POLITECNICO DI TORINO

Master's Degree in Electronic Engineering

Master's Degree Thesis

Nanosheet-GAAFETs modeling and circuit performance evaluation

Supervisor
Prof. Gianluca PICCININI

Candidate
Martina AMATO

Aprile 2021

To my beloved mother, who has always sustained me and tried to lighten the burden of these years.

Summary

During the last decade, three-dimensional electronic devices, known as Fin field-effect transistors (FinFETs), have been developed to pursue continuous technology scaling, by improving device performance while reducing short-channel effects (SCEs). However, FinFETs are currently facing many challenges in terms of performance, layout and cost for further scaling beyond the 7-nm node.
As a matter of fact, nowadays, very thin and tall fin structures would be required to maintain the benefits of such a 3D device, thus raising concerns for both performance and fabrication process.
In this scenario, silicon nanosheet gate-all-around field-effect transistors (NSGAAFETs) have been recognized as excellent candidates to replace fin devices for sub-7nm nodes, due to superior channel electrostatic control and great drive current.
The first part of this work presents the most common NSGAAFET fabrication process, underlying the peculiar steps compared to the already known FinFET flow. Furthermore, a deep exploration of NSGAAFET structure is provided, in order to understand how to effectively model such a novel device.
Then, the unified multi-gate BSIM-CMG model is illustrated and analysed. Firstly, its core section is described, stressing BSIM-CMG capability to capture the behaviour of different multigate devices, including a very basic version of a quadruple gate FET. After that, the attention moves to the modeling of parasitics (resistances and capacitances), which acquire huge importance in very scaled devices.

The second part of the present work, instead, focuses on adapting the existing BSIM-CMG model to an advanced single- and multi-stacked NSGAAFET. To this purpose, new parameters are introduced and parasitics modeling is carefully modified and improved. Then, the new model is applied to investigate a high performance (HP) three-stacked nanosheet gate-all-around field effect transistor (3-NSGAAFET).
In terms of DC performance, main figures of merit, such as I_{dsat}, I_{sub}, SS and $DIBL$, have been extracted using Cadence Virtuoso. The obtained values show excellent drive current, as required for a HP device, and optimal channel electrostatic control. This key aspect is further explored for rising nanosheets widths,

together with the impact of process variations on I_{dsat}.
As regards the AC performance, a five-stage ring oscillator has been simulated and its oscillation frequency f_{osc} extracted. Moreover, the influence of process variations and of nanosheets size on f_{osc} is inspected.
As expected, simulations underline an increase of the oscillation frequency with the raise of nanosheets size, till a maximum value. This is identified as the optimum one for ensuring high performance.
Finally, the developed model is applied to the design of some basic cells (inverter, NAND, NOR gates). The implemented gates are analysed in terms of worst-case delay, for different nanosheets widths.
For each gate, an optimum nanosheet size is found, representing a good trade-off between high performance and occupation area.

Acknowledgements

I must thank all the people who have guided me in the exploration of this novel technology. All the meetings with Professor Piccinini and PhD students Chiara, Fabrizio and Yuri allowed me to increase my knowledge, giving me precious ideas to carry out this thesis. I would like to give special thanks to Chiara, who has always encouraged me to give my best, and to Fabrizio, who has followed my work with care and great passion as no one could have done better.

Contents

List of Tables

List of Figures

List of acronyms

GAAFET	Gate-All-Around Field Effect Transistor
NSGAAFET	Nanosheet Gate-All-Around Field Effect Transistor
BOX	Buried Oxide
SIT	Sidewall Image Transfer
CMP	Chemical Mechanical Polishing
RIE	Reactive Ion Etching
PSG	Phosposilicate Glass
ALD	Atomic Layer Deposition
SAC	Self-Aligned Contact
Mug-FET	Multigate Field-Effect Transistor
GCA	Gradual Channel Approximation
DG	Double Gate
TG	Triple Gate
QG	Quadruple Gate
CG	Cylindrical Gate
SDE	Source/Drain Extension
EOT	Equivalent oxide thickness
ILD	Inter layer dielectric
IL	Interfacial layer
HP	High performance
3-NSGAAFET	Three-stacked Nanosheet Gate-All-Around Field Effect Transistor
DIBL	Drain induced barrier lowering
SS	Subthreshold slope
FO	Fan-out

Chapter 1

Introduction to Nanosheet-GAAFETs

The purpose of this first chapter is to provide an overview of the nanosheet gate-all-around field effect transistor (NSGAAFET) state of the art and to illustrate the main fabrication process, underlying the peculiar steps required for its realization. In addition, NSGAAFET structure is illustrated, with the aim to introduce some key geometrical parameters that will be of concern in the rest of the thesis.

1.1 State of the art

The recent development of nanosheet gate-all-around field effect transistors has promoted them to the most promising candidates for next generation CMOS devices beyond FinFET [14]. These novel structures, indeed, provide optimal gate controllability and so high immunity to short channel effects, thus enabling ultimate device scaling. Furthermore, they show excellent drive current and so are indicated for high performance applications. In addition, they are compatible with FinFET fabrication, making it possible to reuse most of the steps of the already known process flow.

1.2 Fabrication process

The most commonly used process for Nanosheet-GAAFETs fabrication is known as "Nanosheet last", since the silicon nanosheets are released only in the last part of the flow, that is after the dummy gate removal.
The entire process is now described in details and partioned in four macro-sections, starting from an already prepared SOI substrate [2].

- The process begins with the epitaxial growth of $Si_{0.7}Ge_{0.3}/Si$ multilayers. This involves $SiH_2Cl_2 + GeH_4$ for SiGe layers and SiH_4 for the Si ones. The layers thickness typically ranges between 7nm to 12nm.
  Then, individual and dense arrays of $Si_{0.7}Ge_{0.3}/Si$ fins are patterned using the advanced Sidewall Image Transfer (SIT) technique, which allows to define the nanosheet width on the basis of the etching process parameters.
  This is followed by the deposition and planarization of a conformal layer of polysilicon (CMP) on the whole structure.
  A selective etching of the polysilicon, performed through reactive ion etching (RIE), is now essential to define the sacrificial gate. Successively, $SiOCN$ main spacers are deposited and patterned along the dummy gate with an anisotropic etching process.
  The choice of $SiOCN$ as spacer material, instead of the traditional SiN, is justified by its lower dielectric constant ($\epsilon \simeq 5$) with an almost unchanged robustness.

- The following steps are mainly devoted to the realization of the internal spacers, which will be fundamental to avoid unwanted shortcircuits between the future interchannel metal gates and the source/drain (S/D) epitaxial structures. To this purpose, a first vertical anisotropic etching of the portions of $Si_{0.7}Ge_{0.3}/Si$ layers placed outside the main spacers, is carried out. This is done maintaining just few nanometers of the first layer on the bottom for the following epitaxial growth of source/drain regions.
  At this point, the $Si_{0.7}Ge_{0.3}$ layers below the spacers are etched away, leaving some cavities that will be later filled with $SiOCN$, giving origin to the so called inner spacers.
  On the basis of the depth of the etching process, the inner spacers can result more or less aligned to main sidewalls, with consequences on NSGAAFET AC performance (section 5.2).

- Source and drain regions are now epitaxially grown until the upper nanosheet, merging with the interchannels. A salicidation can be successively performed on the top of the S/D structures in order to reduce their resistance.
  Then, a phosposilicate glass (PSG) is deposited on the structure as insulating layer, followed by a planarization process that leaves just the dummy gate exposed. This last can now be removed and, after that, also the sacrificial $Si_{0.7}Ge_{0.3}$ layers can be etched away. During this process, inner spacers act as etch-stop layers.
  Si channels appear now suspended between the source and drain structures.

- The gaps left by $Si_{0.7}Ge_{0.3}$ removal are filled with interfacial layer (SiO_2), high-k dielectric (HfO_2) and metal gate stack, through the atomic layer deposition (ALD).

Then, tungsten is deposited on the whole structure to reach the gate contact and a CMP is carried out to have just the gate trench filled with tungsten. Finally a self-aligned contact (SAC) module can be implemented for source/drain contacts, with the aim to avoid unwanted possible shortcircuits between S/D and gate contacts related to problems of misalignment.

The process illustrated results to have great similarities with the FinFET one, except for four peculiar steps: $Si_{0.7}Ge_{0.3}/Si$ epitaxial growth, $Si_{0.7}Ge_{0.3}/Si$ fin patterning, inner spacer definition and sacrificial $Si_{0.7}Ge_{0.3}$ layers removal.
The high compatibility with FinFET process makes this approach the mainstream for NSGAAFET fabrication.

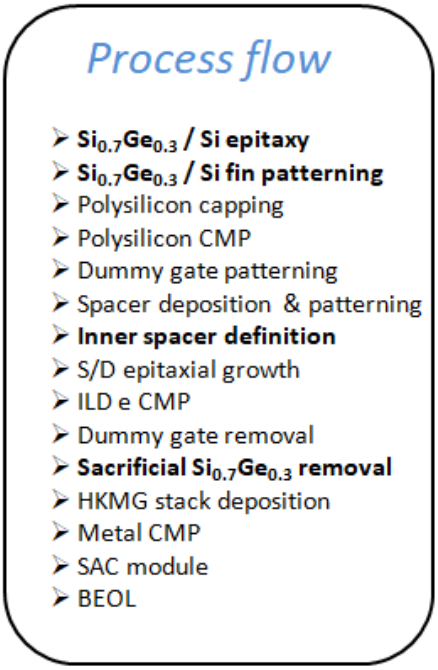


Figure 1.1: Fabrication process of stacked NSGAAFET.

1.3 Device structure

The final NSGAAFET structure appears as a stack of N_{sh} nanosheets, which are wide and thin silicon sheets, characterized by a width W_{sh} and a height H_{sh}. They are layered one upon the other, with the metal gate and oxide material surrounding each of them on all the sides.
As a consequence, the effective channel width for a nanosheet gate-all-around field effect transistor is given by: $W_{eff} = N_{sh} \times [2 \cdot W_{sh} + 2 \cdot H_{sh}]$.
Since gate oxide and interchannel metal gates fill the space between two successive

nanosheets, these ones result to be vertically distanced by a quantity T_{sp}.
Inner spacers of length $L_{sp,in}$ are inserted between gates and source/drain structures, as illustrated in fig. 1.2. Due to etching process limitations, these low-k insulating layers can be not perfectly aligned with the main spacers having length L_{sp}.

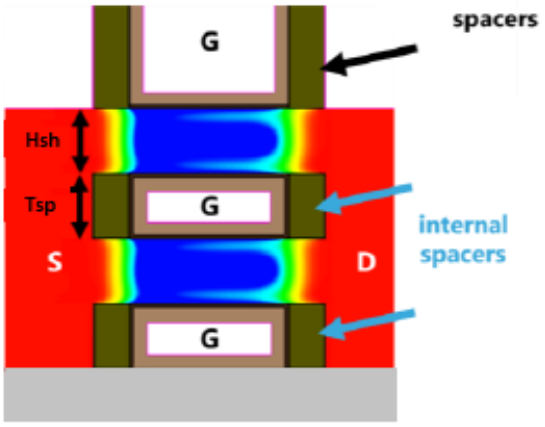


Figure 1.2: Multi-stacked NSGAAFET

Chapter 2

BSIM-CMG model for MugFETs

The aim of this second chapter is to present and analyse the industry-standard compact multigate model BSIM-CMG, developed at Berkeley University.
In section 2.2, the core part is illustrated, stressing its capability to capture the behaviour of different multigate devices, including a basic version of nanosheet gate-all-around field transistor.
Then, in sections 2.3 and 2.4, the attention moves to the analysis of parasitics modeling, since they acquire huge importance in very scaled devices.

2.1 Model overview

Multi-gate field-effect transistors (Mug-FETs) of different shapes can be modelled using the surface potential-based model BSIM-CMG, which garantees fast speed, numerical robustness and good accuracy.
Its strength is the versatiliy since it results valid for Mug-FETs of different shapes such as Double gate (DG) FinFET, triple gate (TG) FinFET, quadruple gate (QG) GAAFET (basic version of the NSGAAFET) and cylindrical gate (CG) GAAFET.
BSIM-CMG model is composed of two main components: the first, which is the basic core section providing the charge and drain current models, and the second, which includes a set of real-device effects submodules.
The core model is obtained by solving Poisson's equation using a long channel assumption, which is the Gradual Channel Approximation (GCA), and assuming Boltzmann's statistics for the carriers. All the other physical effects are neglected.
The second section contains the advanced physical effects which are later added, as correction terms, to the core part in order to obtain an accurate modeling of a real device.
The main real-device effects included are: short channel effects, channel length modulation, quantum mechanical effects, geometrically scalable parasitic resistances

and capacitances, vertical and horizontal electric-field dependent mobility, temperature dependence, current saturation, self-heating, gate leakage, etc [7].

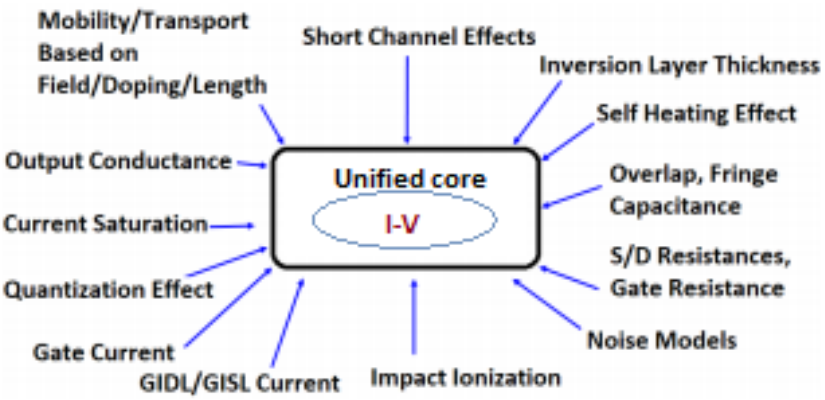


Figure 2.1: BSIM-CMG general structure [7].

