Effects of Scaling on Modeling of Analog RF MOS Devices

Y. Liu, S. Cao, T.Y. Oh[1], B. Wu, O. Tornblad[2] and R.W. Dutton

Center for Integrated Systems
Stanford University, Stanford, CA, USA, yangliu@gloworm.stanford.edu
[1]LG Electronics, Seoul, Korea
[2]Infineon Technologies, Morgan Hill, CA, USA

ABSTRACT

This paper uses advanced TCAD tools to investigate the scaling effects on noise and linearity performance of analog RF MOS devices, as well as the implications on modeling those effects at the compact level. Impedance field method and hydrodynamic transport model are used to study the thermal noise behavior in aggressively scaled MOS devices. A flicker noise model is implemented; low frequency noise performance of devices based on alternative high-k gate dielectrics is also evaluated. The large signal distortion analysis is conducted based on the harmonic balance technique. The effects of doping profiles and their changes with technology scaling on the distortion performance are examined.

Keywords: Impedance field method, thermal noise, hydrodynamic model, flicker noise, harmonic balance, large signal distortion

1 INTRODUCTION

Scaling of MOS devices for digital applications continues at an impressive rate, yielding ongoing improvements in density in spite of limited improvements in drive current and virtually constant supply voltages. Device performance for analog and particularly wireless (RF) applications is reaching limits in terms of noise and distortion, owing to technology constraints imposed primarily by digital scaling. In this work, we aim to explore some of these scaling effects on analog RF performance of MOS transistors. The analysis is based on detailed device-level simulations using advanced TCAD tools. The implications of these simulations and scaling effects in terms of compact models are also discussed.

It has been shown that device-level analysis of drain and gate thermal noise provides both insight and modeling guidance in support of power-constrained optimization of low-noise RF circuits [1]. From the TCAD perspective, the thermal noise analysis uses the impedance field method [2] and hydrodynamic transport modeling of short-channel effects in scaled MOS devices [3]. The IFM method, implemented numerically in device simulations [4], has an equivalent circuit-level interpretation that can help to explain the physical results as well as provide useful guidance in developing compact models. It is demonstrated from device-level simulations that the noise partitioning along the channel changes with scaling of gate length. Large excess gate and drain noise figures are predicted for aggressively scaled devices based on the HD model.

Low frequency noise performance is also impacted by the aggressive scaling of device dimensions. The use of alternative gate dielectrics with high dielectric constants is proposed to avoid severe gate leakage when the equivalent oxide thickness (EOT) is below 2nm. Large trap densities are usually associated with those high-k materials, which may lead to significant degradation of the flicker noise performance [5]. In this work, the simulated impedance field is combined with local flicker noise source modeling to evaluate the low frequency noise performance of devices with hafnium-based gate stacks. The developed numerical model is used to investigate the effect of gate length scaling on flicker noise and corner frequency.

Signal distortion is a key limitation in analog RF systems, resulting from device nonlinearities. Device-level harmonic balance (HB) analysis has been shown to provide an excellent TCAD-based tool for linearity analysis in RF LDMOS devices [6]. HB simulation capabilities [7,8] will be used in this work to explore both the macro- and microscopic implications of device scaling on distortion behavior. Results presented here show effects of doping profiles and their changes with technology scaling on the distortion performance of scaled MOS devices. These results provide useful information concerning the impact of technology scaling on device performance as well as macro-level information that is needed for compact modeling at the behavior level.

2 NOISE SIMULATION AND ANALYSIS

2.1 Impedance Field Method (IFM) and Hydrodynamic (HD) Model

The noise simulations presented in this work are based on the impedance field method and hydrodynamic transport model implemented in a general semiconductor device simulator, PROPHET. The IFM approach was originally developed by Shockley et al. [2]; and it has been adopted for multi-dimensional MOS transistor RF noise simulations based on drift-diffusion (DD) [9] or hydrodynamic (HD) [3] transport models. The general idea is to find the noise transfer function from an arbitrary point in the device to the

