Design and optimization of FinFETs for ultra-low-voltage analog applications


Published in:
IEEE Transactions on Electron Devices

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Design and Optimization of FinFETs for Ultra-Low-Voltage Analog Applications

Abhinav Kranti and G. Alastair Armstrong

Abstract—In this paper, we analyze the enormous potential of engineering source/drain extension (SDE) regions in FinFETs for ultra-low-voltage (ULV) analog applications. SDE region design can simultaneously improve two key analog figures of merit (FOM)—intrinsic dc gain ($A_{\text{VO}}$) and cutoff frequency ($f_T$)—for 60 and 30 nm FinFETs operated at low drive current ($I_{ds} = 5 \mu A/\mu m$). The improved $A_{\text{VO}}$ and $f_T$ are nearly twice compared to those of devices with abrupt SDE regions. The influence of the SDE region profile and its impact on analog FOM is extensively analyzed. Results show that SDE region optimization provides an additional degree of freedom apart from device parameters (fin width and aspect ratio) to design future nanoscale analog devices. The results are analyzed in terms of spacer-to-straggle ratio—a new design parameter for SDE engineered devices. This paper provides new opportunities for realizing future ULV/low-power analog design with FinFETs.

Index Terms—Capacitances, cutoff frequency, Early voltage, FinFETs, intrinsic voltage gain, source/drain extension (SDE) region engineering, transconductance-to-current ratio, ultra-low-voltage (ULV) analog design.

I. INTRODUCTION

OVER THE past few years, low-power low-voltage silicon-on-insulator MOS technology has emerged as a leading candidate for highly integrated mixed-mode circuits for wireless applications. While digital system design has continually pushed for the increased speed of minimum size devices, analog designers have often employed longer channels to avoid short-channel effects (SCEs) and achieve higher voltage gain. However, in the nanoscale regime, upcoming CMOS technologies face many technological challenges [1], the most crucial being the SCEs that tend to degrade the analog figures of merit (FOM) such as Early voltage ($V_{\text{EA}}$), transconductance-to-current ratio ($g_m/I_{ds}$), intrinsic dc gain ($A_{\text{VO}} = g_m/g_{ds} = g_m/I_{ds} \times V_{\text{EA}}$) and cutoff frequency ($f_T = g_m/2\pi C_{\text{gs}}$) where $g_m$ is the transconductance and $C_{\text{gs}}$ is the total gate capacitance) [2], [3]. To overcome the degradation in analog FOM, certain techniques such as HALO implants and laterally asymmetric channel (LAC) or graded-channel (GC) design [4], [5] have been proposed. However, in nanoscale devices, the control of the dopant profile at the source end of the channel (enabling a feasible LAC/GC or HALO concept) is a technological challenge. Therefore, alternative and innovative techniques are required for improving the analog FOM of sub-100 nm MOS devices.

The weak inversion (WI) region in MOSFETs offers relatively higher $A_{\text{VO}}$ along with lowest power dissipation and lower harmonic distortion [6]. The drawbacks of operating in WI are SCEs and relatively slow devices. Due to SCEs, $f_T$ deteriorates (in addition to $A_{\text{VO}}$) as the gate weakens its control over the channel. SCEs in nanoscale devices can be suppressed by using multigate devices. FinFET, a multigate architecture, has received considerable attention in recent years owing to the suppression of SCEs and excellent scalability, and has thus been regarded as a possible candidate for device scaling at the end of International Technology Roadmap for Semiconductors [1].

We use the concept of source/drain extension (SDE) region engineering (also known as gate-underlap design), which has been studied for digital applications [7]–[10] to significantly improve the analog FOM of 60 and 30 nm FinFETs. In the gate-underlap architecture, the SDE region profiles are designed such that the FinFET channel and SDE regions adjacent to the gate are without any dopant. Our initial work on underlap design in double gate MOSFETs [11] and FinFETs [12] has shown promising results to improve analog FOM. In this paper, we focus specifically on ultra-low-voltage (ULV) analog applications focusing on device design and optimization, analyzing important device parameters such as fin width and aspect ratio apart from the SDE region parameters, and targeting applications for the low-voltage/low-frequency base-band applications.

II. SIMULATIONS

Undoped FinFETs [Fig. 1(a)] analyzed here have been simulated using the 3-D simulator ATLAS [13] with gate length ($L_{g}$) of 60 and 30 nm, gate oxide thickness ($T_{\text{ox}}$) of 2.2 nm, and fin height ($H_{\text{fin}}$) of 60 nm. Fin width ($T_{\text{fin}}$) was varied from 22 to 42 nm for 60 nm devices whereas $T_{\text{fin}}$ was varied from 12 to 24 nm for 30 nm FinFETs. Drain voltage ($V_{\text{ds}}$) was fixed at 0.2 V whereas gate bias ($V_{\text{gs}}$) was always maintained below threshold voltage ($V_{\text{th}}$) in order to analyze the potential of FinFETs for ULV analog applications. Analog/RF FOM are extracted at current density ($J_{\text{ds}}$) of 5 $\mu A/\mu m$. The simulations have been performed with the Lombardi mobility model [13], which accounts for surface roughness scattering, acoustic...
and high (1.2 V) drain bias. The Subthreshold slope (S-slope) as a function of 

device structure is shown in (b). Notations: $\Delta - \Delta - \Delta \sigma = 10$ nm, and 
o--o--o $\sigma = 15$ nm with $d = 5$ nm/dec.