2.2 Core section: Unified charge model

The core section represents the basic module on which the whole BSIM-CMG model is founded. It is derived from the resolution of Poisson's equation, under the use of GCA condition and Boltzmann's statistics, for two specific cases: DG- and CG-FET. This is done with the aim to obtain two accurate charge models for the two aforementioned structures and, finally, to generalize them to a universal charge model for other Mug-FETs.

- Solving Poisson's equation for double gate FinFET brings to the expression 2.1, which relates the mobile electron charge and the applied gate voltage [3].

$$V_G - V_{FB} + Q_{d,dg}\frac{t_{ox}}{2\epsilon_{ox}} - V = -Q_{e,dg}\frac{t_{ox}}{2\epsilon_{ox}} + v_t \cdot ln\frac{Q_{e,dg}(Q_{e,dg}+Q_{d,dg})/(4v_t\epsilon_{ch}/T_{fin})}{q\frac{n_i^2}{N_{ch}}T_{fin}\Big[1-exp\Big(\frac{T_{fin}}{4v_t\epsilon_{ch}}(Q_{e,dg}+Q_{d,dg})\Big)\Big]} \tag{2.1}$$

where V_G is the gate voltage, V_{FB} is the flat-band voltage, V is the electron quasi fermi potential and v_t is the thermal voltage. $Q_{e,dg}$ and $Q_{d,dg}$ are the inversion charge and depletion charge per unit area, respectively. N_{ch} is the doping concentration in the channel, n_i is the intrinsic carrier concentration and ϵ_{ch} is the dielectric constant of the channel. ϵ_{ox} is the dielectric constant

of the gate oxide and t_{ox} is its thickness. The fin has width T_{fin} and height H_{fin}.

- Similarly, for cylindrical gate FET, the charge model derived from the solution of Poisson's equation is reported in 2.2 [3].

$$V_G - V_{FB} + Q_{d,cg}\frac{ln(1+t_{ox}/R)}{2\pi\epsilon_{ox}} - V = -Q_{e,cg}\frac{ln(1+t_{ox}/R)}{2\pi\epsilon_{ox}} + v_t \cdot ln\frac{-Q_{e,cg}}{q\frac{n_i^2}{N_{ch}}\pi v_t R^2} + v_t \cdot ln\left[\frac{-(Q_{e,cg}+Q_{d,cg})/4\epsilon_{ch}\pi}{1-exp\left(\frac{Q_{e,cg}+Q_{d,cg}}{4\epsilon_{ch}\pi v_t}\right)}\right] \tag{2.2}$$

where $Q_{e,cg}$ and $Q_{d,cg}$ are the inversion charge and depletion charge per unit length, respectively. R is the channel radius.

By carefully analysing equations 2.1 and 2.2, it can be noticed that they are applicable to derive the charge model for other multi-gate structures.
As a matter of fact, the two equations can be generalized and written in the same closed form, which is the one implemented in BSIM [4]. It is reported in the following, with all the terms normalized with respect to the thermal voltage v_t.

$$v_G - v_o - v_{ch} = -q_m + ln(-q_m) + ln\left(\frac{q_t^2}{e^{q_t} - q_t - 1}\right) \tag{2.3}$$

with:

$$v_o = v_{FB} - q_{dep} - ln\left(\frac{2qn_i^2 A_{ch}}{v_t C_{ins} N_{ch}}\right) \tag{2.4}$$

$$q_t = (q_m + q_{dep}) \cdot \frac{A_{ch}C_{ins}}{\epsilon_{ch}W_{eff}^2} \tag{2.5}$$

where q_m and q_{dep} are the normalized inversion (mobile) and depletion charges, respectively.

Equation 2.3 defines the mobile carrier concentration in the channel for all bias conditions in a continuous and smooth manner, crucial for circuit simulation [4].
It is important to notice that, in equation 2.3, there are no more parameters like radius of wire or thickness of fin but the generalized parameters A_{ch} (area of the channel), N_{ch} (doping in the channel), W_{eff} (channel effective width) and C_{ins} (insulator capacitance per unit length).
As a consequence, the expression 2.3 can be used as a universal charge model for multi-gate field-effect transistors of different shapes, by mapping the proper device parameters.
On the basis of the chosen Mug structure, the device parameters are computed in the model with the formulas reported in table 2.1. The dimensions (width and height) of the Mug-FETs are referred as T_{fin} and H_{fin}, using a Finfet-like terminology.

Table 2.1: Model parameters for different Mug-FETs [7].

	DG	TG	QG	CG
W_{eff}	$2H_{fin}$	$2H_{fin}+T_{fin}$	$2[H_{fin}+T_{fin}]$	2πR
C_{ins}	$W_{eff} \cdot \frac{\epsilon_{ox}}{t_{ox}}$	$W_{eff} \cdot \frac{\epsilon_{ox}}{t_{ox}}$	$W_{eff} \cdot \frac{\epsilon_{ox}}{t_{ox}}$	$\frac{2\epsilon_{ox}\pi}{ln(1+\frac{t_{ox}}{R})}$
A_{ch}	$H_{fin}T_{fin}$	$H_{fin}T_{fin}$	$H_{fin}T_{fin}$	$H_{fin}T_{fin}$

2.3 Sub-modules: Parasitic S/D resistances

Among the various real-device effects submodules that can be added to the core, an important role is occupied by the modeling of parasitic resistances. Indeed, they become even more important as the device scaling goes on.
BSIM-CMG model divides the total source/drain resistance $R_{s/d}$ of a Mug-FET (DG-, TG-, QG- or CG-FET) into two main components: the diffusion resistance R_{geo}, bias independent component including the contact R_{con} and the spreading R_{sp} contributions, and the extension resistance R_{ext}, which is a bias dependent component.
They are all considered in series, as schematized in fig. 2.2, so the total $R_{s/d}$ is modeled as the summation of these three resistances.

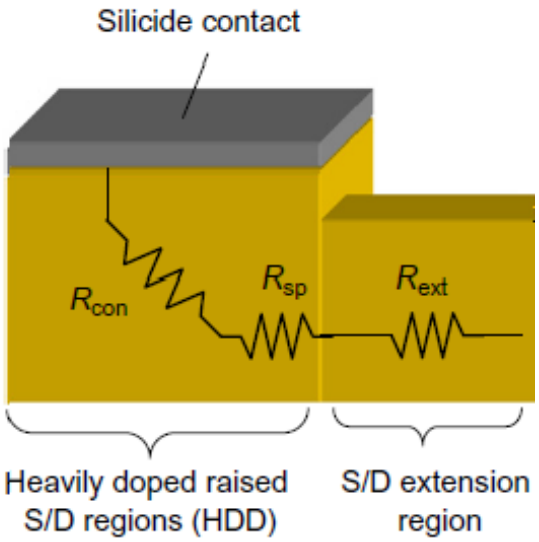


Figure 2.2: Various components of the source-drain resistance of a Mug-FET (DG-, TG-, QG- or CG-FET) [5].

For the computation of the diffusion resistance, the physically derived model is

chosen, among the two available in the model, and adopted. It allows to capture the complex dependency of both R_{con} and R_{sp} on the device geometry, as described in the following.

- Contact resistance accounts for both the resistance of the source/drain region and the one of the silicon/silicide interface. As a consequence, R_{con} mainly depends on the doping level of S/D regions and on the quality of the interface. Its computation is based on the transmission line model, which brings to the following final expression [6].

$$R_{con} = L_T \cdot \frac{\rho_{SD}}{A_{rsd}} \cdot \coth \alpha \tag{2.6}$$

where ρ_{SD} the S/D resistivity and L_T is the so-called penetration length, that is the distance over which the current occurs under the contact. Its resistivity, in $\Omega \cdot \mathrm{m}^2$, is indicated as ρ_c.

$$L_T = \sqrt{\frac{\rho_c \cdot A_{rsd}}{FP \cdot \rho_{SD}}} \tag{2.7}$$

The parameter α is, instead, the ratio between the S/D length (LRSD) and the penetration length L_T ($\alpha = \frac{LRSD}{L_T}$).
Finally, A_{rsd} is the area of the source/drain structure.

$$A_{rsd} = FP \cdot H_{fin} + T_{fin} \cdot H_{epi} + C_r \cdot (FP - T_{fin}) \cdot H_{epi} \tag{2.8}$$

A_{rsd} is computed by considering a source/drain geometry like the one reported in fig. 2.3. The width of S/D corresponds to the fin pitch FP, while H_{epi} represents the height of S/D above the top of the fin.
The structure is not perfectly rectangular but presents two corners on the top, whose shape can be modified by acting on C_r parameter. For triangular corners, C_r must be set to 0.5.

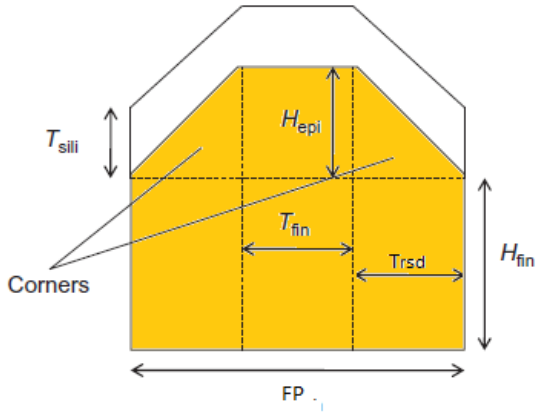


Figure 2.3: Source/drain epitaxial structure.

Equation 2.6 is obtained assuming that just the top part of source/drain structures is covered by silicide.
BSIM-CMG model offers also the possibility to cover with silicide the frontal and rear sections. This can be done by modifying equation 2.6, through the addition of the η parameter [6]:

$$R_{con} = L_T \cdot \frac{\rho_{SD}}{A_{rsd}} \cdot \frac{\cosh\alpha + \eta \cdot \sinh\alpha}{\sinh\alpha + \eta \cdot cosh\alpha} \tag{2.9}$$

- Spreading resistance takes into account the increase in S/D resistance due to the current spreading from the thin source/drain extension (SDE) to the large S/D structures and viceversa.
  As a matter of fact, when current flows from SDE into the S/D region, it spreads out gradually crowding inside the S/D region in the so called spreading region. This phenomenon is also known as current crowding [5].
  Assuming a costant spreading angle θ_{sp}, fixed to 55 degrees, the final formula implemented in BSIM-CMG for R_{sp} is given by [6]:

$$R_{sp} = \frac{\rho_{SD} \cdot cot(\theta_{sp})}{\sqrt{\pi}} \cdot \left(\frac{1}{\sqrt{A_{fin}}} - \frac{2}{\sqrt{A_{rsd}}} + \frac{\sqrt{A_{fin}}}{A_{rsd}} \right) \tag{2.10}$$

  where A_{fin} is the area of the S/D extension, which is simply equal to the product $H_{fin} \cdot T_{fin}$ for the Mug-FETs taken into account in the model.

Modeling of the extension resistance R_{ext} requires a different kind of analysis.
Since it represents the resistance in the extension region under the spacers, its

value depends on the doping profile inside the SDE which, in turn, varies on the basis of the process condition. Futhermore, R_{ext} is also influenced by the surface accumulation of the charge due to the fringe field originating from the gate.
Therefore, the extension resistance is a technology- and bias-dependent component. To simplify the computation, some assumptions about the spacer configuration and the doping profile are made [5]:

1. the total spacer of length L_{sp} is assumed to consist of an offset spacer and an effective spacer;

2. the doping is supposed to be uniform under the effective spacer but it decays, following a Gaussian profile, under the offset spacer.

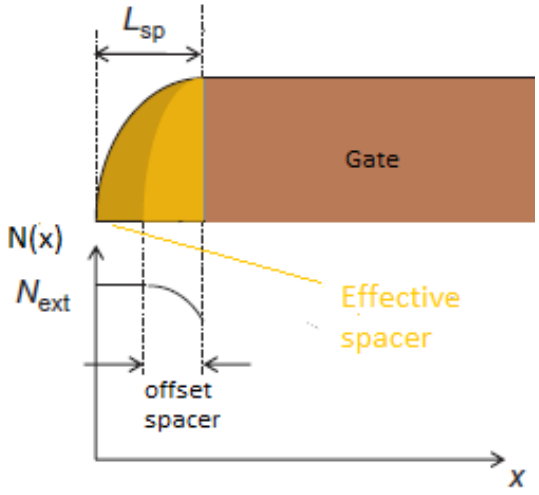


Figure 2.4: Spacer configuration and doping profile taken into account for R_{ext} modeling [5].

Under these hypotheses, R_{ext} can be considered composed of:

- a bias-dependent accumulation resistance R_{acc}, due to the charge collected at the surface of the extension, induced by gate fringe fields. This accumulation region is assumed to be under the offset spacer and also partially under the effective spacer;

- two bias-independent bulk resistance components: $R_{s/de1}$, underneath the surface accumulation region of the SDE, and $R_{s/de2}$, in the uniformly doped region, so under the effective spacer, but far away from the surface accumulation part of the extension (effective length $L_{sp} - \Delta L_{s/de}$).