terminal where the noise measurement is carried out. This approach is schematically illustrated in Fig. 1. The impedance field $A_i(\vec{r})$ is defined as the ratio of current fluctuation at the i-th electrode and the injected current at position $\vec{r}$ in the device. The local noise source generated within a small segment $[\vec{r}, \vec{r}+d\vec{r}]$ is equivalently modeled by a small-signal current injected at $\vec{r}$ and pulled out at $\vec{r}+d\vec{r}$. Therefore, the contribution of this segment to the noise at the i-th electrode is $\left|\nabla A_i(\vec{r})\right|^2 \cdot S_{in} \cdot d\vec{r}$, where S_{in} is the power spectral density of the local noise source. For an ideal long channel MOS transistor, the generation of impedance field can be represented using the equivalent circuit model given in Fig. 2. Fig. 2(a) illustrates the partition of a saturated transistor into a linear transistor and a saturated transistor. The partition is at the point of the channel where the current is injected. The small signal circuit of the two transistors in common gate configuration is also shown. By writing down their MOS current equations, one can show that the impedance field for drain current noise is given by [4]:

$$A_d = R_s / (R_s + 1/g_m),$$

which increases linearly from 0 to 1 in the channel up to the pinch-off point. Its gradient $\left|\nabla A_d\right|$ is therefore approximately proportional to the local AC resistance. The impedance field for gate current noise is generated through the capacitive coupling between the channel and the gate, as represented by the distributed RC network in Fig. 2(b). Its profile usually has a maximum in the middle of the channel. For deep sub-micron transistors, second order effects become more significant and 2D simulations are usually necessary.

As transistor channel length reduces to deep sub-micron regime, carrier transport tends to be more ballistic, which generally requires the consideration of higher order moments in the Boltzmann Transport Equation (BTE) for accurate modeling. The HD model implemented in PROPHET follows that developed in [10]. In addition to the Shockley semiconductor equations, two energy balance equations are also solved, where the carrier velocity is controlled by the energy relaxation rate rather than limited by saturation velocity. The thermal diffusion current component is also added to the drift-diffusion current. An important correction for HD-based noise simulation is on the evaluation of the local noise source. In the drift-diffusion model, the local noise source can be modeled under the near-equilibrium assumption as $S_{in} = 4qn\mu(\vec{E})k_B T_0$, where $\mu(\vec{E})$ is field dependent mobility and T_0 is the lattice temperature. In the hydrodynamic model, the near-equilibrium assumption is no longer valid; and an empirical correction has been adopted for the local noise source based on Monte Carlo simulation results [11].

2.2 Scaling Effects on Thermal Noise

We simulate aggressively scaled n-type MOS transistors with the device structure and doping profiles similar to those proposed in [12]. Such devices exhibit good electrostatic control and allow scaling of gate length down to 15nm. We firstly simulate impedance field distributions for an NMOS with 30nm gate length based on both the DD and HD models. Both the impedance field A_d and its gradient $\left|\nabla A_d\right|^2$ along the channel are shown in Fig. 3. For the DD case, we can see that A_d increases monotonically from source to drain with an abrupt slope change at the source side metallurgical junction. This translates to the small peak in the profile of $\left|\nabla A_d\right|^2$, indicating the effect of the source to channel resistance. The HD simulation gives a quite different impedance field profile. The peak of $\left|\nabla A_d\right|^2$ at the source side junction is enhanced. This indicates the ballistic carrier transport makes the source side resistance more significant compared to the channel resistance. An overshoot of A_d is observed near the drain side, which leads to the shift of the central peak of $\left|\nabla A_d\right|^2$ to the source side and an additional small peak in the drain side. This is believed to be due to carrier velocity effects. Fig. 4 shows the simulated impedance field for gate noise A_g and its gradient $\left|A_g\right|^2$. Both DD and HD results show a maximum of A_g in the channel. The maximum is shifted further to the source side in the HD case, which is also attributed to the increasing of the relative importance of the source side resistance.

The 2D spatial distribution of the noise contribution to the drain and gate, as determined by $\left|\nabla A_i(\vec{r})\right|^2 \cdot S_{in}$, are plotted in Fig. 5 and Fig. 6, respectively. Three cases are simulated for comparison purposes: DD with velocity saturation model, DD without velocity saturation model and HD model. The difference between the HD model and the DD model with velocity saturation is evident, which is consistent with that observed in Fig. 3 and Fig. 4. The impedance profiles from the DD model without velocity saturation show much more similarity with that of the HD case. This is because the carrier velocity overshoot effect is artificially achieved in the DD simulations by removing the velocity saturation model.