phonon scattering and optical intervalley scattering. Quantum 
effects will not be significant in this paper as undoped FinFETs, with 
$H_{\text{fin}}$ and $T_{\text{fin}} > 10$ nm, are optimized for operation in the 
WI region [14]. The source/drain profile was modeled using the expression 
$N_{\text{SD}}(x) = (N_{\text{SD}}(x))_{\text{peak}} \exp(-x^2/\sigma^2)$, where 
$(N_{\text{SD}})_{\text{peak}}$ is the peak source/drain doping. $\sigma$ (lateral 
straggle) defines the roll-off [10] of the source/drain profile as 
$\sigma = \sqrt{2sd/\ln(10)}$, where $s$ is the spacer width and $d$ is the 
source/drain doping gradient [9], [10] evaluated at the gate edge 
($d = 1/[dN_{\text{SD}}(x)/dx]$) was varied from 3 to 9 nm/decade. The 
lateral straggle parameter $\sigma$ was varied from 5 to 15 nm and the 
spacer widths corresponding to these values of $\sigma$ lie in the 
range of 3–90 nm.

### III. RESULTS AND DISCUSSION

Fig. 2 (a) and (b) show the simulated and experimental [15] 
$\frac{I_{\text{ds}}}{V_{\text{gs}}}$ characteristics for 60 nm FinFETs at low (50 mV) 
and high (1.2 V) drain bias. The Subthreshold slope (S-slope) 
values for these devices are given in Table I. The agreement 
of our simulation with experimental results [15] provides a 
reasonable basis for our analysis. Since our interest is to analyze 
FinFETs for ULV analog applications, we focus our attention to 
reasonable basis for our analysis. Since our interest is to analyze 
our simulation with experimental results [15] provides a 

<table>
<thead>
<tr>
<th>$T_{\text{fin}}$ (nm)</th>
<th>S-slope (mV/dec) @ $V_{\text{ds}} = 50$ mV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>Simulated</td>
</tr>
<tr>
<td>22</td>
<td>68.3</td>
</tr>
<tr>
<td>32</td>
<td>73.5</td>
</tr>
<tr>
<td>42</td>
<td>91.5</td>
</tr>
</tbody>
</table>

The $S_{\text{fin}}$ was varied from 2 to 22 nm in 10 nm steps. Symbols indicate experimental 
$T_{\text{fin}}$ is varied from 22 to 42 nm in 10 nm steps. Notations: $\square$–$\square$ $\sigma = 7.5$ nm, 
$\Delta - \Delta \sigma = 10$ nm, $\times - \times \sigma = 12.5$ nm and o--o $\sigma = 15$ nm.

With selective deposition are available, the reduction of access 
resistance is still a technological challenge. This parasitic re-
sistance will not significantly impact the device performance 
at low $J_{\text{ds}}$. In the subsequent discussion, we will focus our 
attention on devices with $T_{\text{fin}} = 32$ nm and analyze the ULV 
analog behavior by varying the SDE region parameters.

Fig. 3(a) shows the method used in this paper to extract the 
effective channel length ($L_{\text{eff}}$) of gate-underlap devices. Please 
note that $L_{\text{eff}}$ is gate-bias-dependent and therefore different in 
weak and strong inversion regions. As our region of interest is 
in the subthreshold region, we evaluate $L_{\text{eff}}$ in the WI region. 
In the underlap region, the depletion layer boundary (i.e., the 
position in the channel where the electron concentration is 
lower than the net doping) at the source (or drain) edge depends 
on $s$ and $d$. Therefore, as shown in Fig. 3(a), the contribution of 
the SDE region to $L_{\text{eff}}$ can be taken as the distance from the 
depletion layer boundary (at source/drain end) to the gate edge 
[18]. Fig. 3(b) and (c) show the extracted values of $L_{\text{eff}}$ as 
a function of $s$ and $d$ for 60 and 30 nm gate length FinFETs. Devices 
designed with wider $s$ and steeper $d$ result in longer 
$L_{\text{eff}}$ whereas shorter $s$ with gradual $d$ lead to shorter $L_{\text{eff}}$. 

Fig. 4(a) and (b) show the variation of threshold voltage ($V_{\text{th}}$) 
and $S_{\text{fin}}$ as a function of $s$ for various $\sigma$ values. $V_{\text{th}}$ was 
extracted from simulation as the gate bias when the normalized 
drain current ($J_{\text{ds}}/(W_{\text{g}}/L_{\text{g}})$) reaches 400 nA, where $W_{\text{g}}$ is 

Fig. 1. (a) Schematic diagram of a FinFET analyzed in this paper, and 
(b) Variation of source doping profile for various $\sigma$ values along the cut-plane 
along the channel as indicated by dashed lines. Please note that only half of 
the device structure is shown in (b). Notations: $\Delta - \Delta - \Delta \sigma = 10$ nm, and 
o--o--o $\sigma = 15$ nm with $d = 5$ nm/dec.

Fig. 2. $I_{\text{ds}}-V_{\text{gs}}$ curves of FinFET at (a) $V_{\text{ds}} = 50$ mV and (b) $V_{\text{ds}} = 1.2$ V. 
$T_{\text{fin}}$ is varied from 22 to 42 nm in 10 nm steps. Symbols indicate experimental 
data whereas lines refer to 3-D simulation results. Device parameters: $H_{\text{fin}}$ = 
60 nm, $L_{\text{g}}$ = 60 nm and $T_{\text{ox}}$ = 2.2 nm.