Combining these three contributions into the resistive network of fig. 2.5, the final expression implemented in BSIM is obtained:

$$R_{ext} = \frac{\frac{R_{s/de10}}{H_{fin} \cdot T_{fin}}}{1 + \frac{R_{s/de10}}{R_{acc0} \cdot T_{fin}}(V_{gs/d} - V_{fbsd})} + \frac{R_{s/de20} \cdot (L_{sp} - \Delta L_{s/de})}{H_{fin} \cdot T_{fin}} \tag{2.11}$$

where R_{acc0}, $R_{s/d10}$ and $R_{s/d20}$ are technology-dependent parameters that must be extracted from a physical model.

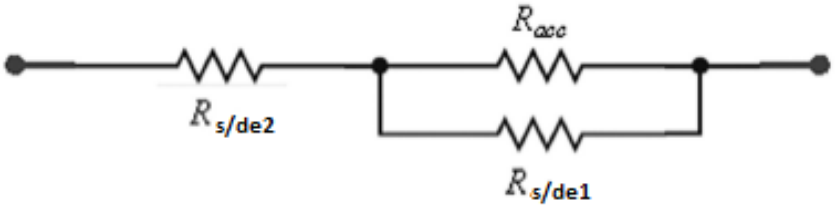


Figure 2.5: Circuit schematic for the computation of the source/drain extension resistance.

2.4 Sub-modules: Parasitic capacitances

In order to accurately evaluate AC performance of a Mug-FET, also the modeling of parasitic capacitances is added to the core.
For a SOI Mug-FET, BSIM-CMG includes three contributions: a bias-independent fringe capacitance, a bias-dependent overlap capacitance and a substrate capacitance.

2.4.1 Fringe capacitance

Fringe capacitance $C_{FR,geo}$ arises due to the close proximity of Mug-FET elements such as gate, extensions, S/D regions.
For the computation, the physically based model is chosen, among the three available in BSIM. It is the most accurate, since it addresses the complex dependencies of $C_{FR,geo}$ on the device geometry [6].
It splits fringe capacitance into three components:

- a top component $C_{FR,top}$ between the top part of the extension/of source-drain region and the gate;
- a side component $C_{FR,side}$ between the sidewall of the extension/of source-drain region and the gate;

- a corner component C_{corner} between the corner part of the S-D structure and the gate.

$C_{FR,top}$ calculation is based on a 2D fringe capacitance model which partitions top fringe capacitance into [5]:

- a top extension-to-gate capacitance $C_{fg,top}$;

- a top source/drain structure-to-gate capacitance $C_{cg,top}$, which is further split into two components ($C_{cg1,top}$ and $C_{cg2,top}$).

The three aforementioned contributions ($C_{fg,top}$, $C_{cg1,top}$, $C_{cg2,top}$) are expressed per unit of width. They have distinct electric field lines trajectories (fig. 2.6) which bring to different capacitive expressions, as reported in the following.

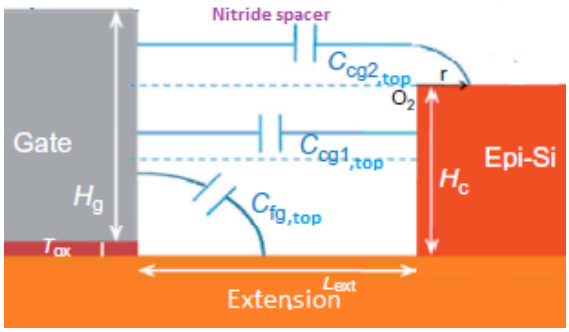


Figure 2.6: Top fringe components: $C_{fg,top}$, $C_{cg1,top}$, $C_{cg2,top}$ [5].

- The top fin extension-to-gate capacitance $C_{fg,top}$ is associated with electric field lines between the gate and the top surface of SDE, which is located under the spacer.
  The length d of each electric field line is the perimeter of a quarter ellipse, having a major axis l and a minor axis $tox + h$. The former lies along the extension length (L_{sp}) and ranges between 0 and L_{max}; the latter lies along the gate height (H_g) and ranges between t_{ox} and $t_{ox} + H_{max}$ [5].

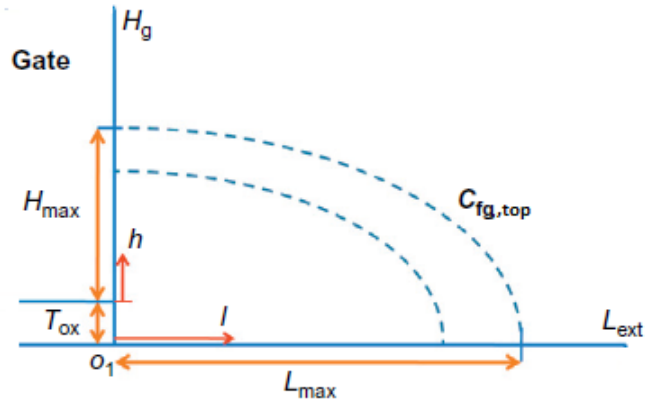


Figure 2.7: Electric field lines related to $C_{fg,top}$.

Once Euler approximation is applied to define d, the total $C_{fg,top}$ is computed by summing infinitesimal capacitors. Each of these is equal to $\Delta C = \epsilon_{sp} \cdot \frac{\Delta A}{d}$, where ϵ_{sp} is the dielectric constant of the spacer and ΔA is the area of the infinitesimal capacitor.

The infinitesimal summation is calculated for two different cases, leading to two expressions, per unit width, referred as:

- $C_{fg,topsat}$: obtained under the assumption that gate height (H_g) is greater than H_{max}, and, as a result, the capacitance does not change with H_g variation. This situation is denoted as the saturation condition.
- $C_{fg,toplog}$: obtained assuming that H_g is less than H_{max}. In this case, the entire inner part of the gate belongs to the top extension-to-gate region and so $C_{cg1,top}$=0. In this case, the capacitance is, to first order, a logarithmic function of H_g. This situation is denoted as the log condition.

Using a smoothing function with a fitting parameter δ to describe the transition from $C_{fg,topsat}$ to $C_{fg,toplog}$, the total expression, per unit of width, is obtained [5]:

$$C_{fg,top} = C_{fg,topsat} - \frac{(C_{fg,topsat} - C_{fg,toplog} - \delta) + \sqrt{(C_{fg,topsat} - C_{fg,toplog} - \delta)^2 + 4\delta C_{fg,topsat}}}{2} \tag{2.12}$$

where $C_{fg,topsat}$, $C_{fg,toplog}$ are linked to H_g, H_{max}, t_{ox}, L_{sp} and ϵ_{sp}.

- $C_{cg1,top}$ is a simple parallel plate capacitance between the gate and the top part of the S/D structure.

The final formula covers not only the case in which $H_g > H_{max}$, but also the possibility that $H_g < H_{max}$, which pushes $C_{cg1,top}$ to 0.
In addition, equation 2.13 takes into account that S/D structure could be not as tall as the gate.
The expression, per unit of width, is given by [5]:

$$C_{cg1,top} = \frac{1}{CNON} \cdot ln\left[1 + exp\left(CNON \cdot \epsilon_{sp} \cdot \frac{min(H_c, H_g + t_{ox}) - H_{max}}{L_{sp}}\right)\right] \tag{2.13}$$

where CNON is a fitting parameter and $H_c = H_{epi} + T_{sili}$ is the height of S/D structure over the fin, considering also a silicide layer of thickness T_{sili} above.

- $C_{cg2,top}$ is the capacitance due to the electric field lines which, starting from the gate, travel a distance L_{sp} horizontally and then follow a quarter cicle until reaching the top of S/D structure.
  Assuming that the quarter cicle has a radius r centered at the corner of the S/D contact, the ultimate expression is achieved by summing all the infinitesimal capacitances from 0 to a R value.
  The formula, per unit of width, is reported in equation 2.14 [5].

$$C_{cg2,top} = \frac{2\epsilon_{sp}}{\pi} \cdot ln\left(\frac{L_{sp} + 0.5\pi R}{L_{sp}}\right) \tag{2.14}$$

  where R is a geometry-dependent parameter linked to H_g, H_c and t_{ox}.

In conclusion, the total $C_{FR,top}$, at each drain/source side, is given by:

$$C_{FR,top} = C_{fg,top} \cdot T_{fin} + C_{cg1,top} \cdot T_{fin} + C_{cg2,top} \cdot T_{fin} \tag{2.15}$$

The three components are multiplied by T_{fin} since both extensions and the top part of S/D regions have width T_{fin}.

As the top fringe component, also $C_{FR,side}$ is based on a 2-D fringe capacitance model which partitions it into:

- a side extension-to-gate capacitance $C_{fg,side}$;
- a side source-drain structure to gate capacitance $C_{cg,side}$, which is further separated into two components ($C_{cg1,side}$ and $C_{cg2,side}$).

The three aforementioned contributions ($C_{fg,side}$, $C_{cg1,side}$, $C_{cg2,side}$) are expressed per unit of height.

- The side extension-to-gate capacitance $C_{fg,side}$ is associated with the electric field lines between the side of the SDE and the gate.
  The length of each electric field line is the perimeter of a quarter ellipse, having a major axis l and a minor axis $t_{ox} + W$. The former lies along the extension length (L_{sp}) and ranges between 0 and L_{max}; the latter lies along the gate wing (W_g) and ranges between t_{ox} and $t_{ox} + W_{max}$.

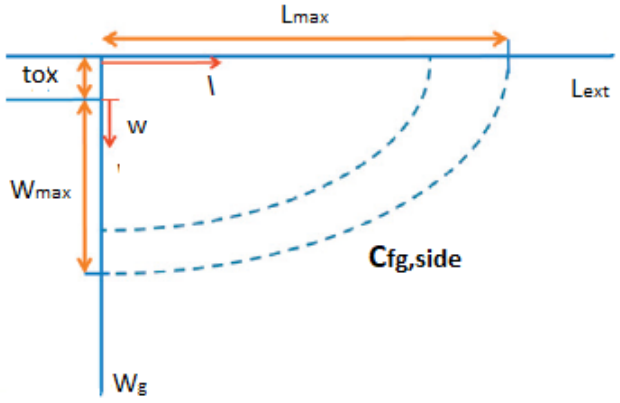


Figure 2.8: Electric field lines related to $C_{fg,side}$.

Since the computation procedure is equivalent to that of $C_{fg,top}$, the final formula for $C_{fg,side}$ has the same form. The only difference is that the gate height H_g is replaced with the gate wing W_g, and the upper limit H_{max} with W_{max}.
Equation 2.16 represents extension-to-gate capacitance, per unit of height:

$$C_{fg,side} = C_{fg,sidesat} - \frac{(C_{fg,sidesat} - C_{fg,sidelog} - \delta) + \sqrt{(C_{fg,sidesat} - C_{fg,sidelog} - \delta)^2 + 4\delta C_{fg,sidesat}}}{2} \tag{2.16}$$

where $C_{fg,sidesat}$ and $C_{fg,sidelog}$ are linked, in this case, to W_g, W_{max}, t_{ox}, L_{sp} and ϵ_{sp}.

- $C_{cg1,side}$ is a simple parallel plate capacitance between the gate sidewall and the lateral side of the extension.
  The same reasoning applied for $C_{cg1,top}$ brings to an equivalent $C_{cg1,side}$ expression, per unit of height.

$$C_{cg1,side} = \frac{1}{CNON} \cdot ln\left[1 + exp\left(CNON \cdot \epsilon_{sp} \cdot \frac{min(T_{rsd}, W_g + t_{ox}) - W_{max}}{L_{sp}}\right)\right] \tag{2.17}$$

In equation 2.17, since sidewalls are investigated, S/D lateral width T_{rsd} (fig. 2.3) and gate wing W_g are present. They substitute H_c and H_g, respectively.

$$T_{rsd} = \frac{1}{2} \cdot (FP - T_{fin}) \tag{2.18}$$

$$W_g = T_{rsd} - t_{ox} \tag{2.19}$$

- $C_{cg2,side}$ is the capacitance due to the electric field lines which, originating from the gate, travel a distance L_{sp} horizontally and then follow a quarter cicle, until reaching the side part of S/D structure.
  With the same procedure applied for $C_{cg2,top}$, but replacing H_g with W_g and H_c with T_{rsd}, expression 2.20 is obtained.
  This is the final formula, per unit of height:

$$C_{cg2,side} = \frac{2\epsilon_{sp}}{\pi} \cdot ln\left(\frac{L_{sp} + 0.5\pi R}{L_{sp}}\right) \tag{2.20}$$

  where R is a geometry-dependent parameter linked, in this case, to W_g, T_{rsd} and t_{ox}.

In conclusion, the total $C_{FR,side}$, for each drain/source side, is given by:

$$C_{FR,side} = C_{fg,side} \cdot H_{fin} + C_{cg1,side} \cdot H_{fin} + C_{cg2,side} \cdot H_{fin} \tag{2.21}$$

The three components are multiplied by H_{fin} since both extensions and the side part of S/D regions have height H_{fin}.

The corner capacitance C_{corner} is a simple parallel plate capacitance between the gate and the corners of the S/D structure.

$$C_{corner} = \frac{\epsilon_{sp}}{L_{sp}} \cdot A_{corner} \tag{2.22}$$

where $A_{corner} = (FP - T_{fin}) \times (H_{epi} \cdot C_r + T_{sili})$ is the total area of the two edges of S/D.