In order to investigate the scaling effects, drain and gate thermal noise figures have been simulated for devices with the gate length ranging from 30nm to 1.2 μm. The excess drain noise parameter γ is given by [13]

$$\gamma = S_{id} / 4k_B T g_{d0},$$

where S_{id} is simulated drain noise power spectral density and g_{d0} is the drain to source conductance at zero drain bias. The excess gate noise parameter δ is given by [13]

$$\delta = \frac{5 g_{d0} S_{ig}}{4 k_B T \omega^2 C_{gs}^2},$$

where S_{ig} is simulated gate noise power spectral density and C_{gs} is the source to gate capacitance. The analytical results of γ and δ for long channel devices under saturation are 0.66 and 4/3, respectively. The simulated results of γ and δ from both DD and HD simulations are plotted in Fig. 7(a) and (b). It can be seen that DD model only predicts a very moderate increase of the two parameters as the channel length decreases. However, the HD model gives better account of the hot carrier effects and predicts much higher values of γ and δ for devices with aggressively scaled channel length.

2.3 Scaling Effects on Low Frequency Noise

In order to maintain tolerable gate leakage current while keeping the reduction of the equivalent oxide thickness, alternative gate insulating materials with high dielectric constants have been under extensive study. One promising candidate is Hf-based materials, including HfO_2 and HfSiON. However, large trap densities are usually associated with high-k materials due to the poor interface quality or, alternatively, degraded bulk properties of the high-k layer. As a result, low frequency noise performance in high-k based devices may be significantly degraded. To investigate the scaling effects on low frequency noise performance, we have incorporated flicker noise modeling in our simulations. Similar to the treatment of thermal noise, the contribution of the flicker noise from a channel location to drain is given by $\left|A_d(\vec{r})\right|^2 S_{in}^f$. The current noise source S_{in}^f is related to the carrier density fluctuation noise source S_n^f by [14]

$$S_{in}^f = q^2 \mu^2 E^2 (1 + \mu \sqrt{n_{2D}} / \mu_{c0})^2 S_n^f,$$

where n_{2D} is the 2D carrier density and μ_{c0} is a scattering parameter to account for the mobility fluctuation induced by the carrier number fluctuation. S_{in}^f is computed as the superposition of trapping-detrapping noise for spatially distributed trap centers in the gate stacks [13]. A schematic plot is given in the inset of Fig. 9, describing the trapping/de-trapping processes by tunneling between the channel and the trap centers with varying depth into the gate oxide. The tunneling rates are computed using the WKB method. In the calculations, the different barrier height, dielectric constants and tunneling effective mass of the high-k layer and the interfacial SiO_2 layer are considered. Only those traps with energy close to the channel quasi-Fermi level can contribute to the noise fluctuations [13].

We simulated NMOS devices with HfO_2 gate dielectric and varying gate length from 30nm to 1.2μm. The HfO_2 layer thickness is 3.5nm and the interfacial SiO_2 layer is 0.8nm; the EOT is 1.3nm. We assumed a uniform trap density of 5e17$cm^{-3}eV^{-1}$ in the SiO_2 and 1.5e19$cm^{-3}eV^{-1}$ in the HfO_2. A value of 1.5e8 cm/Vs is used for μ_{c0} according to [5]. Fig. 9 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 a function of channel length in Fig. 10. It can be seen that, as the channel length reduces, the magnitude of flicker noise increases much faster than that of the thermal noise.

3 DISTORTION ANALYSIS BASED ON HARMONIC BALANCE METHOD

The harmonic balance (HB) method is a device simulation technique capable of performing large-signal analysis in the steady state. In this method, time-invariant nonlinear systems with periodic responses are transformed and solved for a given number of harmonic components in the frequency domain. In this section, we simulate the distortion behavior of aggressively scaled MOS transistors with 40nm nominal gate length. We carry out one-tone HB simulations using the HB simulation capabilities implemented in the PISCES II-HB device simulator [7].

In order to investigate the impact due to doping profiles in scaled devices, we simulate three device structures as depicted in Fig. 11. Case A is a MOS transistor with 40nm gate length and two halos at both source and drain. Case B is a reverse scaled version that has a gate length of 68nm. Case C is the same as the case B, except that the drain side halo doping is removed. The figure of merit, HM3 dBc, is defined as the difference between the fundamental output power component HM1 and the third-order component HM3 at the same input power level.

The simulated HM3 in dBc as a function of the output power level are plotted for the three devices in Fig. 12. Simulation results for two operation frequencies, f=2GHz and 100MHz, are shown in Fig. 12(a) and (b), respectively. It can be seen that the device with longer channel length and two halos (Case B) exhibits the lowest HM3, indicating the best linearity performance. This trend is more prominent for the case of 2GHz operation frequency,

strongly suggesting the influence of the intrinsic capacitance and associated nonlinearities.

