W. Shockley et al, in Quantum Theory of Atoms, Molecules and the Solid-State, Academic, NY, 1966, pp. 537-563

[4] PROPHET Technical Information. [Online]. Available: http://www-tcad.stanford.edu/~prophet

[5] T.-Y. Oh, C. Jungemann, and R. Dutton, Proc. SISPAD, pp.87-90, 2003

[6] J.P. Nougier, IEEE Trans. Elec. Dev., Vol. 41, pp.2034-2049, 1994

[7] F. Bonani and G. Ghione, Solid-State Elec., Vol. 43, pp.285-295, 1999

[8] F.-C. Hou, G. Bosman, and M. E. Law, IEEE Trans. Elec. Dev.., Vol. 50, pp. 846-852, 2003

[9] E. Vandamme and L. Vandamme, IEEE Trans. Elec. Dev., Vol. 47, 2146-2152, 2000

[10] G. Reimbold, IEEE Trans. Elec. Dev., Vol. 9, pp. 1190-1198, 1984

[11] J.-W. Wu et al, IEEE Trans. Elec. Dev., Vol. 51, pp. 1262-1266, 2004