The three components C_{corner}, $C_{FR,top}$ and $C_{FR,side}$ are summed to get the total fringe capacitance $C_{FR,geo}$ for each drain/source side.

$$C_{FR,geo} = C_{corner} + C_{FR,top} + 2 \cdot CGEOE \cdot C_{FR,side} \tag{2.23}$$

where CGEOE is a fitting parameter and the multiplication by "2" takes into account both the sidewall contributions of each S/D structure.

2.4.2 Overlap and substrate capacitance

Source-drain overlap capacitance C_{ov} is a parasitic element that originates due to the encroachment of source-drain implant profile under the gate region.
As a matter of fact, the post-implant thermal processing steps cause lateral diffusion of dopants under the gate and so the overlap of source-drain extensions in the final device structure.
Depending on the bias voltages applied, the extensions can be in accumulation or in depletion condition. This leads to a different amount of overlap charge and so to different C_{ov} values.
In BSIM-CMG the accurate modeling, illustrated below, is implemented [6].

- At source side, the overlap capacitance is evaluated as the derivative of the overlap charge $Q_{gs,ov}$ with respect to V_{gs}. $Q_{gs,ov}$ is expressed as the sum of two terms: overlap charge in heavily doped source region and overlap charge in the ligthly-doped source extension.

$$Q_{gs,ov} = W_{eff} \cdot CGSO \cdot V_{gs} + W_{eff} \cdot CGSL \cdot V_s^* \tag{2.24}$$

  where:

  - CGSO is the overlap capacitance per unit of channel width between gate and the heavily doped source region. For a LDD device, CGSO can be set to 0 since most of the gate overlapped source region is the LDD;
  - CGSL is the overlap capacitance per unit of channel width between gate and lightly doped source region. It is equal to $\frac{\epsilon_{SiO2}}{EOT} \cdot l_{ov}$, where l_{ov} is the overlap length;
  - V_s^* is a function of the effective potential in the overlap region at source side ($V_{gs,ov}$), which, in turn, depends on V_{gs}.

- At drain side, the overlap capacitance is evaluated as the derivative of the overlap charge $Q_{gd,ov}$ with respect to V_{gd}. $Q_{gd,ov}$ is expressed as the sum of two terms: overlap charge in heavily doped drain region and overlap charge in the ligthly-doped drain extension.

$$Q_{gd,ov} = W_{eff} \cdot CGDO \cdot V_{gd} + W_{eff} \cdot CGDL \cdot V_d^* \tag{2.25}$$

  where CGDO and CGDL are the equivalent parameters at drain side. V_d^* is a function of the effective potential in the overlap region at drain side, which, in turn, depends on V_{gd}.

At this point, the last evaluation concerns substrate capacitances.
For a SOI multi-gate-FET, they are taken from source/drain to the substrate through the buried oxide [6].
The source-substrate component $C_{s,box}$ is given by:

$$C_{s,box} = C_{BOX} \cdot ASEO \tag{2.26}$$

The drain-substrate component $C_{d,box}$ is given by:

$$C_{d,box} = C_{BOX} \cdot ADEO \tag{2.27}$$

where ASEO and ADEO are the source/drain-to-substrate overlap area through BOX equal to the product LRSD · FP.
C_{BOX} is the buried-oxide capacitance per unit area:

$$C_{BOX} = \frac{\epsilon_{BOX}}{EOT_{BOX}} = \frac{\epsilon_{r,BOX} \cdot \epsilon_0}{EOT_{BOX}} \tag{2.28}$$

Chapter 3

Complete model for multi-stacked NSGAAFETs

The goal of this chapter is to present how the universal BSIM-CMG model has been modified to effectively capture a complete multi-stacked NSGAAFET. To this purpose, improved resistive and capacitive models are explained.

3.1 Additional key parameters

Among the four multi-gate field effect structures that can be captured using BSIM-CMG model, QG-FET is the one of main interest for the aim of this work.
Indeed, it represents a very basic version of the NSGAAFET, since it is composed of just one silicon nanosheet of width T_{fin} and height H_{fin}. It results that T_{fin} and H_{fin} parameters stand for the usual width (W_{sh}) and height (H_{sh}) of the nanosheet, respectively.

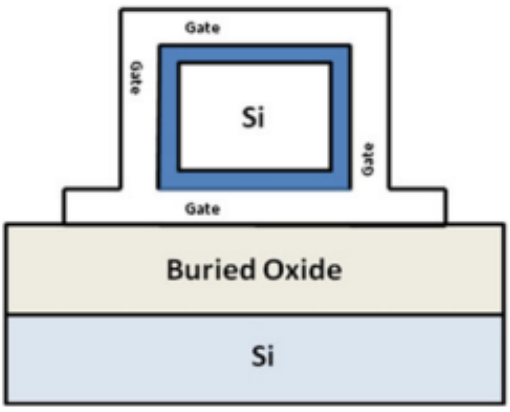


Figure 3.1: Quadruple-gate FET.

In order to effectively describe a real single and multi-stacked NSGAAFET, some key parameters must be included in the model:

- number of nanosheets per stack N_{sh} ;
- vertical spacing between nanosheets T_{sp};
- total height of the stack H_{stack}.

This allows to define a new effective width, given by:

$$W_{eff} = N_{sh} \times [2 \cdot T_{fin} + 2 \cdot H_{fin}] \tag{3.1}$$

and a total height of the stack equal to:

$$H_{stack} = N_{sh} \cdot H_{fin} + N_{sh} \cdot T_{sp} \tag{3.2}$$

Moreover, inner spacers dielectric constant ϵ_{spin} and length L_{spin} are added in order to model these structures.
The inclusion of all these new parameters is the starting point to modify the original modeling of parasitic resistances and capacitances.

3.2 Advanced parasitic S/D resistive network

In order to model the parasitic resistances of a multi-stacked NSGAAFET, the new resistance scheme of fig. 3.2 must be implemented.

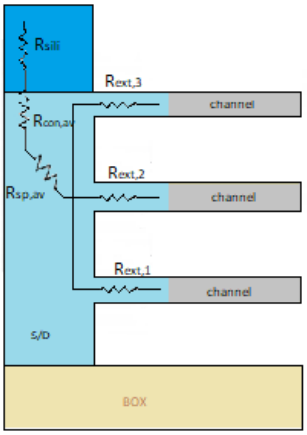


Figure 3.2: Cross-sectional view of a three-stacked nanosheet GAAFET.

- Contact resistance of the stacked structure $R_{con,av}$ must take into account not only the silicon/silicide interface contribution but also the resistance in the epitaxial S/D structure underneath the contact.
  Since the contact resides at the top of S/D epitaxial structure, carriers start from the top and flow towards the various channels. The penetration length of these current paths increases as the nanosheet considered is closer to the bottom. In order to capture this behaviour inside BSIM-CMG model, an average single contact resistance is modeled. More in details, $R_{con,av}$ considers an average current path corresponding to one half of the total height of the stack.
  $$R_{con,av} = L_{T,av} \cdot \frac{\rho_{SD}}{A_{rsd,av}} \cdot \coth \alpha_{av} \tag{3.3}$$
  where:
  $$A_{rsd,av} = FP \cdot H_{av} + T_{fin} \cdot H_{epi} + C_r \cdot (FP - T_{fin}) \cdot H_{epi} \tag{3.4}$$
  with the average height H_{av} equal to $H_{stack}/2$.
  Considering a η parameter different from zero, eq. 3.3 becomes:
  $$R_{con,av} = L_{T,av} \cdot \frac{\rho_{SD}}{A_{rsd,av}} \cdot \frac{\cosh \alpha_{av} + \eta \cdot \sinh \alpha_{av}}{\sinh \alpha_{av} + \eta \cdot cosh\alpha_{av}} \tag{3.5}$$

- Spreading resistance of the stacked structure $R_{sp,av}$ describes the resistance due to the current spreading from the S/D epitaxial structure to the S/D extensions.
  Considering that carriers spread into the various extensions following a path whose length depends on the channel position, the same discussion done for $R_{con,av}$ is applied.
  As a consequence, an average spreading resistance is modeled. It considers an average height H_{av} of the S/D structure and of the extension equal to $H_{stack}/2$.
  $$R_{sp,av} = \frac{\rho_{SD} \cdot cot(\theta_{sp})}{\sqrt{\pi}} \cdot \left(\frac{1}{\sqrt{A_{fin,av}}} - \frac{2}{\sqrt{A_{rsd,av}}} + \frac{\sqrt{A_{fin,av}}}{A_{rsd,av}} \right) \tag{3.6}$$
  $$A_{fin,av} = T_{fin} \cdot H_{av} \tag{3.7}$$

- Total Extension resistance $R_{ext,tot}$ of the structure is computed as the parallel combination of the extension resistive components related to each single nanosheet. All the contributions are considered equal since all the extensions are doped in the same way and are all subjected to the fringe fields of the two interchannel metal gates located above and below each nanosheet.
  For a three-stacked NSGAAFET:
  $$R_{ext,tot} = R_{ext,1} || R_{ext,2} || R_{ext,3} \tag{3.8}$$

Since no physical model in TCAD is available for stacked NSGAAFET until now, default values are taken for the evaluation of each extension resistance.

- Additional silicide resistance R_{sili} has been included in the model. It takes into account the resistance of the silicide contact, having thickness T_{sili}, width FP and length LRSD.
$$R_{sili} = \frac{\rho_{sili} \cdot T_{sili}}{FP \cdot LRSD} \tag{3.9}$$
where ρ_{sili} is the contact resistivity expressed in $\Omega \cdot m$.

3.3 Improved parasitic capacitive network

In order to capture the main parasitic capacitive contributions of a multi-stacked NSGAAFET, the original model illustrated in section 2.4 must be properly modified.

- For fringe capacitance of a multi-stacked NSGAAFET $C_{FR,geost}$, each of its three component (top, side and corner) must be accurately investigated.
Top fringe component $C_{FR,topst}$ takes into account the capacitive contributions between the top part of the upper extension/of source-drain structure and the gate, through the spacer material.
Since the upper part of the device structure is the same of the simple MugFeTs (DG-, TG-, QG-, CG-FET), top fringe component expression is unchanged.
$$C_{FR,topst} = C_{fg,top} \cdot T_{fin} + C_{cg1,top} \cdot T_{fin} + C_{cg2,top} \cdot T_{fin} \tag{3.10}$$
Side fringe component of a multi-stacked NSGAAFET $C_{FR,sidest}$ includes the capacitive contributions between the sidewall of the SDEs/of source-drain region and the gate side, though the spacer material.
In a NSGAAFET, not just one source/drain extension but N_{sh} extensions H_{fin}-tall are placed one upon the other. In addition, the side of S/D structure has height corresponding to the total stack H_{stack}.
It results that $C_{FR,sidest}$ is given by:
$$C_{FR,sidest} = C_{fg,side} \cdot (N_{sh} \cdot H_{fin}) + C_{cg1,side} \cdot H_{stack} + C_{cg2,side} \cdot H_{stack} \tag{3.11}$$
Corner fringe component $C_{cornerst}$ is taken between the corner of S/D structure and the gate, through the spacer material.
Assuming that source-drain edges are unchanged, the expression remains the usual one:
$$C_{cornerst} = \frac{\epsilon_{sp}}{L_{sp}} \cdot A_{corner} \tag{3.12}$$

The three components $C_{cornerst}$, $C_{FR,topst}$ and $C_{FR,sidest}$ are summed to get the total fringe capacitance $C_{FR,geost}$ for each drain/source side:

$$C_{FR,geost} = C_{cornerst} + C_{FR,topst} + 2 \cdot CGEOE \cdot C_{FR,sidest} \tag{3.13}$$

- Overlap capacitances $C_{ov,sst}/C_{ov,dst}$ arise due to the lateral diffusion of dopants from the SDEs towards the channel regions.
  For a NSGAAFET, extensions are N_{sh} in number and so equations 2.24 and 2.25 must be modified by inserting the new effective width of the multistacked device. Considering $W_{eff} = N_{sh} \times [2 \cdot T_{fin} + 2 \cdot H_{fin}]$, S/D overlap charges are:

$$Q_{gs,ovst} = W_{eff} \cdot CGSO \cdot V_{gs} + W_{eff} \cdot CGSL \cdot V_s^* \tag{3.14}$$

$$Q_{gd,ovst} = W_{eff} \cdot CGDO \cdot V_{gd} + W_{eff} \cdot CGDL \cdot V_d^* \tag{3.15}$$

  Source/drain overlap capacitances are obtained by deriving $Q_{gs,ovst}/Q_{gd,ovst}$ with respect to V_{gs}/V_{gd}. This corresponds to sum all the overlap components, related to the encroachment of the various extensions, to get the final one.