The critical feedback capacitance, $C_{gd,}$ and the source to drain capacitance, Csd, are simulated at different drain biases for the three devices, as shown in Fig. 13. C_{gd} is usually critical for linearity, both in absolute value and in nonlinearity. In Fig. 13(a), it can be seen that in Cases A and B both devices exhibit smaller C_{gd} compared to that of Case C, which is directly linked to the presence of the drain side halo doping in those two devices. In Fig. 13(b), it is observed that Case B also exhibits the smallest magnitude and best linearity in C_{sd} among the three devices, mainly due to the largest effective channel length in that device. Combining these observations, it is suggested that the linearity performance for the longer device with two halos is affected the least by the nonlinear intrinsic capacitances.

4 SUMMARY

We have numerically investigated the thermal noise performance in deep sub-micron MOS transistors, by using advanced IFM and HD simulation capabilities. The equivalent circuit interpretation of the impedance field offers good insight into the changes of noise contribution along the channel as device gate length is aggressively scaled. Large excess drain and gate noise parameters are predicted by the HD transport model for devices with very short channel lengths. We have also implemented a flicker noise model and investigated the scaling effects on the low frequency noise performance of hafnium-based devices. The HB-based distortion analysis reveals the impact of halo doping and channel length on the linearity performance of scaled MOS devices. The relation to the nonlinearity of intrinsic capacitances is also examined.

The support of MARCO MSD is gratefully acknowledged.

REFERENCES

[1] J.-S. Goo et al. Proc. SISPAD, pp.269-272, 2000
[2] W. Shockley, J. A. Copeland, and R.P. James, in Quantum Theory of Atoms, Molecules and the Solid-State, Academic, NY, 1966, pp. 537-563
[3] C. Jungemann et al., IEDM pp.871-4, 2003
[4] T.-Y. Oh, "AC and noise analysis of deep-submicron MOSFETs," Ph.D. Dissertation, Stanford Univ., 2004
[5] B. Min et al., IEEE T. Electron Dev. Vol. 51, pp.1679-1687, 2004
[6] O. Tornblad, et al., Proc. SISPAD, pp. 243-6, 2005
[7] B. Troyanovsky, "Frequency Domain Algorithms for Simulating Large Signal Distortion in Semiconductor Devices", PhD Dissertation, Stanford University, 1997
[8] B. Troyanovsky et al., Proc. SASIMI '96 1996.
[9] S. Donati et al., Tech. Dig. IEDM pp.81-84 1998
[10] B. Meinerzhagen and W.L. Engl, IEEE T. Electron Dev., vol. 35, pp. 689-697, 1998
[11] T.-Y. Oh, C. Jungemann, R.W. Dutton, Proc. SISPAD 2003, pp.87-90
[12] D. Connelly et al., IEEE T. Electron Dev. Vol. 53, pp. 146-152, 2006
[13] A. van der Ziel, Noise in Solid State Devices and Circuits, Wiley, NY, 1986
[14] E.P. Vandamme and L.K.J. Vandamme, IEEE T. Electron Dev. Vol. 47, pp.2146-2152, 2000

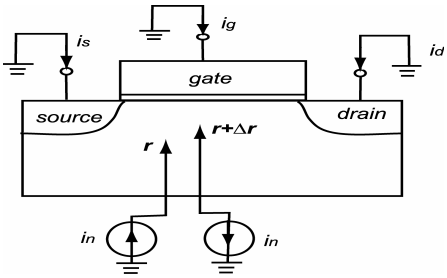


Fig. 1: Schematic plot of noise modeling in MOS transistors using the Impedance Field Method.

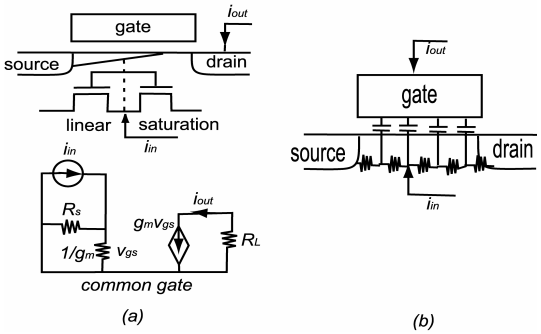


Fig. 2: Equivalent circuit models of the impedance field for the drain (a) and gate (b) current noise.