Fig. 3. (a) Method to determine the effective channel of gate-underlap devices 
in the WI region. Dependence of $L_{\text{eff}}$ on $s$ and for (b) $L_{\text{g}} = 60$ nm and 
(c) $L_{\text{g}} = 30$ nm. For each $\sigma$ curve, the lowest $L_{\text{eff}}$ value corresponds to 
$d = 9$ nm/dec whereas the highest value represents $d = 3$ nm/dec. $d$ is varied 
from 3 to 9 nm/dec in steps of 2 nm/dec. Notations: $\square$–$\square$ $\sigma = 7.5$ nm, 
$\Delta - \Delta \sigma = 10$ nm, $\times - \times \sigma = 12.5$ nm and o--o $\sigma = 15$ nm.

Fig. 4. (a) Method to determine the effective channel of gate-underlap devices 
in the WI region. Dependence of $L_{\text{eff}}$ on $s$ and for (b) $L_{\text{g}} = 60$ nm and 
(c) $L_{\text{g}} = 30$ nm. For each $\sigma$ curve, the lowest $L_{\text{eff}}$ value corresponds to 
$d = 9$ nm/dec whereas the highest value represents $d = 3$ nm/dec. $d$ is varied 
from 3 to 9 nm/dec in steps of 2 nm/dec. Notations: $\square$–$\square$ $\sigma = 7.5$ nm, 
$\Delta - \Delta \sigma = 10$ nm, $\times - \times \sigma = 12.5$ nm and o--o $\sigma = 15$ nm.
the gate width \((= 2H_{\text{fin}} + T_{\text{fin}})\). FinFETs with wider spacers and larger \(\sigma\) values show suppressed SCEs resulting in near ideal values of \(S\)-slope. As a rough design guide, Fig. 4(b) shows that \(S\)-slope \(\leq 70\) mV/decade are observed for devices designed with \(\sigma \geq 10\) nm, \(d \leq 5\) nm/dec and \(s > 25\) nm. FinFETs designed with lower \(\sigma\) value perform worse than those with abrupt SDE regions as lower \(\sigma\) with higher \(d\) value (at a given \(s\)) result in a shorter \(L_{\text{eff}}\) and degrade the performance [10]. Please note that larger values of \(\sigma\) can be obtained by 1) increasing \(d\) at a constant \(s\) or 2) increasing \(s\) for a fixed \(d\). The case 1) is undesirable as it would result in shorter \(L_{\text{eff}}\) (significant SCEs). The condition 2) is feasible as it yields longer \(L_{\text{eff}}\) (reduced SCEs). In this paper, larger \(\sigma\) values refer to case 2), i.e., increasing \(s\) at a constant \(d\) to obtain the higher \(\sigma\) values.

Fig. 5(a) and (b) show the variation of \(f_T\) and \(A_{\text{VO}}\) with \(s\) for various values of lateral straggles \((\sigma)\), along with the results of FinFETs with abrupt SDE regions. \(f_T\) was extracted as the frequency at which short-circuit current gain \((h_{\text{21}})\) is 0 dB. Results show that SDE region optimization in 60 nm FinFETs can result in high values of \(f_T\) and \(A_{\text{VO}}\) of 40 GHz and 35 dB, respectively, an improvement of \(\sim 1.5-2\) compared to devices with abrupt SDE regions. FinFETs with gradual gradients \((d = 7-9\) nm/dec\) at lower \(\sigma\) values perform worse than those with abrupt SDE regions due to SCEs as \(L_{\text{eff}} < L_g\). An increase in \(\sigma\) from 5 to 15 leads to an improvement in both \(A_{\text{VO}}\) and \(f_T\). This is somewhat surprising as it indicates an increase in \(f_T\) with increase in \(L_{\text{eff}}\). We will show in subsequent discussion that the key issue to improve \(f_T\) is the effectiveness of underlap design in reducing the fringing capacitance, along with the improvement in \(g_m\). Also, devices designed with large \(\sigma\) values along with much larger spacer widths \((\sigma = 15\) with \(s = 86\) nm\) lead to a reduction in the \(f_T\) due to the additional parasitic series resistance associated with wider spacers. It is important to note that in a conventional design with abrupt SDE regions, it is not possible to attain a simultaneous improvement in \(A_{\text{VO}}\) and \(f_T\), as an increase in \(A_{\text{VO}}\) requires a longer \(L_g\) (assuming a linear dependence of \(V_{\text{EA}}\) on \(L_g\)), which would compromise \(f_T\) because of a reduction in \(g_m\) (\(g_m\) is approximately inversely proportional to \(L_g\)). As shown in Fig. 5(a) and (b), \(f_T\) and \(A_{\text{VO}}\) values appear to saturate at about 40 GHz and 35 dB, respectively, in devices designed with \(\sigma \geq 10\) with \(s\) lying in the range of 40–80 nm. The limitation of too wide a spacer or too large a straggles is extra parasitic series resistance effect. These values of \(s\) suggest that to achieve significant improvement in both \(A_{\text{VO}}\) and \(f_T\), the maximum \(s\) value should be limited to 60 nm. However, FinFETs designed with optimal \(\sigma\) values and wider spacers within this limit result in higher values of both \(f_T\) and \(A_{\text{VO}}\).