- Substrate capacitance, considering a SOI multistacked-NSGAAFET, is evaluated from source/drain towards the substrate through the BOX.
  Assuming that S/D to substrate overlap area (ASEO/ADEO) through BOX are equal:

$$C_{d,boxst} = C_{s,boxst} = C_{BOX} \cdot ASEO \tag{3.16}$$

- The additional inner spacer capacitance $C_{add,tot}$ has been included in the model. It is peculiar to the NSGAAFET structure.
  Indeed, in this novel device, metal gate, which is placed between nanosheets, and S/D regions face each other across the inner spacer. This gives origin to $C_{add,tot}$, which arises between the source/drain structures and the interchannel metal gates, through the inner spacer material.
  Using a parallel plate model, the inner spacer capacitance involving a single interchannel metal gate:

$$C_{add} = \frac{\epsilon_{spin}}{L_{insp}} \cdot (T_{sp} - 2t_{ox}) \cdot T_{fin} \tag{3.17}$$

  where ϵ_{spin} and L_{spin} are the dielectric constant and the length of the inner spacers, respectively. The height of the interchannel metal gate is obtained by subtracting from the spacing T_{sp} the two layers of gate oxide above and below. Considering that interchannel metal gates are N_{sh} in number, the total $C_{add,tot}$, for each source/drain side, is obtained by summing all the contributions related to each of them.

$$C_{add,tot} = \frac{\epsilon_{spin}}{L_{insp}} \cdot (T_{sp} - 2t_{ox}) \cdot T_{fin} \cdot N_{sh} \tag{3.18}$$

Chapter 4

DC performance of a HP 3-NSGAAFET

The goal of this chapter is to present the DC simulation results obtained by exploiting the new modified model of chapter 3 for a three stacked nanosheet gate-all-around field-effect-transistor (3-NSGAAFET).
In particular, in section 4.2, important device parameters, such as I_{dsat}, I_{sub}, SS, $V_{th,sat}$ and $DIBL$, are evaluated for a n-type three-stacked NSGAAFET.
Then, the analysis moves to the influence of process variations on DC performance and of the nanosheets size on the channel electrostatic control.

4.1 Simulation parameters

The 3-NSGAAFET under test is characterized by a SOI substrate.
This means that it presents a buried oxide layer (BOX) beneath the source/drain regions capable to stop bottom leakage currents.
Channel regions and source-drain structures have a doping concentration of 1×10^{16} cm^{-3} and 1×10^{21} cm^{-3}, respectively. S/D regions are supposed to have two triangular corners at the top ($C_r = 0.5$) and a titanium silicide layer on the top, on the front and on the rear.
Contact resistivity at S/D-silicide interface is fixed to 1.5×10^{-13} Ω·m^2.
Interfacial layer (IL) and high-k gate oxide are made of S_iO_2 and HfO_2, respectively. Equivalent oxide thickness (EOT) is 0.7 nm, which consists of 0.5 nm-thick IL and 1.5 nm-thick HfO_2.
Low-k spacers and inner spacers are made of SiOCN that has a dielectric constant of 5 (ϵ_{sp}=ϵ_{spin}).
Table 4.1 shows the values of the geometrical parameters of a NSGAAFET for sub-7 nm nodes.

Table 4.1: Geometrical parameters of NSGAAFET.

Parameters	Definitions	Values
L_g	Gate length	14 nm
H_{fin}=H_{sh}	Nanosheet height	7 nm
T_{fin}=W_{sh}	Nanosheet width	30 nm
T_{sp}	Vertical spacing	10 nm
FP	Fin pitch	56 nm
H_{epi}	Height of S/D above the stack	10 nm
$LRSD$	S/D structures length	17 nm
T_{sili}	Silicide thickness	10 nm
H_g	Gate Height	35 nm
L_{sp}	Spacer length	5 nm
$L_{sp,in}$	Inner spacer length	5 nm
l_{ov}	Overlap length	3.5 nm
N_{sh}	Number of nanosheet per stack	3

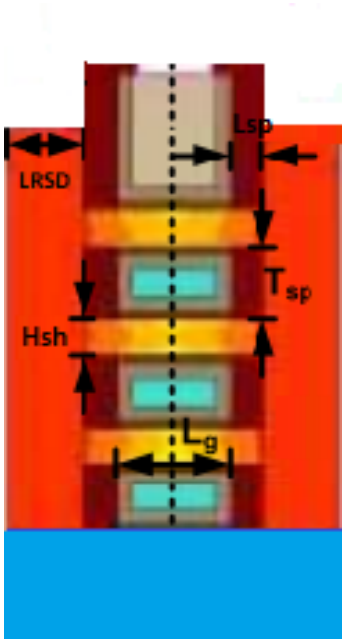


Figure 4.1: Three-stacked NSGAAFET.

4.2 DC characteristics

The NSGAAFET explored in this section is a n-type high performance (HP) device. It is characterized by the parameters of section 4.1, in particular H_{sh} = 7 nm and

W_{sh} = 30 nm, and it is composed of three nanosheets, as common in the current literature [11].
The DC analysis of this device has been performed on Cadence Virtuoso environment ®, using Spectre as simulator. A DC voltage source between gate and source (with parametric value V_{gs}) and another one between drain and source (with parametric value V_{ds}) have been inserted. The investigation starts with the simulation

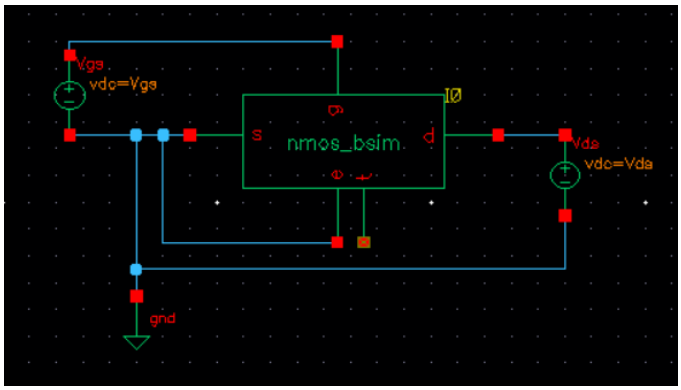


Figure 4.2: Schematic of N-type NSGAAFET.

of I_d vs V_{ds} characteristics for different V_{gs} values. Assuming a supply voltage V_{DD}= 0.7 V, the upper limit for V_{gs} corresponds to 0.7 V.

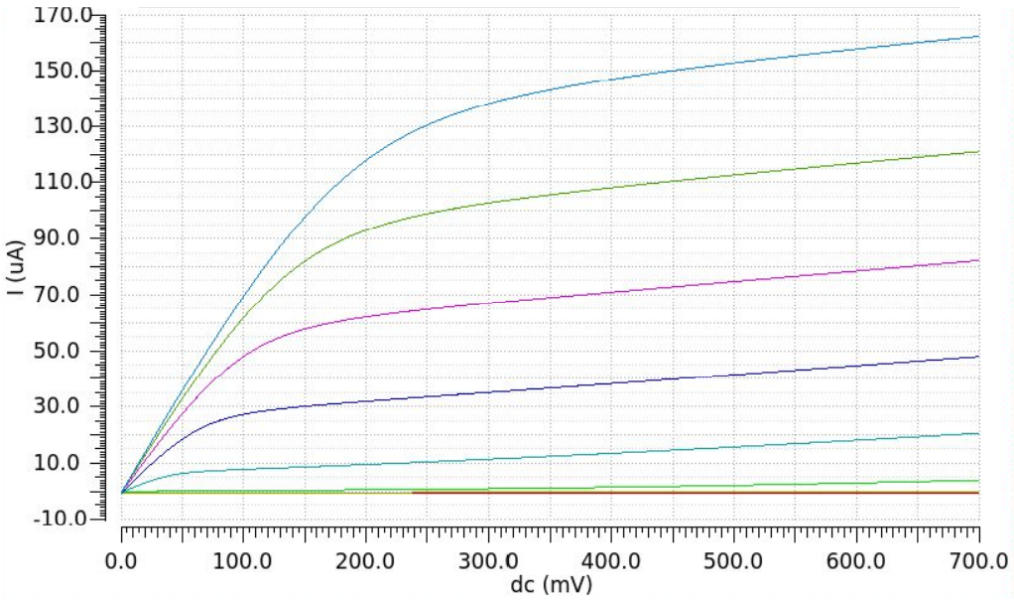


Figure 4.3: I_d vs V_{ds} characteristics for V_{gs} from 0 to 0.7 V.

The first value to be measured is I_{dsat}, defined as the value assumed by the drain current when both V_{gs} and V_{ds} are equal to V_{DD}. It results to be 162.38 µA.

This is more clearly visible in fig. 4.4, where I_d vs V_{gs} characteristic, for $V_{ds} = 0.7\,\mathrm{V}$, is reported.

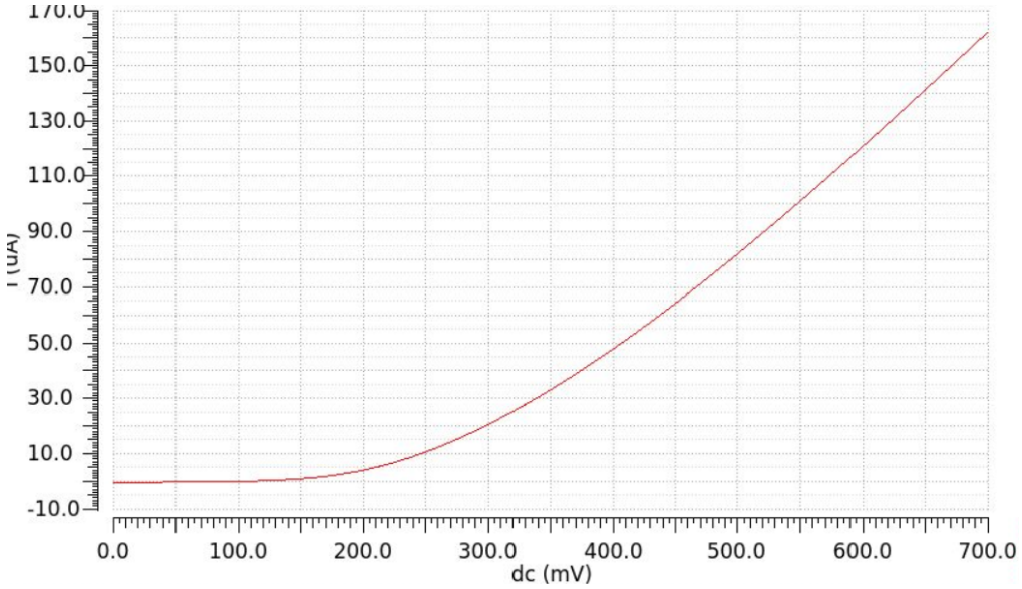


Figure 4.4: I_d vs V_{gs} characteristic for V_{ds}=V_{DD}, in a n-type 3-NSGAAFET.

If the attention moves to the subthreshold region, the I_{dsub} value can be computed. It is measured at very low V_{gs} (almost 0 V) and for $V_{ds} = V_{DD}$.
Being a HP device, it results, not surprisingly, to be around 15.19 nA.

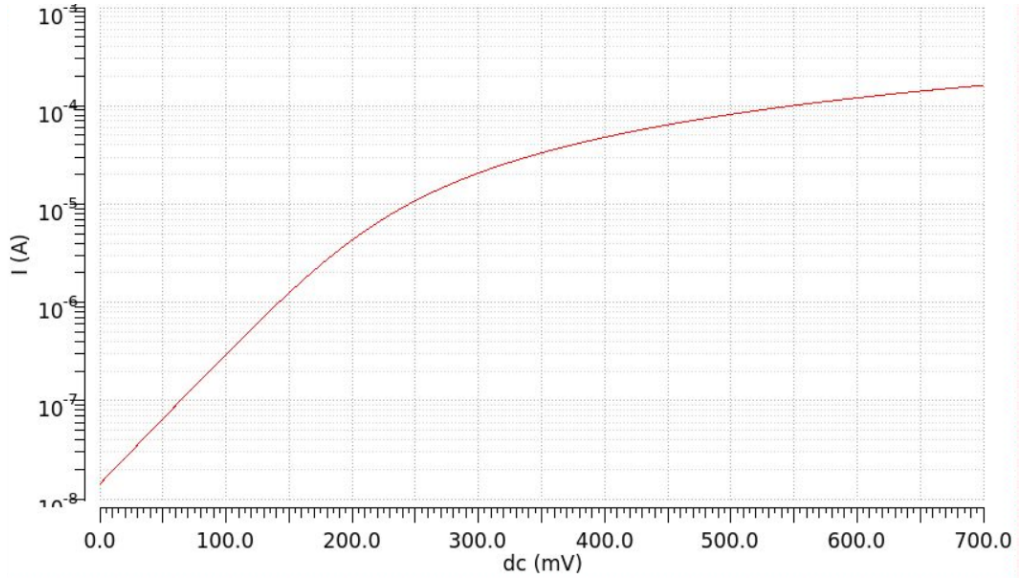


Figure 4.5: $\log(I_d)$ vs V_{gs} characteristic for V_{ds}=V_{DD}, in a n-type 3-NSGAAFET.

From fig. 4.5 another important figure of merit can be extracted. This is the subthreshold slope (SS), defined as the inverse of the slope of the $log(I_d) - V_{gs}$ characteristic in subthreshold region, for high V_{ds}.

$$SS = \left[\frac{d[log(I_d)]}{dV_{gs}}\right]^{-1} \tag{4.1}$$

It quantifies how many mV of V_{gs} are required to have a variation of the subthreshold current by one decade. As a consequence, it is a measure of the speed of the device to be switched on/off.
For the stacked structure under test, the extracted SS, at room temperature, is around 78 mV/dec.

Among the SCEs, drain-induced-barrier-lowering (DIBL) is one of the most relevant. Indeed, in scaled devices, the drain is so close to the source region that V_{ds} can effectively lower the source-channel barrier. This effect becomes increasingly important as V_{ds} rises, causing great V_{th} degradation at high V_{ds} values.
As a consequence, DIBL can be computed as the shift of the threshold voltage ΔV_{th} at two different V_{ds} values: 0.05 V (low drain bias) and 0.7 V (high drain bias corresponding to V_{DD}).

$$DIBL = \frac{\Delta Vth}{\Delta Vds} \tag{4.2}$$

In order to compute the threshold voltage V_{th}, the second derivative method is applied. It evaluates V_{th} as the V_{gs} value for which the derivative of the trasconductance gm ($= \frac{dI_d}{dV_{gs}}$) with respect to V_{gs} is maximum.
Fig 4.6 shows the two curves representing $\frac{d^2I_d}{dV_{gs}^2}$ vs V_{gs} for low (in red) and high (in yellow) V_{ds}.