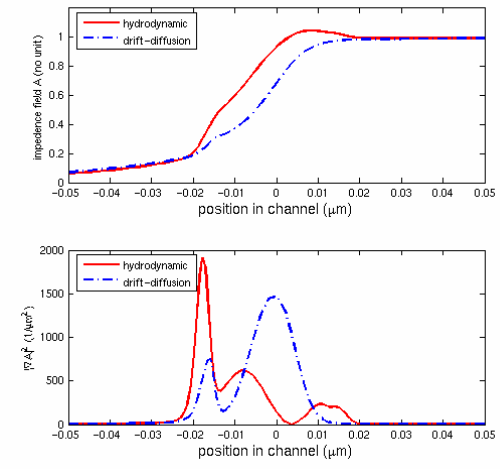


Fig. 3: Simulated drain impedance field A_d and its gradient $|\nabla A_d|^2$ along the channel. Both DD and HD simulation results are shown. The gate length is 30nm. Vd=1.0V and Vg=0.9V.

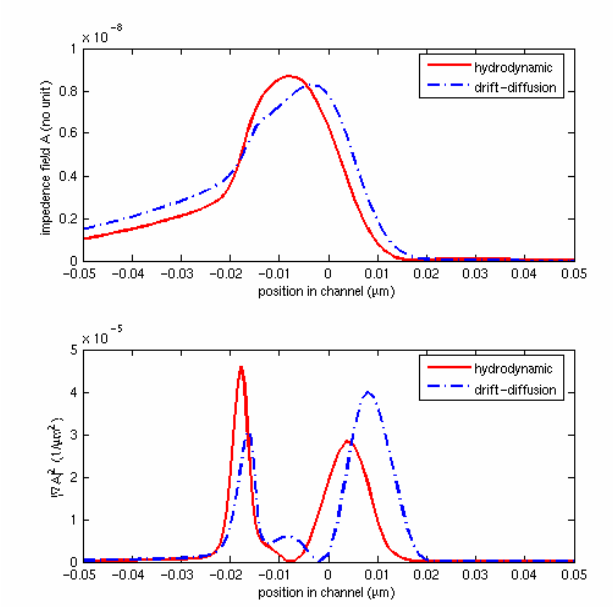


Fig. 4: Simulated gate impedance field A_g and its gradient $\left|\nabla A_g\right|^2$ along the channel. Both DD and HD simulation results are shown. The gate length is 30nm. Vd=1.0V and Vg=0.9V.

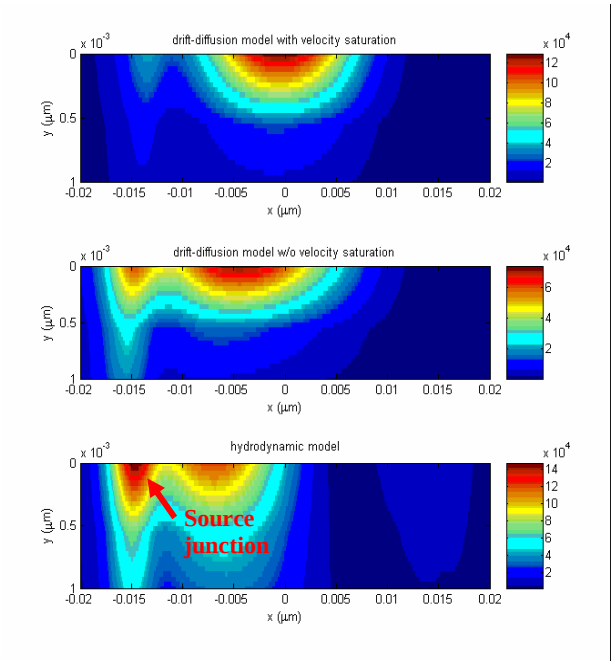


Fig. 5: Simulated noise contribution to the drain noise. Three cases are simulated (from top to bottom): DD with velocity saturation, DD without velocity saturation and HD. The gate length is 30nm. Vd=1.0V and Vg=0.9V.