To further investigate the potential of underlap design for ULV analog applications in short channel FinFETs, we examined the extent of improvement in \(f_T\) and \(A_{\text{VO}}\) in 30 nm devices. As shown in Fig. 6(a) and (b), for a substantial improvement \((\sim \times 2.5\) as compared with abrupt SDE devices\) in \(f_T\) and \(A_{\text{VO}}\), FinFETs should be designed with \(\sigma\) values lying between 10–15 nm along with wider spacers \((\sim 50\) nm\). The extent of improvement in \(A_{\text{VO}}\) and \(f_T\) is much higher in 30 nm devices \((\sim \times 2.5)\) as compared to 60 nm FinFETs \((\sim \times 1.5)\) thus suggesting that SDE engineering will become even more suitable at ultra-short gate lengths. A crossover is observed between the data relating to \(d = 3\) and 5 nm/dec for \(\sigma = 12.5\) and 15 nm, respectively, in the graphs of \(f_T\) and \(A_{\text{VO}}\). As \(s\) is increased beyond a certain critical value \((\approx 55\) nm\), parasitic series resistance effect associated with wider spacers and steeper gradient becomes important and eventually the rate of improvement of analog FOM decreases. For the case of \(\sigma = 12.5\) nm, \(f_T\) and \(A_{\text{VO}}\) increase sharply up to \(d = 5\) nm/dec \((s \approx 36\) nm\). However, at \(d = 3\) nm/dec \((s \approx 60\) nm\), parasitic series resistance dominates and the rate of improvement in analog FOM decreases. For \(\sigma = 15\) nm, \(f_T\) and \(A_{\text{VO}}\) improve sharply with an increase in spacer width up to \(d = 5\) nm/dec \((s \approx 50\) nm\) and analog FOM does not degrade. However, the rate of improvement of \(f_T\) and \(A_{\text{VO}}\) degrade when spacer is increased beyond 55 nm \((\sigma = 15\) nm, \(d = 3\) nm/dec and \(s \approx 85\) nm\). The different values of spacer width that lead to the onset of parasitic series resistance effect in the case of \(\sigma = 12.5\) nm \((d = 3\) nm/dec\) and \(\sigma = 15\) nm \((d = 5\) nm/dec\)
causes the crossover between the data points. Also, as shown in Fig. 3(b) and (c), \( I_{\text{eff}} = 12.5 \text{ nm/dec} > (I_{\text{eff}}) = 15 \text{ nm/dec} \) thus suggesting that wider spacers for \( \sigma = 12.5 \text{ nm} \) will degrade the device performance as compared to that designed with \( \sigma = 15 \text{ nm} \) (\( d = 5 \text{ nm/dec} \)).

Fig. 7(a) shows the \( I_{\text{ds}}-V_{gs} \) characteristics at drain bias of 0.2 V for 30 nm devices. An increase in \( \sigma \) or \( s \) for a given \( d \) increases \( I_{\text{eff}} \), which results in a reduction in the off-current \( (I_{\text{off}} = I_{\text{ds}} \text{ at } V_{gs} = 0 \text{ V}) \). As shown in Fig. 7(b), for \( J_{\text{ds}} < 50 \mu\text{A/\mu m} \), higher \( g_m \) values are exhibited by underlap devices designed with larger \( \sigma \) values whereas at larger \( J_{\text{ds}} (50 \mu\text{A/\mu m} < J_{\text{ds}} < 100 \mu\text{A/\mu m}) \), FinFETs with \( \sigma < 12.5 \text{ nm} \) perform better. The peak \( g_m \) sharply reduces with an increase in \( \sigma \) due to the parasitic series resistance associated with wider spacer regions. As our interest is in the ULV analog operations, we will concentrate at \( J_{\text{ds}} = 5 \mu\text{A/\mu m} \) (\( V_{ds} = 0.2 \text{ V} \)). The results at higher \( J_{\text{ds}} \) and higher drain voltages are discussed later in this paper.