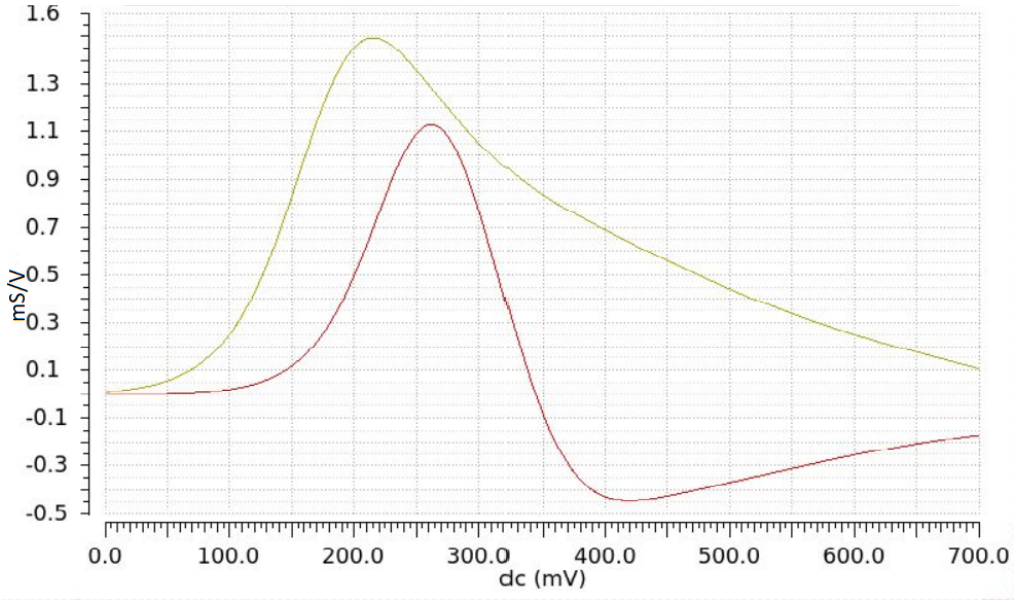


Figure 4.6: $\frac{d^2 I_d}{dV_{gs}^2}$ vs V_{gs} for V_{ds}=0.05 V (red curve) and for $V_{ds} = V_{DD}$ (yellow curve).

As expected in real scaled device, the two peaks are not aligned so DIBL effect is present.
In particular, since $V_{th,lin}$= 262 mV and $V_{th,sat}$ = 215 mV, DIBL appears reasonably low DIBL (72 mV/V). It proves that the presence of the gate all around the nanosheets ensures an optimum channel electrostatic control.

Table 4.2: Table of the main results obtained for the 3-stacked NSGAAFET.

I_{dsat}	I_{sub}	SS	DIBL
162.38 µA	15.19 nA	78 mV/dec	72 mV/V

4.3 Impact of process variations on I_{dsat}

Starting from the three-stacked NSGAAFET of section 4.2, some key parameters are changed inside a physically acceptable range. This causes a variation of the impact of resistive contributions on I_{dsat} value. To this purpose, for each new value, the saturation current is evaluated, being a fundamental figure of merit for a HP device.

The first element to be explored is the doping concentration of S/D structures (NSD). As it increases, ρ_{SD} reduces and, as a consequence, the two contributions

of diffusion resistance ($R_{sp,av}$ and $R_{con,av}$) lower.
This makes I_{dsat} sensibly rise: the upper bound is $\times 2.90$ times the lower one.

Table 4.3: I_{dsat} variation with source-drain doping concentration.

NSD	$2{\cdot}10^{25}$ m^{-3}	$5{\cdot}10^{25}$ m^{-3}	$2{\cdot}10^{26}$ m^{-3}	$5{\cdot}10^{26}$ m^{-3}	$1{\cdot}10^{27}$ m^{-3}
I_{dsat}	56.07 µA	82.92 µA	128.84 µA	148.89 µA	162.38 µA

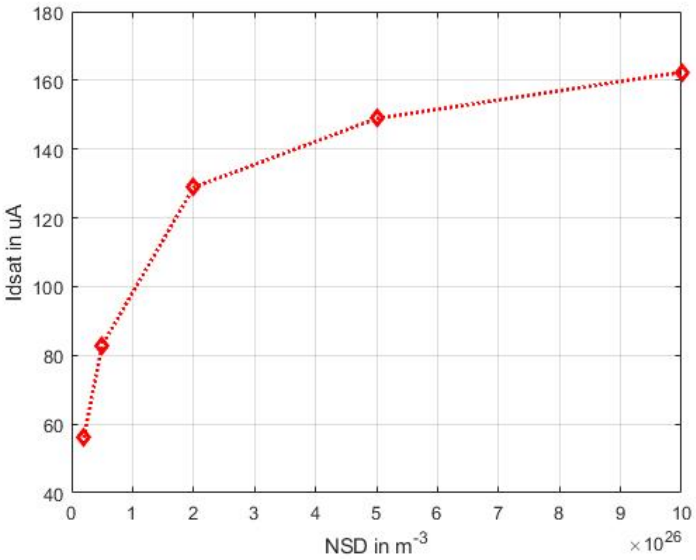


Figure 4.7: I_{dsat} variation with source-drain doping concentration.

Resistivity of the contact (ρ_c) has a crucial role, since it determines the value of contact resistance $R_{con,av}$, as well as the R_{sili} one.
Increasing ρ_c, I_{dsat} greatly deteriorates: a reduction by $\simeq 54\%$ is registered inside the range.

Table 4.4: I_{dsat} variation with contact resistivity.

ρ_c	$1.5{\cdot}10^{-13}$ Ω·m^2	$3.5{\cdot}10^{-13}$ Ω·m^2	$5.5{\cdot}10^{-13}$ Ω·m^2	$7.5{\cdot}10^{-13}$ Ω·m^2
I_{dsat}	162.38 µA	116.22 µA	91.17 µA	75.21 µA

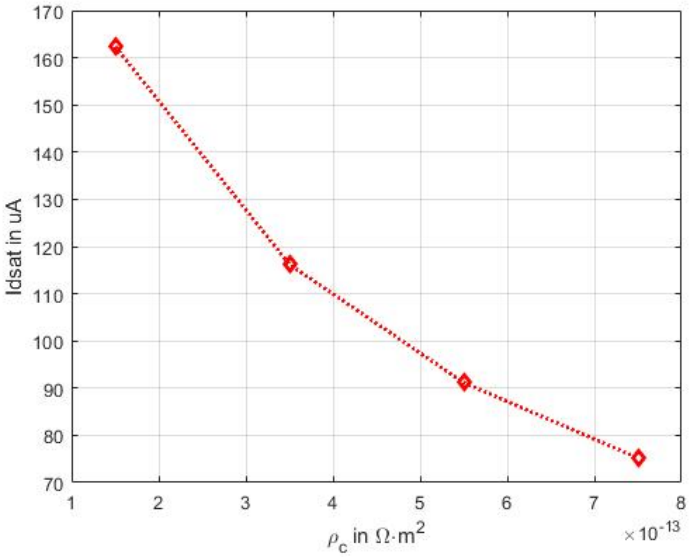


Figure 4.8: I_{dsat} variation with contact resistivity.

A silicide layer of thickness T_{sili} is originated once the salicidation process of the S/D structures has finished. Rising values of T_{sili} lead to higher R_{sili}, resulting in a small lowering of I_{dsat}.

Table 4.5: I_{dsat} variation with silicide thickness.

T_{sili}	5 nm	10 nm	15 nm	20 nm	25 nm
I_{dsat}	163.09 µA	162.38 µA	161.68 µA	160.99 µA	160.30 µA

The same T_{sili} variation considered above is taken into account. However, in this case, silicide layer is supposed to be present just on the top of S/D regions, not on the front, nor on the rear.
Compared to the previous situation, I_{dsat} decreases by 33%.

Table 4.6: I_{dsat} variation with silicide thickness present just on the top S/D structures.

T_{sili}	5 nm	10 nm	15 nm	20 nm	25 nm
I_{dsat}	109.09 µA	108.79 µA	108.49 µA	108.19 µA	107.89 µA

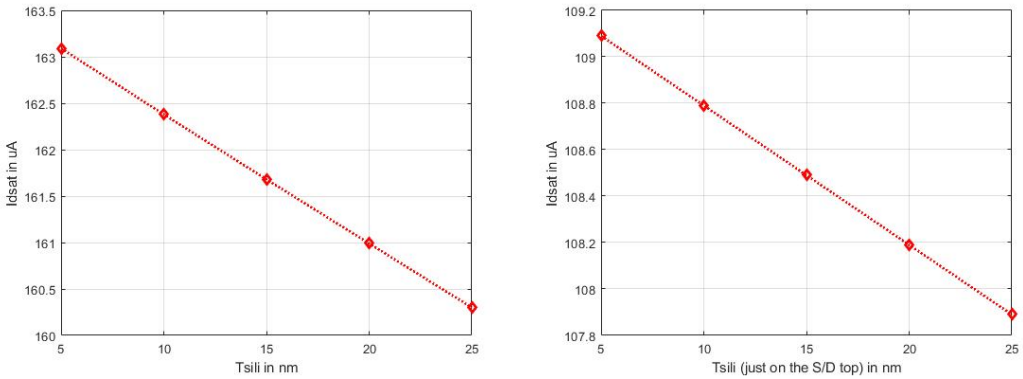


Figure 4.9: I_{dsat} variation with silicide thickness in two different situations: silicide just on S/D top (plot on the right) or also on front and rear (plot on the left).

The height of the S/D epitaxial structures above the stack (H_{epi}) contributes to the evaluation of S/D area (A_{rsd}). As H_{epi} rises, A_{rsd} increases and this ends up in the reduction of both R_{sp} and R_{con}. As a result, I_{dsat} increases.

Table 4.7: I_{dsat} variation with the height of S/D regions above the stack.

H_{epi}	5 nm	10 nm	15 nm	20 nm
I_{dsat}	157.99 µA	162.38 µA	166.21 µA	169.58 µA

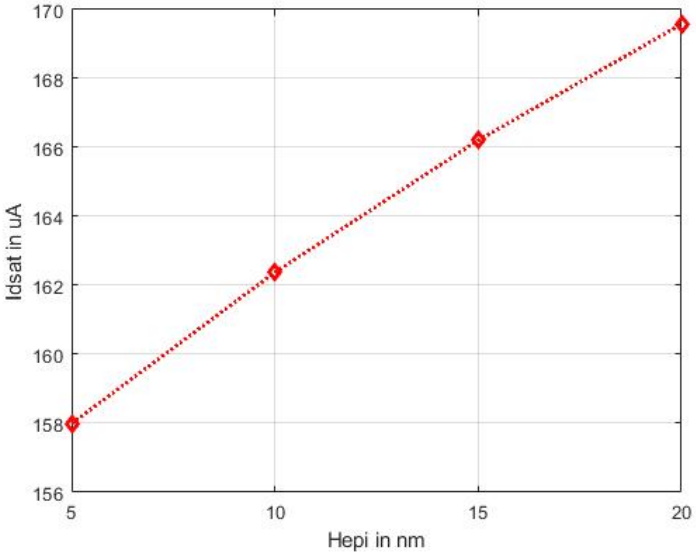


Figure 4.10: I_{dsat} variation with the height of S/D regions above the stack.

The length of S/D regions (LRSD) enters in $R_{con,av}$ computation.
As LRSD increases, α parameter (in eq.3.3) rises. The consequence is that $R_{con,av}$ slightly lowers and, in turns, I_{dsat} rises.

Table 4.8: I_{dsat} variation with the length of S/D regions.

LRSD	11 nm	14 nm	17 nm
I_{dsat}	160.88 µA	161.74 µA	162.38 µA

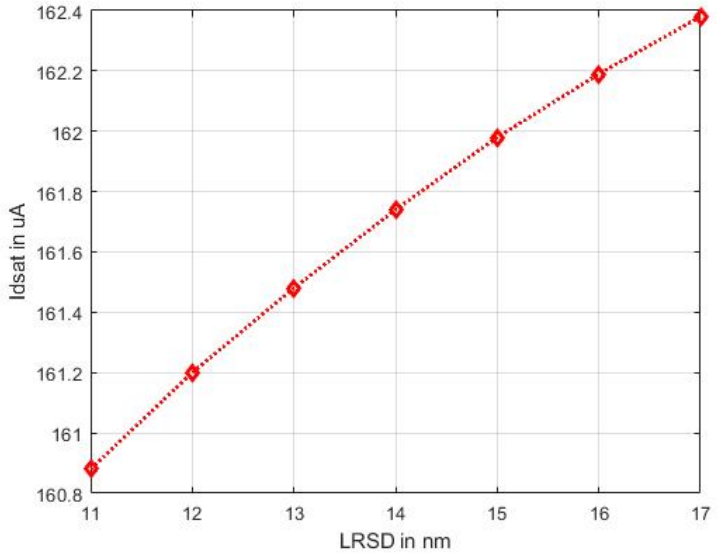


Figure 4.11: I_{dsat} variation with source/drain length.

