Narrow Width Effects on High Frequency Performance and RF Noise of Sub-40nm Multi-finger nMOSFETs

Kuo-Liang Yeh, Chih-Shiang Chang, and Jyh-Chyurn Guo
Institute of Electronics Engineering, National Chiao Tung University, Hsinchu, Taiwan
Tel: +886-3-5131368, Fax: +886-3-5724361, E-mail: jeguo@mail.nctu.edu.tw

Abstract
The impact of narrow width effects on high frequency performance parameters like $f_T$, $f_{MAX}$, and RF noise in 35 nm multi-finger n-MOSFETs is investigated in this paper. Multi-OD devices with reduced width under fixed finger number ($N_F$) lead to higher $R_g$ and suffer the penalty in $f_T$, $f_{MAX}$, and $N_F_{min}$. On the other hand, narrow-OD MOSFET with larger $N_F$ can yield lower $R_g$ and higher $f_{MAX}$. However, these narrow-OD devices even with lower $R_g$ suffer lower $f_T$ and higher $N_F_{min}$. The mechanisms responsible for narrow width effects on $f_T$, $f_{MAX}$, and noise parameters will be presented to offer an important guideline of MOSFET layout for RF circuits design using nanoscale CMOS technology.

I. Introduction
Nanoscale CMOS MOSFETs with multi-finger layout have been extensively used for higher $f_T$ and $f_{MAX}$, and lower RF noise driven by gate length scaling and gate resistance ($R_g$) reduction from multi-finger structure [1-3]. Unfortunately, the continuous reduction of finger width ($W_F$) and increase of finger number ($N_F$) for smaller $R_g$ may lead to the penalty of lower transconductance ($g_{m}$) and larger parasitic capacitances. The former one comes from stress induced mobility degradation and the latter one stems from gate related fringing capacitances [4-5]. Both can not be scalable with device scaling and the impact may dominate high frequency characteristics in nanoscale devices. The potential impact from parasitic capacitances and the trade-off with $R_g$ becomes a critical factor governing the specified RF performance parameters and has to be considered seriously in devices layout for RF circuit design.

II. Experimental
The multi-finger n-MOSFETs were fabricated in 65nm CMOS process with 35nm physical gate length and the total gate width fixed at 64 μm ($W_{OD}=W_{XN}=64μm$). Fig.1(a)-(c) illustrate multi-finger MOSFET layouts, namely standard, narrow-OD, and multi-OD devices in which $σ$ and $σ_i$ denote the longitudinal and transverse stresses introduced from STI. S-parameters were measured by Agilent network analyzer E8364B up to 40GHz. In particular, open and short deembedding to the bottom metal, i.e. M1, were performed to remove the parasitic capacitances from the pads as well as interconnection lines [5]. Four noise parameters, such as $N_{F_{min}}$, $R_n$, $Re(Y_{opt})$, and $Im(Y_{opt})$ can be measured by using ATN-NP5B from 1GHz up to 18GHz.