To analyze the reason for the enormous improvement in \( A_{\text{VO}} = g_m/I_{\text{ds}} \times V_{\text{EA}} \) and \( f_T = g_m/(2\pi C_{\text{gs}}) \), we evaluate other important analog parameters, such as transconductance-
to-current ratio \( (g_m/I_{\text{ds}}) \), Early voltage \( V_{\text{EA}} \), transconductance \( (g_m) \) and total input capacitance \( (C_{\text{gs}}) \) focusing on 30 nm FinFETs. Fig. 8(a) and (b) shows the variation of \( g_m/I_{\text{ds}} \) with normalized drain current \( (I_{\text{ds}}/(W_g/L_g)) \) for various \( \sigma \) values. \( g_m/I_{\text{ds}} \) ratio is a measure of the efficiency to translate current (hence power) into transconductance [2]. The \( g_m/I_{\text{ds}} \) parameter does not depend on device dimensions (to the first order) and its value is inversely proportional to the channel inversion level. An increase in \( \sigma \) shifts the source/drain doping away from the gate edge, thus minimizing the influence of drain on the channel region, leading to higher values (in WI region) of \( g_m/I_{\text{ds}} \sim 36 \text{ V}^{-1} \) as compared to \sim 30 \text{ V}^{-1} \) for abrupt SDE regions. This reduction in SCEs is reflected in an increase in current ratio \( ((I_{\text{ds}})/(I_{\text{ds}}_{\text{SDE}}))/(I_{\text{ds}}_{\text{Abrupt}})) \) to 2 (as compared to devices with abrupt SDE regions) at \( g_m/I_{\text{ds}} = 25 \mu\text{A/\mu m} \) for \( \sigma = 15 \text{ nm} \). This factor of \sim 2 times improvement in drain current translates into a higher \( A_{\text{VO}} \) at lower \( J_{\text{ds}} \). Please note that the reduction of \( (g_m/I_{\text{ds}})_{\text{peak}} \) in the WI region from the ideal value of \sim 38 \text{ V}^{-1} \), signifies SCEs as \( (g_m/I_{\text{ds}})_{\text{peak}} = (\ln(10))/S \)-slope. SCE region optimization is particularly advantageous at low \( V_{gs} \), as the current flow is mainly due to diffusion of carriers, where a wider spacer region (\sim 45 nm; \( \sigma = 15 \text{ nm} \)) does not degrade the device performance. However, as shown in Fig. 8(b), in strong inversion region (\sim 5 \text{ V} \_1), larger \( \sigma \) values and wider spacers introduce additional parasitic series resistance, which degrades the performance (current ratio \( ((I_{\text{ds}})/(I_{\text{ds}}_{\text{SDE}}))/(I_{\text{ds}}_{\text{Abrupt}})) < 1 \) results in the degradation in \( g_m \).

Fig. 9(a) shows the variation of total gate capacitance \( (C_{gs} = C_{gs} + C_{gd} + C_{gb}) \), where \( C_{gs}, C_{gd}, \text{ and } C_{gb} \) represent gate-to-source, gate-to-drain and gate-to-substrate capacitances, respectively) with spacer widths for various \( \sigma \) values. An increase in \( s \) shifts the source/drain doping away from the gate edge, resulting in a significant reduction in parasitic fringing capacitance, which leads to the decrease in \( C_{gs} \). This reduction is nearly 60% (\( \sigma \geq 10 \text{ with } s \geq 40 \text{ nm} \) when compared with abrupt SDE FinFETs. The total fringing capacitance, composed of internal \( (C_{ii}) \) and external \( (C_{ic}) \) capacitances [Fig. 9(b)], is directly proportional to the distance of the gate edge from the source/drain doping profile. For an increase in spacer width, \( C_{gs} \) initially decreases sharply and thereafter linearly as the capacitance is dominated by...
the fringing component ($C_f$). This causes nearly comparable $C_{gg}$ values for two different $\sigma$ values (12.5 and 15 nm) at different $d$ values (3 and 5 nm/dec). It is important to note that SDE engineered FinFETs with shorter spacers and gradual gradients achieve nearly the same value of $C_{gg}$ as compared to devices with abrupt SDE regions. The data corresponding to $C_{gg} > 0.8 \hspace{1em}\Omega\hspace{1em}fF/\mu m$, i.e., $C_{gg}$ SDE/$C_{gg}$ ABRUPT $> 1$, represents gate-overlap architecture rather than the desirable gate-underlap design. Thus, the optimization of the SDE region is extremely important to minimize $C_{gg}$ associated with structure. It is important to note that such a drastic reduction in parasitic capacitance is not possible in other approaches [4], [5] that have been suggested for improving analog FOM.

Fig. 10(a)–(c) shows the variation of transconductance ($g_m$), output conductance ($g_{ds}$) and Early voltage ($V_EA$ for $I_{ds} = I_{ds}/g_{ds}$) with $s$ for various $\sigma$ values. An increase in $\sigma$ with the range 5–15 nm improves the gate control (due to reduction in SCEs) and increases $g_m$. The increase in $g_m$ is in the range of 30%–60% for $\sigma \geq 10$ nm and $s \geq 40$ nm. An increase in $\sigma$ results in higher $g_m/I_{ds}$ values (Fig. 8) which implies an increased $g_m$ at the same current. Multigate devices (FinFETs) achieve higher values of $g_m$ than single gate devices at lower $V_{gs}$ due to the volume inversion effect [3]. Also, it is important to note that SDE devices designed with larger $\sigma$ ($\geq 12.5$ nm) along with wider spacers ($\geq 60$ nm) result in the degradation of $g_m$ and $f_T$ due to parasitic series resistance effect. The increase of 30%–60% in $g_m$ translates into an equivalent relative increase in Gain Bandwidth product ($GBW = g_m/2\pi C_L$, where $C_L$ is the load capacitance). The most significant aspect increasing $s$ in underlap design is the reduction in $g_{ds}$ leading to the improvement in $V_{EA}$. Due to an increase in $s$, reduction in the net dopant concentration causes a substantial reduction in the peak electric field at the gate edge yielding higher values of $V_{EA}$ [12]. Underlap design results in a relative reduction of $g_{ds}$ by a factor of $\sim 12$ for $\sigma \geq 12.5$ nm with $s \geq 40$ nm. The large reduction in $g_{ds}$ combined with an improvement in $g_m$ is responsible for the higher values of $A_{VO}$, shown in Figs. 5 and 6. Please note that $g_{ds}$ and $V_{EA}$ are governed by (i) channel resistance, i.e., region under the gate and (ii) resistance of the spacer region. For large $s$ values, $g_{ds}$ is dominated by the resistance of the spacer region. This results in comparable values of $g_{ds}$ and $V_{EA}$ for $\sigma = 12.5$ nm ($d = 3$ nm/dec) and 15 nm ($d = 5$ nm/dec) and a crossover between these data points.