In the computation done, the length of the extensions (L_{sp}) has no influence on I_{dsat}. Indeed, since no physical model is available for stacked NSGAAFET until now, default values are taken for $R_{ext,tot}$ evaluation. These values do not take into account the dependence of $R_{ext,tot}$ (so of I_{dsat}) on L_{sp}.

4.4 Impact of nanosheets size on electrostatics

Another interesting aspect to be investigated is how the nanosheets width affects the gate capability to control the channel in a NSGAAFET.
To this purpose, DIBL effect is evaluated for different W_{sh} values in the three-stacked structure with parameters of section 4.1.

Simulations reported below show the curves $\frac{d^2I_d}{dV_{gs}^2}$ vs V_{gs} for low and high drain bias. They are performed for a nanosheet width W_{sh} which varies from 15 nm to 35 nm.

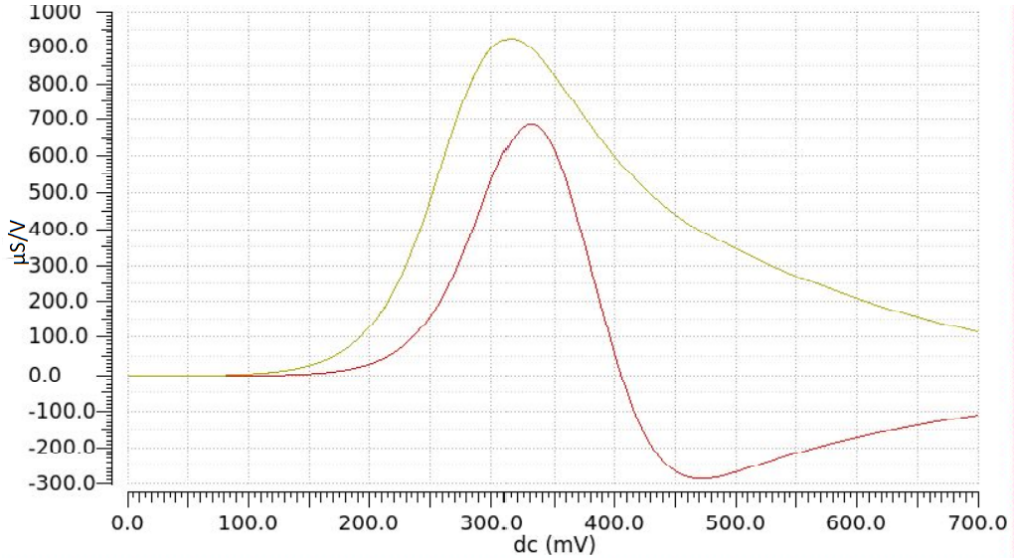


Figure 4.12: $\frac{d^2I_d}{dV_{gs}^2}$ vs V_{gs} for V_{ds}=0.05 V (red curve) and for $V_{ds} = V_{DD}$ (yellow curve) in a three-stacked NSGAAFET with W_{sh}=15 nm.

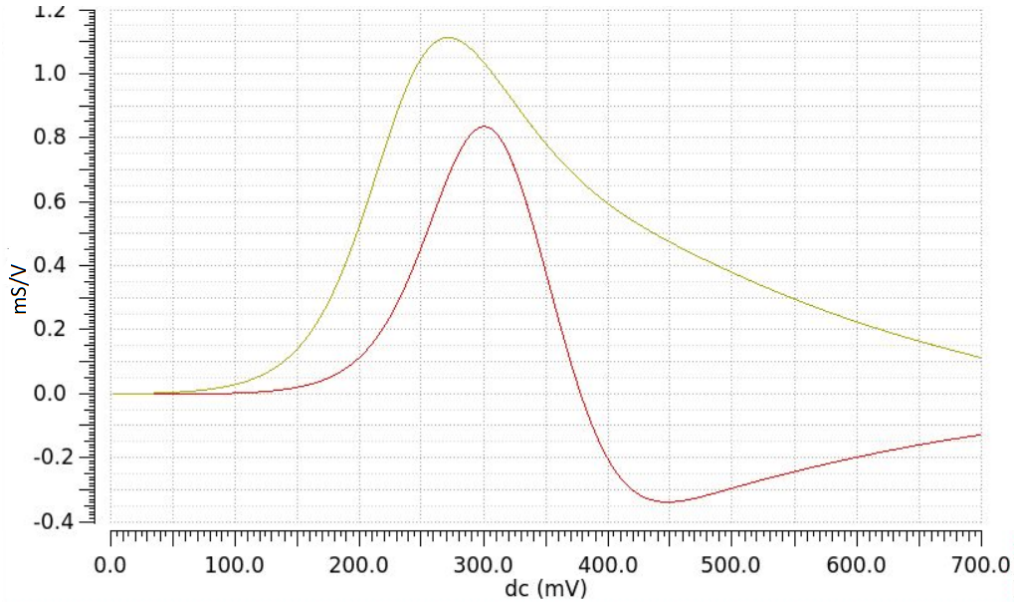


Figure 4.13: $\frac{d^2I_d}{dV_{gs}^2}$ vs V_{gs} for V_{ds}=0.05 V (red curve) and for $V_{ds} = V_{DD}$ (yellow curve) in a three-stacked NSGAAFET with W_{sh}=20 nm.

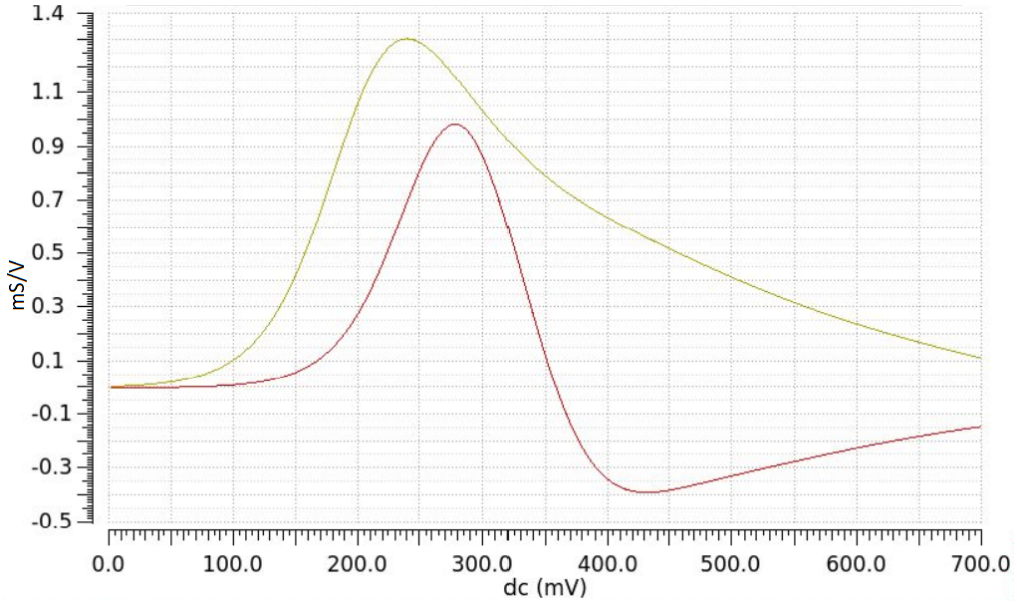


Figure 4.14: $\frac{d^2I_d}{dV_{gs}^2}$ vs V_{gs} for V_{ds}=0.05 V (red curve) and for $V_{ds} = V_{DD}$ (yellow curve) in a three-stacked NSGAAFET with W_{sh}=25 nm.

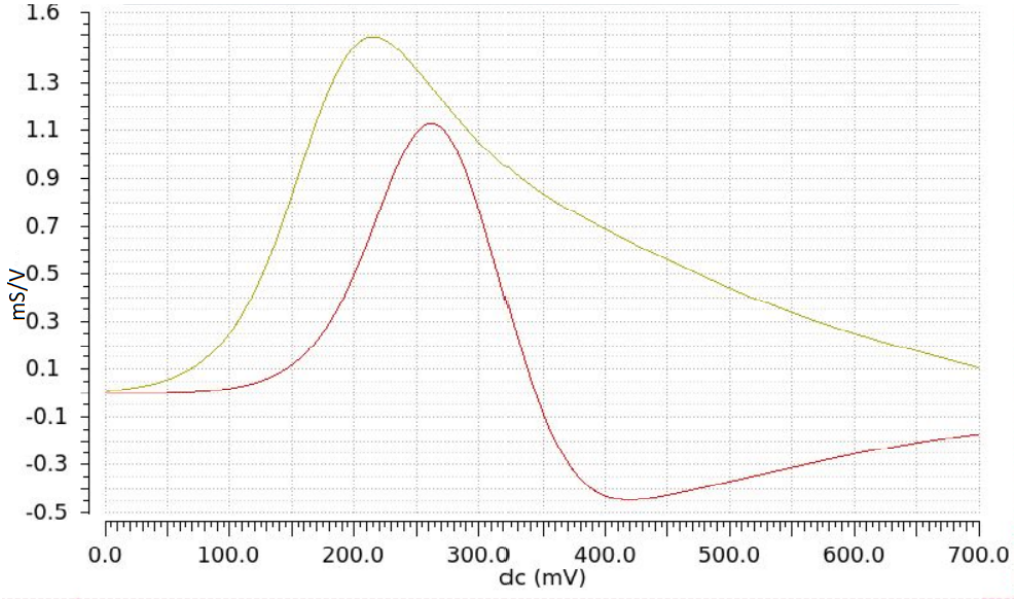


Figure 4.15: $\frac{d^2I_d}{dV_{gs}^2}$ vs V_{gs} for V_{ds}=0.05 V (red curve) and for $V_{ds} = V_{DD}$ (yellow curve) in a three-stacked NSGAAFET with W_{sh}=30 nm.

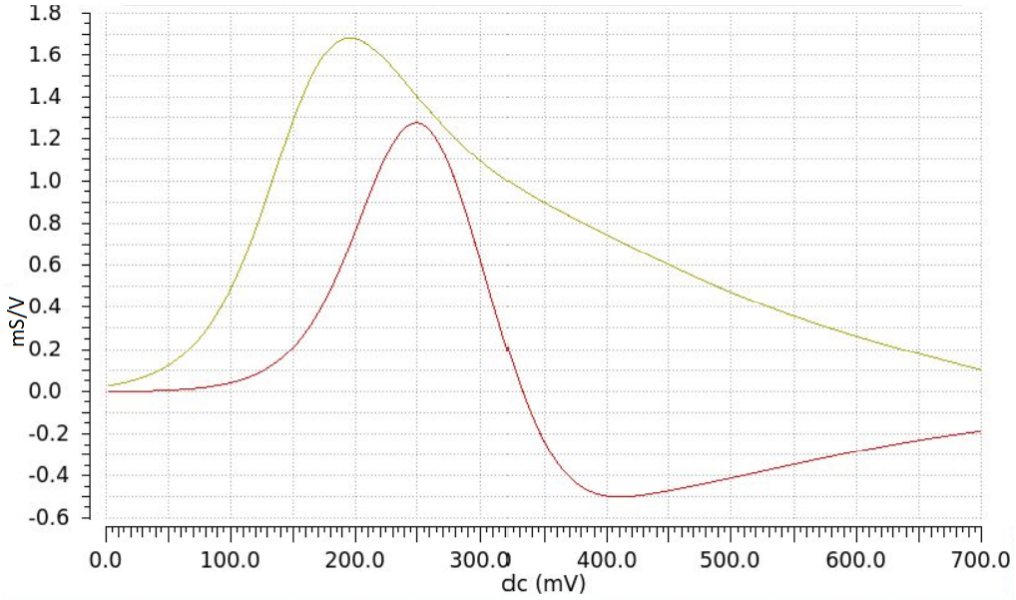


Figure 4.16: $\frac{d^2I_d}{dV_{gs}^2}$ vs V_{gs} for V_{ds}=0.05 V (red curve) and for $V_{ds} = V_{DD}$ (yellow curve) in a three-stacked NSGAAFET with W_{sh}=35 nm.

Results proove that, as W_{sh} rises, the peaks of the two curves become even more distanced and so DIBL effect increases. This is an expected behaviour since, by increasing W_{sh}, the sidewall gates have less impact on the channel charge. It derives that the gate reduces its control while the drain role becomes more evident.

Table 4.9: Table of DIBL results for the 3-stacked NSGAAFET with different W_{sh}.

W_{sh}	15 nm	20 nm	25 nm	30 nm	35 nm
DIBL	25 mV/V	43 mV/V	58 mV/V	72 mV/V	81 mV/V

Chapter 5

AC performance of a HP 3-NSGAAFET

The goal of this chapter is to study the AC performance of a three-stacked NS-GAAFET, by exploiting the new modified model implemented. To this purpose, a ring oscillator (RO) has been simulated, stressing the impact of the some process variations and of nanosheets width on the oscillation frequency.

5.1 Ring oscillator

The first logic circuit used to evaluate the AC performance of multi-stacked NS-GAAFETs is the ring oscillator. It is composed of an odd number of CMOS inverters: the output of each one is used as input for the next one. The last output is fed back to the first inverter.
The oscillation frequency depends on the number N and on the delay time τ of each inverter as follows:

$$f_{osc} = \frac{1}{2N\tau} \tag{5.1}$$

In particular, a five-stage ring oscillator composed of n-type and p-type three-stacked nanosheet GAAFETs has been simulated.
Both the types are characterized by the same parameters reported in section 4.1; the only difference involves the nanosheets width. It is fixed to 30 nm for the n-type, while the p-type is sized so that pull-up and pull-down networks of the inverters have almost equivalent effective resistance. As a consequence, following what explained in subsection 5.1.1, the p-type nanosheets width is set to 26 nm.