III. Results and Discussion
Fig. 2(a) demonstrates $g_{m}$ versus $V_{GS}$ ($V_{GS}-V_T$) in saturation region measured from narrow-OD and standard nMOS. This monotonic degradation of $g_{m}$ with $W_F$ scaling suggests that the increase of compressive $σ$ is the dominant factor responsible for mobility degradation and the resulted $g_{m}$ reduction [6]. As for multi-OD nMOS shown in Fig. 2(b), the $g_{m}$ of OD4 and OD8 are degraded by 11.2% and 18.9%, as compared to OD1, but the degradation becomes smaller to 8.5% for OD16. According to our recent work [7], the increase of effective width ($W_{eff}$) from STI top corner rounding induced $ΔW$ can compensate mobility degradation. Note that $ΔW$ of nMOS in this process is 38.7 nm. With the increase of $N_{OD}$, $ΔW$ effect will be enhanced and may dominate STI $σ$ effect, which explain the increase of $g_{m}$ when scaling $W_{OD}$ from 0.25μm for OD8 to 0.125μm for OD16. Fig. 3(a) reveals that $W_F$ scaling in narrow-OD nMOS leads to a monotonic $f_T$ degradation. An analytical model given by (1) [8] suggests that $f_T$ degradation can be originated from $g_{m}$ and/or increase of $C_{gg}$. For narrow-OD nMOS, the smallest $g_{m}$ appearing in W05N128 (Fig.2(a)) is considered as one of major factors responsible for the worst $f_T$. Furthermore, the measured $C_{gg}$ (Fig.3(b)) indicates 8.3% increase of $C_{gg}$ in W05N128. The combined effect from lower $g_{m}$ and larger $C_{gg}$ can explain $f_T$ degradation in narrow-OD devices [5]. As for multi-OD nMOS, Fig.4(a) again indicates a monotonic degradation of $f_T$ with $W_{OD}$ scaling. Obviously, OD16 with the smallest width (0.125μm) suffers the lowest $f_T$. As shown in Fig.4(b), OD16 reveals substantially larger $C_{gg}$ compared to OD1. Although OD16 yields higher $g_{m}$ than OD4 and OD8, due to $ΔW$ effect (Fig.2(b)), the much larger $C_{gg}$ offsets the $g_{m}$ increase and leads to $f_T$ degradation.

Fig. 5(a) presents an increase of $f_{MAX}$ with $W_F$ scaling and the highest $f_{MAX}$ achieved by W05N128. The $f_{MAX}$ can be calculated by (2) [8], which indicates that the higher $f_T$ and lower $R_g$ can enhance $f_{MAX}$. Referring to Fig.3(a), W05N128 suffers the lowest $f_T$. However, the smaller $W_F$ and larger $N_F$ in narrow-OD devices can reduce $R_g$ as shown in Fig.5(b). Almost 50% lower $R_g$ realized by W05N128 can over-compensate $f_T$ degradation and contribute to higher $f_{MAX}$. On the other hand, multi-OD nMOS shown in Fig.6(a) indicates a monotonic degradation of $f_{MAX}$ with $W_{OD}$ scaling. Referring to Fig.4(a), the larger $N_{OD}$, i.e. the smaller $W_{OD}$ indeed leads to $f_T$ degradation. As for $R_g$ shown in Fig.6(b), the smaller $W_{OD}$ (larger $N_{OD}$ and fixed $N_F$) suffers the higher $R_g$ and OD16 reveals around 20~25% higher $R_g$. According to (2), the lower $f_T$ and larger $R_g$ in multi-OD devices are two key factors responsible for $f_{MAX}$ degradation with $W_{OD}$ scaling.

Fig. 7(a)-(d) present $N_{F_{min}}$, $R_n$, $Re(Y_{opt})$, and $Im(Y_{opt})$ measured from the standard and narrow-OD nMOS (1~18GHz). W05N128 with the smallest $W_F$ and $R_g$ suffers 0.2~0.5dB higher $N_{F_{min}}$ in 9~18GHz. Fig.7(c) and (d) reveal significant increase of $Re(Y_{opt})$ and $Im(Y_{opt})$ (absolute value) in W05N128 at higher frequencies, above 9GHz. According to (3)-(4) [8], the increase of $R_n$ or $Re(Y_{opt})$ will result in higher $N_{F_{min}}$ and the calculated $N_{F_{min}}$ can fit measured data in terms of frequency and layout dependence as shown in Fig.7(a). This proven model combined with Fig.7(b) and (c) for measured $R_n$ and $Re(Y_{opt})$ indicates that the increase of $Re(Y_{opt})$ is the primary factor responsible for higher $N_{F_{min}}$ in narrow-OD nMOS. According to (5) for $R_n$, the benefit of smaller $R_n$ in narrow-OD device happens to be cancelled out by the increase of the second term due to lower $g_{m}$ from compressive $σ$. As for multi-OD nMOS shown in Fig.8(a)-(d), OD16 reveals the largest value in $Re(Y_{opt})$ and $Im(Y_{opt})$, leaving $R_n$ as the exception. It is interesting to note that OD16 suffers the largest $R_n$ (Fig.6(b)) but achieves $R_n$ lower than OD8. Referring to (5), the higher $g_{m}$ can help reduce $R_n$, due to smaller $g_{m}/g_{m}'$, and reverse the correlation between $R_n$ and $R_g$. It explains why OD16 has the larger $R_n$ but smaller $R_g$ than OD8. The counterbalance between $Re(Y_{opt})$ and $R_n$ results in comparable $N_{F_{min}}$ between OD16 and OD8. All of multi-OD nMOS suffer higher $N_{F_{min}}$ than OD1.
\[ f_t = \frac{g_m}{2 \pi \sqrt{C_{Gd} \cdot C_{Gg}}} \]  
\[ f_{\text{max}} = \frac{g_m}{2 \pi \sqrt{2 R_g (g_{\text{on}} + 2 R_g C_{Gd}) + g_{\text{on}} (R_g + R_s)}} \]  
\[ f_{\text{max}} = 1 + 2 R_g \left| Y_{\text{opt}} \right|^2 + R_s \left| Y_{\text{opt}} \right| \]  
\[ N_{\text{FB}} = 10 \log \left( \frac{R_n}{g_m} \right) \]  
\[ R_n = R_g + \frac{g_{\text{on}}}{g_m} (\gamma > 1 \text{ for short channel devices}) \]