Fig. 11 shows the dependence of $f_T$ and $A_{VO}$ on $T_{fin}$ along with the results of abrupt SDE regions. FinFETs with abrupt SDE regions need to be designed with thinner silicon films to achieve higher values of $f_T$ and $A_{VO}$ whereas SDE engineering provides additional flexibility in the choice of $T_{fin}$. This will be extremely beneficial for nanoscale devices where high quality thin films are required to limit SCEs and provide improvement in analog FOM. The lowest value of $T_{fin}$ (= 12 nm) corresponds to an aspect ratio ($AR = H_{fin}/T_{fin}$) of 5 whereas the largest $T_{fin}$ values refer to an AR of 2.5. SDE engineered FinFETs designed with thicker $T_{fin}$, i.e., lower AR ($AR = 2.5$ at $T_{fin} = 24$ nm) achieve higher values of $f_T$ and $A_{VO}$ as
obtained with abrupt SDE regions with $T_{\text{fin}} = 12 \text{ nm} (\text{AR} = 5)$. Doubling $T_{\text{fin}}$ in SDE-engineered FinFETs reduces $f_T$ by only 20% at lower $J_{ds}$, but the relative improvement in $f_T$ increases.

IV. OPTIMIZATION OF ANALOG FOM

Our results in the previous section has shown that the gate-underlap region is characterized by three parameters—$s$, $d$, and $\sigma$. In order to develop guidelines for optimal device performance, we evaluate the dependence of $f_T$ and $A_{\text{VOC}}$ on spacer-to-straggle ($s/\sigma$). $s/\sigma$ serves as a simple design parameter for the optimization of SDE regions as the $L_{\text{eff}}$ is a function of $V_{gs}$ and thus different in weak and strong inversion regions. As shown in Fig. 12(a) and (b), the lower limit on $s/\sigma$, i.e., $(s/\sigma)_{\text{min}} = 2.5$ must be maintained to avoid SCEs ($L_{\text{eff}} > L_g$) and achieve higher analog FOM as compared to devices with abrupt SDE regions, whereas $(s/\sigma)_{\text{max}} > 4$ signifies a degradation in $g_m$ and $f_T$ due to parasitic series resistance associated with wider spacers. The minimum value of $(s/\sigma)$ of 2.5 translates into a maximum acceptable source/drain doping concentration ($N_{SD})_{\text{peak}}$ of ~3 orders of magnitude reduction at the gate edge. FinFETs designed with $s/\sigma < 2.5$ perform worse than devices with abrupt SDE regions because a large $d$ at lower $\sigma$ causes dopant spill into the channel, leading to lower $A_{\text{VOC}}$ and $f_T$. Since the value of $d$ depends on thermal budget and diffusivity, very small values ($< 3 \text{ nm/dec}$) may be difficult to achieve (very small values of $d$ require a nonstandard process [19], [20] such as solid phase epitaxy or laser thermal annealing), as it may ultimately require control of individual atoms. Therefore, it is more appropriate to increase $s$ to achieve higher $s/\sigma$ values. Our results suggest that $s$ should be approximately equal to $3\sigma$, i.e., the spacer should be wide enough to ensure that more than 99.5% of the SDE region profile should be across the spacer. Also, as lateral straggling is equal to $2\sqrt{D\tau}$, where $D$ is the Diffusion coefficient (in units of square centimeters per second) and $t$ is the time, optimal straggles ($\sigma = 10–15 \text{ nm}$) values can be understood in terms of the diffusion length of the SDE profile. Considering a typical process at $T = 1000 \text{ °C}$ with diffusivity of $5 \times 10^{-15} \text{ cm}^2 \cdot \text{s}^{-1}$, achieving $\sigma$ in the range of 10–15 nm would require an annealing time in the range of 50–110 s.

Fig. 12. Dependence of (a) $f_T$ and (b) $A_{\text{VOC}}$ on spacer-to-straggle ratio $(s/\sigma)$ for different values of $\sigma$ extracted at $J_{ds} = 5 \mu\text{A/um}$ and $V_{ds} = 0.2 \text{ V}$. The horizontal line on the $y$-axis represent $f_T$ and $A_{\text{VOC}}$ values of FinFET designed with abrupt SDE regions. For each $\sigma$ curve, the lowest $f_T$ or $A_{\text{VOC}}$ value corresponds to $d = 9 \text{ nm/dec}$. Notations and device parameters are the same as in Fig. 6.

Fig. 13 compares the analog FOM for two different gate length devices, but with the same total source-to-drain length. In order to generate a $\sigma$ range (8–18 nm), $d$ was varied for both devices. Despite an increase in $s/L_g$ ratio from 0.5 ($L_g = 60 \text{ nm}$) to 1.5 ($L_g = 30 \text{ nm}$) a relative improvement of up to 70% in $f_T$ occurs, for $\sigma$ in the range of 12–16 nm. Interestingly, this is approximately in line with the very simple theory associated with ideal abrupt junctions which predicts $f_T$ is inversely proportional to $L_g$. It is significant that the anticipated improvement in $f_T$ with traditional scaling is only achievable if a relatively optimal value of $\sigma$ is used. Such an optimally designed SDE gives the additional benefit of a very significant improvement in $A_{\text{VOC}}$ (of up to 20 dB), making it possible in principle to design a nanoscale CMOS wideband low power op-amp with $A_{\text{VOC}} \sim 40 \text{ dB}$. Indeed, if the spacer is of sufficient length ($> 3\sigma$), and $\sigma < 14 \text{ nm}$, neither $A_{\text{VOC}}$ nor $f_T$ in the shorter device is critically dependent on the precise value of $\sigma$.