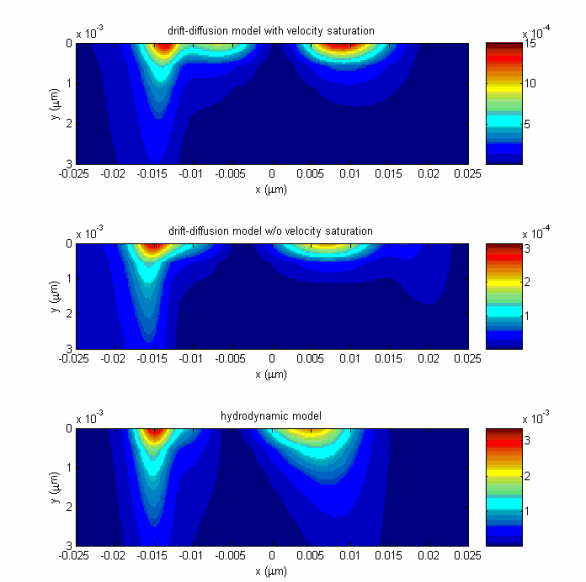
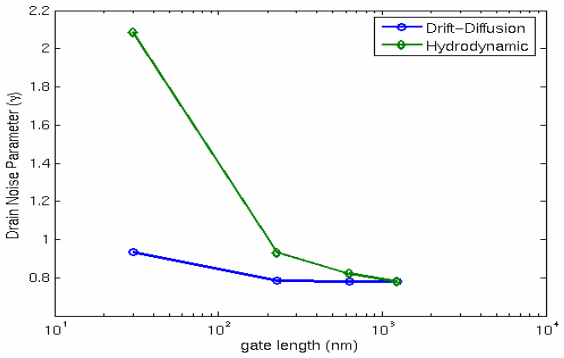
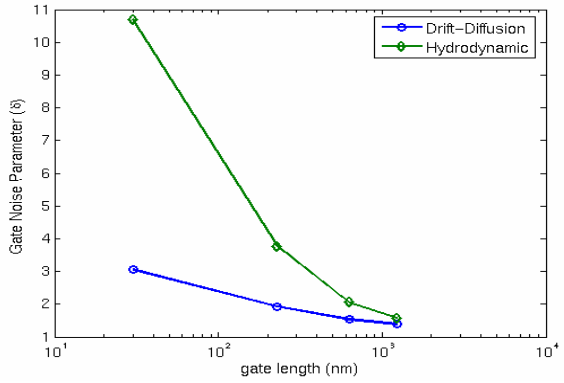


Fig. 6: Simulated noise contribution to the gate noise. Three cases are simulated (from top to bottom): DD with velocity saturation, DD without velocity saturation and HD. The gate length is 30nm. Vd=1.0V and Vg=0.9V.



(a)



(b)

Fig. 7: Simulated excess noise parameters for devices with varying gate length. Both HD and DD simulation results are shown. Vd=1.0V and Vg=0.9V. (a) drain noise parameter γ; (b) gate noise parameter δ.

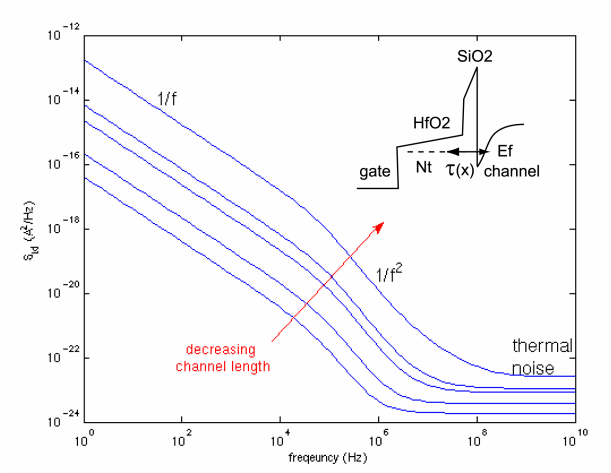


Fig. 9: Simulated low frequency noise at drain as a function of frequency. Devices with varying gate length are simulated: 30, 130, 230, 630 and 1230 nm, respectively. Vd=0.1V and Vg=0.4V.