**Acknowledgement**

This work is supported in part by the National Science Council under Grant NSC98-2221-E-009-166-MY3. Besides, the authors acknowledge the support from NDL for noise measurement and CiC for test chip and device fabrication.

**References**


**Fig. 1** Multi-finger MOSFET layouts with various \( W_F \) and \( N_F \) and STI stresses \( \sigma \) and \( \tau \). (a) standard device: \( W_0 x N_0 =2 \mu m x 32 \) (W2N32), (b) narrow-OD devices: \( W_0 x N_0 =1 \mu m x 64 \) (W1N64), \( 0.5 \mu m x 128 \) (W05N128), and (c) multi-OD devices \( W_0 x O_D x N_{O_D} \): \( W_2 = 2 \mu m : O_D4, O_D8, O_D16 \).

**Fig. 2** The \( g_m \) versus \( V_{\text{G}} \) (VDS=1V) measured from (a) narrow-OD nMOS and (b) multi-OD nMOS. \( f_t \) is calculated from \( g_m \).

**Fig. 3** (a) The measured and calculated \( f_t \) versus \( V_{\text{G}} \) (VDS=1.0V) and (b) the gate capacitances \( C_G \) versus \( V_{\text{G}} \) extracted from Im\( Y_{\text{opt}} \) for narrow-OD and standard nMOS.

**Fig. 4** (a) Measured and calculated \( f_t \) vs. \( V_{\text{G}} \) (VDS=1.0V) (b) \( C_G \) vs. \( V_{\text{G}} \) extracted from Im\( Y_{\text{opt}} \) for multi-OD and standard nMOS.

**Fig. 5** (a) The measured and calculated \( f_{\text{max}} \) versus \( V_{\text{G}} \) (VDS=1.0V) and (b) \( R_n \) versus \( V_{\text{G}} \) for narrow-OD nMOS (W1N64, W05N128) and W2N32.

**Fig. 6** (a) The measured and calculated \( f_{\text{max}} \) versus \( V_{\text{G}} \) (VDS=1.0V) and (b) \( R_n \) versus \( V_{\text{G}} \) for multi-OD nMOS (OD4, OD8, OD16) and OD1.

**Fig. 7** Noise parameters of narrow-OD nMOS (a) measured and calculated \( N_{\text{FB}} \), (b) \( R_n \), (c) Re\( Y_{\text{opt}} \), and (d) Im\( Y_{\text{opt}} \).

**Fig. 8** Noise parameters of multi-OD nMOS (a) measured and calculated \( N_{\text{FB}} \), (b) \( R_n \), (c) Re\( Y_{\text{opt}} \), and (d) Im\( Y_{\text{opt}} \).