Fig. 14 shows the dependence of extra series resistance $(\Delta R = R_{\text{underlap}} - R_{\text{Abrupt}})$ associated with the underlap design. The parasitic resistance was extracted by the method proposed in [17] and the value determined by our simulations (6.15 kΩ) for the abrupt SDE region agrees well with the experimental result of ~6.25 kΩ reported in [17]. For devices designed with $\sigma \leq 12.5 \text{ nm}$ and $s < 40 \text{ nm}$, with $d$ in the range of 5–9 nm/decade, $\Delta R$ is below 0.5 kΩ, whereas it increases sharply to ~1 kΩ for $s > 40 \text{ nm}$. Any additional series resistance introduced by our design technique ($s \leq 3\sigma$) will be significantly less than the overall resistance including contact resistance. The extra resistance will significantly degrade the performance in the strong inversion region. For a device operating in the WI region ($J_{ds} = 5 \mu\text{A/um}$), the extra resistance will not significantly limit the performance except for $s > 60 \text{ nm}$. In this paper, we have focused our attention in optimizing the underlap region parameters to improve analog/RF FOM at low current densities. Analog/RF FinFETs may be operated
at higher $J_{ds}$ corresponding to the peak $g_m$ (or $f_T$) of the device. A recent study by Dickson et al. [21] suggested that 1) peak values of $f_T$, maximum frequency of oscillation ($f_{MAX}$), minimum noise figure ($NF_{min}$) are obtained at 300 $\mu$A/$\mu$m, 200 $\mu$A/$\mu$m and 150 $\mu$A/$\mu$m, respectively for a large range of Si(Ge) (Bi)CMOS technologies and 2) $J_{ds}$ corresponding to peak $f_T$, $f_{MAX}$, and $NF_{min}$ remain invariant over technology node and foundries. In this paragraph, we analyze the behavior of $f_T$ and $f_{MAX}$ at high $V_{ds}$ of underlap and nonunderlap FinFETs. As shown by the comparison between measurements reported by Parvais et al. [22] and our simulations in Fig. 15(a) and (b), nonunderlap FinFETs designed with thicker fins achieve peak $f_T$ and $f_{MAX}$ values at similar $J_{ds}$. However, as $T_{fin}$ is reduced, not only are peak $f_T$ and $f_{MAX}$ values reduced, but the peak $f_T$ and $f_{MAX}$ values occur at lower $J_{ds}$ (200 $\mu$A/$\mu$m and $\sim$100 $\mu$A/$\mu$m $T_{fin}$ = 25 and 17 nm, respectively). This is due to the parasitic source/drain resistance associated with a nonplanar multigate structure [16], [17]. Therefore, the result from [21] regarding the invariance of $J_{ds}$ corresponding to peak $f_T$ and $f_{MAX}$ is appropriate for planar MOSFETs, but only for nonplanar FinFETs with thick fins.

Fig. 16(a) and (b) show the variation of $f_T$ and $f_{MAX}$ with $J_{ds}$ in gate-underlap FinFETs operated at higher $V_{ds}$ (=1.2 V), where the effect of parasitic resistances has been included, with appropriate values of contact resistance taken from [17]. An increase in $\sigma$ shifts 1) peak $f_T$ and $f_{MAX}$ values to lower $J_{ds}$ and 2) peak $f_T$ and $f_{MAX}$ values at lower $J_{ds}$ are significantly improved because of the underlap design. It is important to note that key analog/RF FOMs such as $f_T$ and $f_{MAX}$ do not degrade in an underlap design when operated at $J_{ds} < 50 \mu A/\mu m$ ($V_{ds} = 1.2$ V). The relative improvement in $f_{MAX}$ at low $J_{ds}$ is very similar to that already stated for $f_T$. Please note that although ($f_T$)_{Underlap} < ($f_T$)_{AbruptSDE} for $J_{ds} > 100 \mu A/\mu m$, underlap devices achieve higher $f_{MAX}$ values due to the contribution of the term $\sqrt{g_{ds} + 2\pi f_T C_{gd}}$ ($f_{MAX}$ is inversely proportional to $\sqrt{g_{ds} + 2\pi f_T C_{gd}}$) due to a drastic reduction in $g_{ds}$ and capacitances.

Table II illustrates trends and provides guidance in determining optimal SDE region parameters for different gate lengths. The considerable improvement in key analog FOM at shorter gate lengths reflects the significant potential of SDE region optimization in nanoscale devices for ULV applications. $s/L_q$ and $\sigma/L_q$ ratios must be increased while reducing $L_q$ in order to improve $f_T$ and $A_{VOS}$, $s/L_q$ ratio should be increased from 1.0 to 1.7 as $L_q$ is reduced from 60 to 30 nm. This implies that sub-30 nm devices should be designed with source/drain diffusion length ($\sigma$) up to three quarters of the gate length, along with ultra-wide spacers (e.g., 2.6$L_q$).