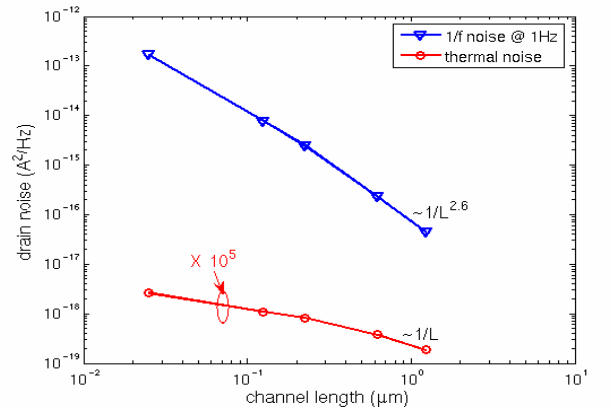


Fig. 10: Simulated flicker noise at 1Hz and thermal noise as functions of channel length. Vd=0.1V and Vg=0.4V.

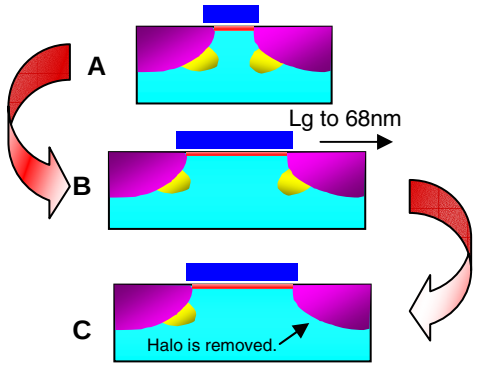


Fig. 11: Schematic plot of three devices used in HB simulations. A) L_g=40nm, 2 halos. B) L_g=68nm, 2 halos. C) L_g=68nm, 1 halo.

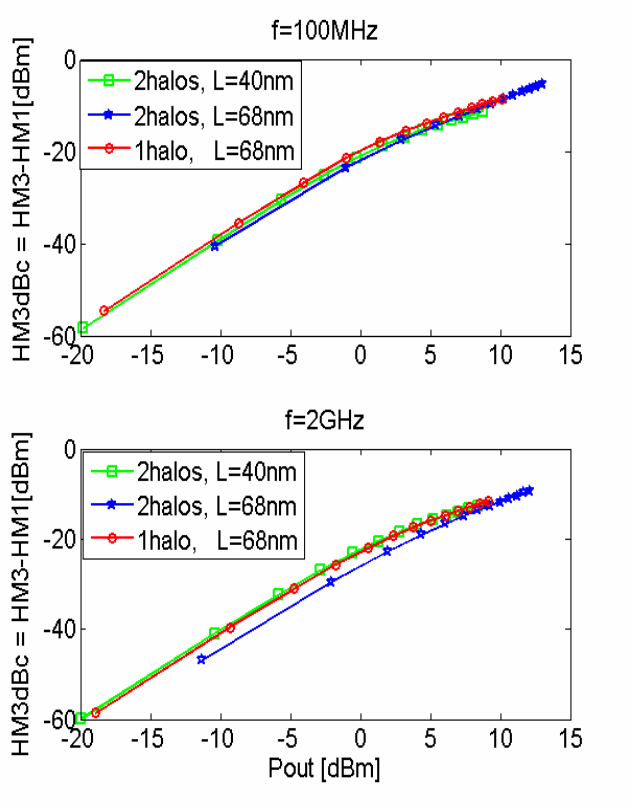


Fig. 12: HM3 from HB one-tone simulations for the three devices in Fig. 11. DC bias Vg=1.2V and Vd=1V. a) f=2GHz. b) f=100MHz

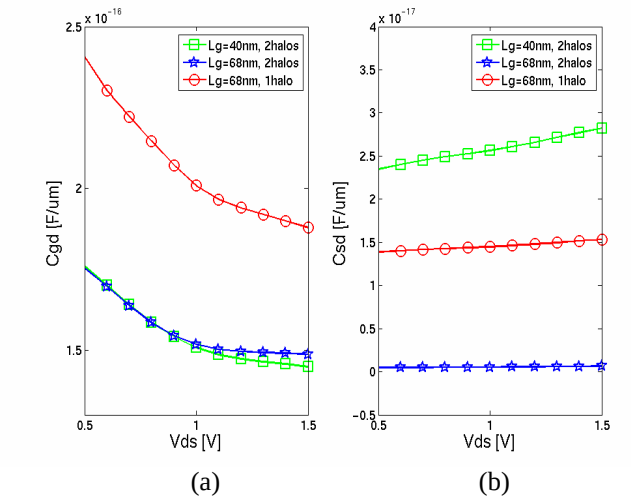


Fig. 13: Small signal C-V simulation results for the three devices in Fig. 11. a) Cgd vs. Vds; b) Csd vs. Vds.