It must be emphasized that we have presented a comparative assessment of FinFETs (designed with abrupt SDE regions and underlap design) at low current levels ($\sim$5 $\mu A/\mu m$ at $V_{ds} = 0.2$ V and $V_{gs} < V_{th}$) intended for ULV applications. Such SDE region engineered devices, with wider than expected spacers, will not be useful for operation at higher $J_{ds}$ unless the spacer is shorter than the gate length, due to the parasitic series resistance that severely degrades the device performance. Also, it must
be noted that due to the uncertainties involved in the s and σ values in nanoscale devices, some of the values reflecting the improvement may be different from those stated in the text. Nevertheless, the overall trend in the design of an optimal low-voltage/low-power device of well-calibrated straggle and progressively increasing spacer with reduction in gate length is very significant.

V. CONCLUSION

The enormous potential of SDE region engineering in FinFETs for analog applications has been extensively analyzed for long and short gate length devices of varying aspect ratio and fin thickness. Gate-underlap FinFETs operated at low current levels are particularly useful for ULV analog applications as both gain and speed of devices can be drastically improved. The enormous potential of SDE region engineered FinFETs for ULV analog applications is attributed to:

1) improvement in transconductance due to suppressed SCEs;
2) reduction in drain conductance and consequent improvement in Early voltage due to reduction in peak electric field;
3) the drastic reduction in parasitic fringing capacitance due to an optimal gate-source/drain underlap design.

The potential to control the spacer width and lateral straggler of the doping profile provides an additional degree of freedom alongside the aspect ratio, to design future analog devices. Optimally-engineered FinFETs do show high tolerance to variations in spacer widths and lateral straggler values, thereby relaxing the process control requirements. Values of spacer-to-straggle ratio lying between 2.5 and 4 represent the lower and upper limit required for achieving an improvement factor > 2 in key analog FOM over a more traditional design with abrupt source–drain junctions. The lower limit on spacer-to-lateral straggler ratio corresponds to an upper limit on the net doping concentration at the gate edge of ~3 orders of magnitude lower than the peak source/drain doping. As gate length is reduced, it is significant that neither the spacer length nor the lateral straggler change significantly, with the optimal spacer lying in the range 40–55 nm and optimal straggler in the range 10–15 nm. The impact of parasitic resistance associated with wider spacers on cutoff frequency and maximum frequency of oscillations is minimal if operated at lower current densities (~50 μA/μm). The designs featured in this paper have focused on an optimal device for ultra-low-power analog applications, but the same principles apply when designing a device to meet a low operating power International Technology Roadmap for Semiconductors specification for a digital application, which will typically require a shorter spacer but very similar values of lateral straggler.

REFERENCES

Abhinav Kranti was born in New Delhi, India, in 1975. He received the B.S. (Honours) and M.S. degrees in physics from St. Stephen’s College, University of Delhi, Delhi, India, in 1996 and 1998, respectively, and the Ph.D. degree in electronic science from University of Delhi, in 2002.

His doctoral work focused on modeling short channel vertical surrounding gate MOSFETs. His paper submitted to IEEE Topical meeting on silicon monolithic integrated circuits in RF systems, USA, 2001 was selected as one of the outstanding student papers. From July 2002 till November 2004, he was a Postdoctoral Researcher at Microwave Laboratory, Université catholique de Louvain, Louvain-la-Neuve, Belgium where he worked on graded channel architecture and multigate MOSFETs for analog/RF applications. Since February 2005, he is working as a Research Fellow at Northern Ireland Semiconductor Research Centre, Queen’s University Belfast, Belfast, U.K., on low-voltage analog/RF device and circuit design with multiple gate MOSFETs. He has also worked on the extraction and optimization of thermal resistance in trench isolated Bipolar transistors. His research interests include semiconductor device physics, device/circuit modeling, novel FET structures for advanced CMOS applications, thermal resistance management in bipolar and MOS devices and GaN-based wide-bandgap devices for microwave applications. In these fields, he has coauthored more than 50 scientific articles in international journals and conferences.

Dr. Kranti was awarded Senior Research Fellowship by Council of Scientific and Industrial Research, Government of India, New Delhi, in May 2001.

G. Alastair Armstrong received the B.Eng. degree in electrical engineering and the Ph.D. degree in electronics from Queen’s University, Belfast, U.K., in 1968 and 1971, respectively.

After a number of years working in industry, he was appointed in 1982 as a Lecturer in Electronics at Queen’s University and a Senior Lecturer in 1989. He currently holds the Chair of electronic engineering at Queen’s University and has more than 25 years experience in simulation and design of semiconductor devices and circuits. As Assistant Director of the Northern Ireland Semiconductor Research Centre, he leads the semiconductor device simulation research group. He has more than 25 years experience in simulation and design of semiconductor devices and circuits including MOS, SOI and bipolar devices. Over the course of his career he has coauthored two textbooks and published more than 130 journal and conference papers on topics involving semiconductor devices, signal processing and medical electronics. During the course of his career, he has worked as a consultant to several electronics companies. His current research interest is in the application of silicon on insulator technologies in nanoelectronics.

Dr. Armstrong has been awarded more than 12 research grants all involving some aspect of semiconductor device simulation for silicon technologies.