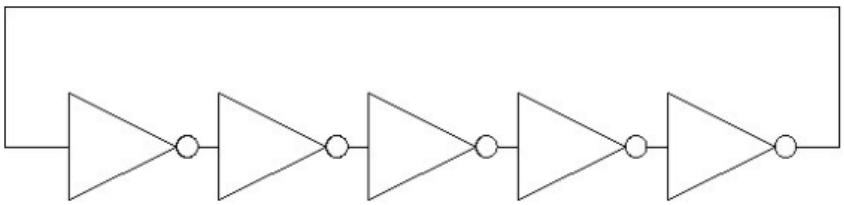


Figure 5.1: Five-stage ring oscillator.

In order to evaluate the oscillation frequency, a transient simulation has been carried out with Cadence Virtuoso, using Spectre as simulator. f_{osc} has been extrapolated by computing the frequency of the signal on the feedback branch, through the frequency function of Virtuoso Calculator.
The resulting output signal of the five-stage RO is reported in fig. 5.2 and has $f_{osc} = 23.33$ GHz.

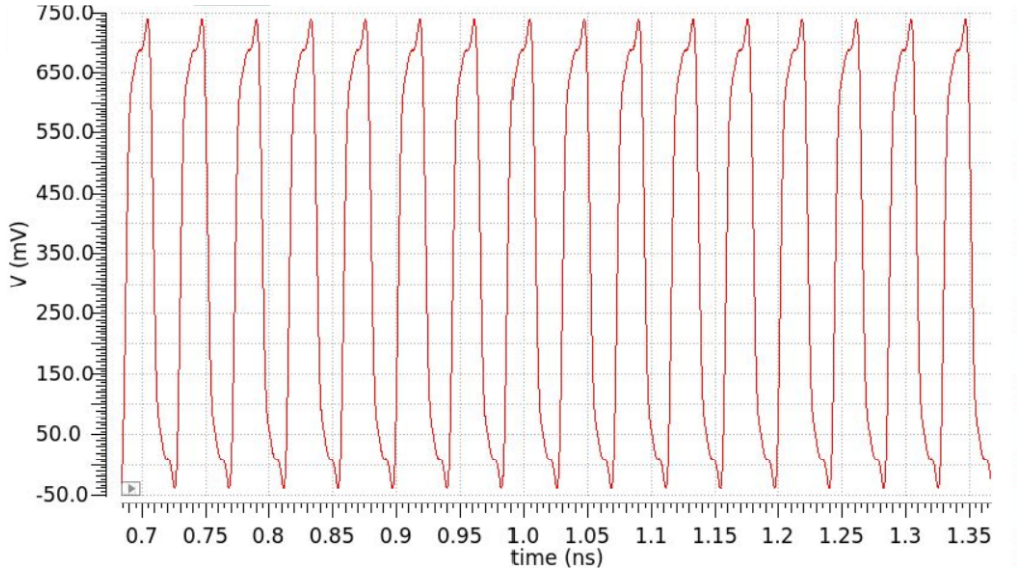


Figure 5.2: Output signal of the five-stage ring oscillator.

5.1.1 P-type 3-NSGAAFET sizing

Differently from traditional planar devices, drive current in a n-type NSGAAFET results to be lower than the one flowing in a p-type NSGAAFET. The reason lies in carriers mobility in highly doped source/drain structures and in the channels.

S/D regions are realized with silicon heavily doped with phosporus, for n-type, and with boron for p-type. At high dopant concentration, both electrons and holes

mobility is dominated by impurity ion scattering. In particular, this coulomb scattering probability is higher for electrons than for holes and it becomes more and more evident for high doping. The result is a more pronounced lowering of electron mobility, which ends up in a higher S/D resistivity for n-type NSGAAFET.

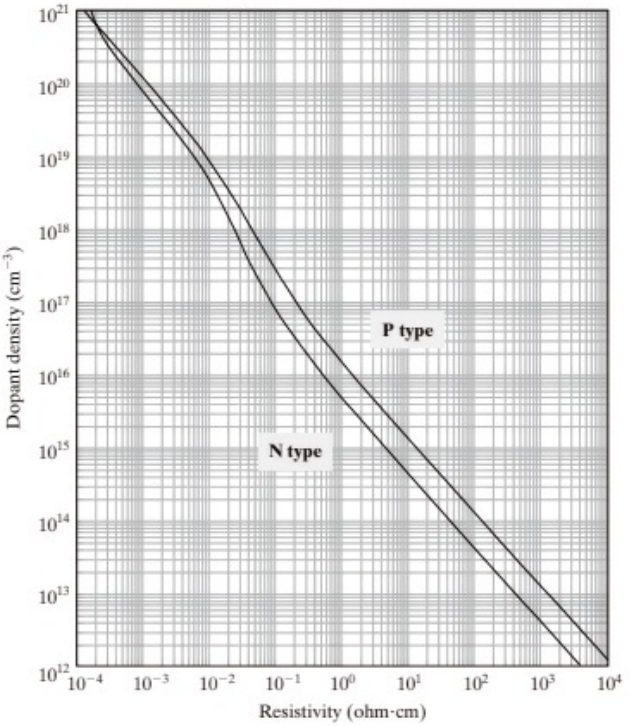


Figure 5.3: Dopant density vs resistivity of silicon at room temperature [10].

In the channel, mobility degradation due to an applied electric field follows distinct rules for electrons and holes. In particular, for high electric field, this degradation occurs at a greater rate for electrons than for holes (fig. 5.4).
This behaviour can be ascribed to surface roughness scattering, which is the main scattering mechanism at high electric field. The electron mobility is more sensitive to surface roughness than the hole mobility, because the inversion-charge centroid for electrons is closer to the silicon-gate oxide interface than for holes [8].

These two combined effects, related to mobility, explain the presence of higher drive current in p-NSGAAFET and are justified by the formulas implemented in the model.
Several simulations have been done considering a p-type 3-NSGAAFET with the same parameters of the n-type, reported in section 4.1.
It has been found that, for the same fixed W_{sh}, the measured current increase is $\simeq$ +19 % with respect to the n-type.

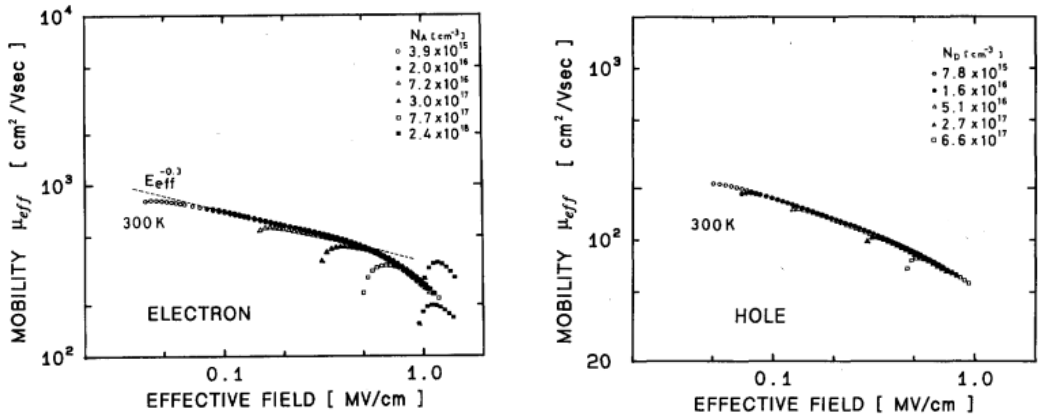


Figure 5.4: Electrons (on the left) and holes (on the right) mobility in inversion layer at 300K versus effective electric field, as a function of substrate doping [9].

In order to maintain the same quantity of current flowing, the p-type nanosheet width should be β times the n-type one, with $\beta \simeq 0.85$.
With 26 nm-nanosheet width for the p-type, it has been measured $|I_{dsat} = 163.45\,\mu A|$, which is almost equivalent to the value obtained with n-type $W_{sh} = 30$ nm.

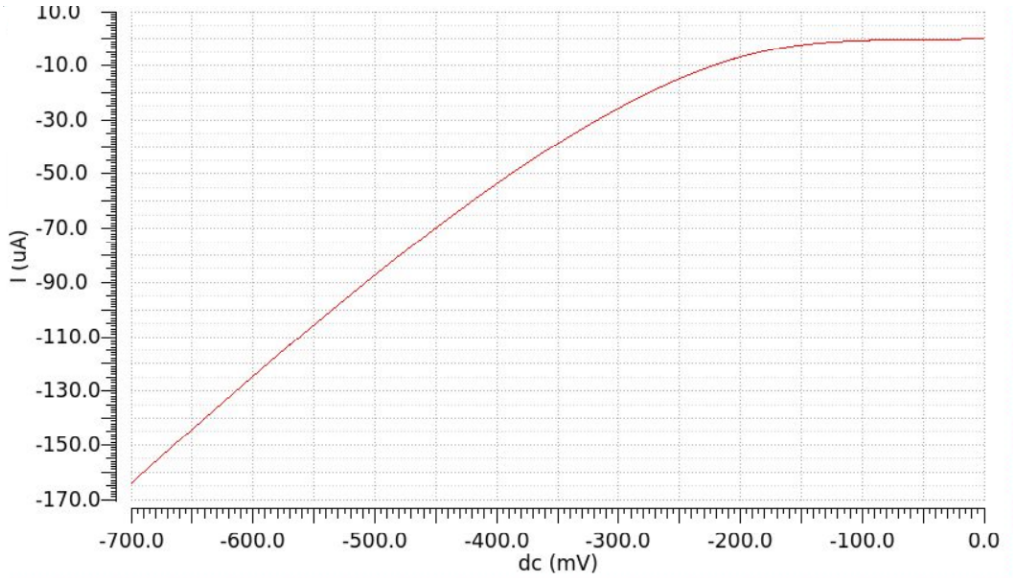


Figure 5.5: I_d vs V_{gs} characteristic for V_{ds}=V_{DD}, in a ptype 3-NSGAAFET with W_{sh}=26 nm.

5.2 Influence of process variations on f_{osc}

Starting from the five-stage ring oscillator of section 5.1, some key parameters of both ntype and ptype 3-NSGAAFET are changed inside a phisically acceptable range. This brings to a variation of the impact of capacitive contributions on f_{osc}. To this purpose, for each new value, the oscillation frequency is extrapolated, being a fundamental figure of merit for HP applications.

The first element to be explored is the material of spacers and inner spacers. Assuming that the same dielectric is used to realize both of them, it is varied between silicon oxycarbonitride ($\epsilon \simeq 5$) and silicon nitride ($\epsilon \simeq 8/9$).
As expected, as the dielectric constant increases, parasitic capacitance contributions $C_{FR,geost}$ and $C_{add,tot}$ enhance. As a consequence, f_{osc} lowers, with a decrease by 13% moving from lower $\epsilon_{sp}/\epsilon_{spin}$ value to the upper one.

Table 5.1: f_{osc} variation with spacer/inner spacer material.

ϵ_{sp}=ϵ_{spin}	5	6	7	8	9
f_{osc}	23.33 GHz	22.47 GHz	21.67 GHz	20.92 GHz	20.22 GHz

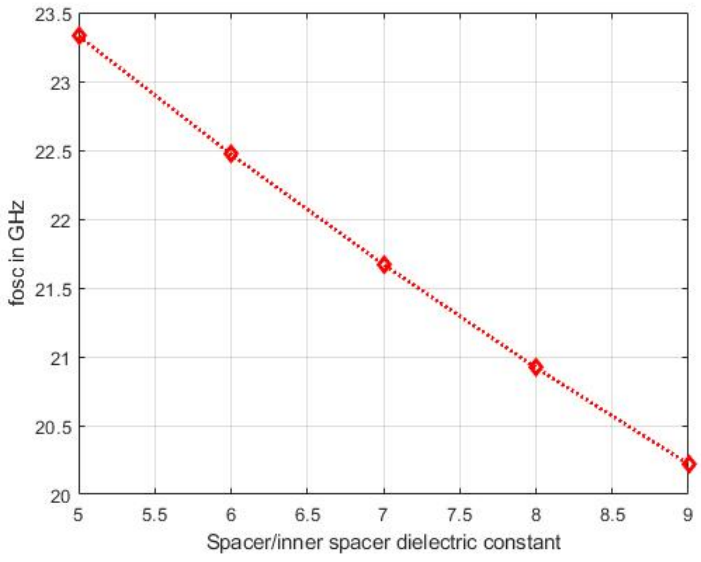


Figure 5.6: f_{osc} as a function of spacer/inner spacer dielectric constant.

Since spacers and inner spacers are realized in two different steps of the technology process (section 1.2), they can be made of different materials.

This second set of simulations assumes that spacers are composed of SiOCN, while the inner spacer material varies between SiOCN and the gate oxide ($\epsilon_{spin} = 25$). Due to the significant rise of the $C_{add,tot}$ contribution, the highest decrease of f_{osc} is reached for ϵ_{insp}= 25. This corresponds to have just the gate oxide separating interchannel metal gates from S/D regions. In particular, in this case, the f_{osc} lowers by 11% comparing to the value obtained considering inner spacers made of SiOCN.

Table 5.2: f_{osc} variation with inner spacer material.

ϵ_{spin}	5	9	13	17	21	25
f_{osc}	23.33 GHz	22.78 GHz	22.25 GHz	21.75 GHz	21.27 GHz	20.81 GHz

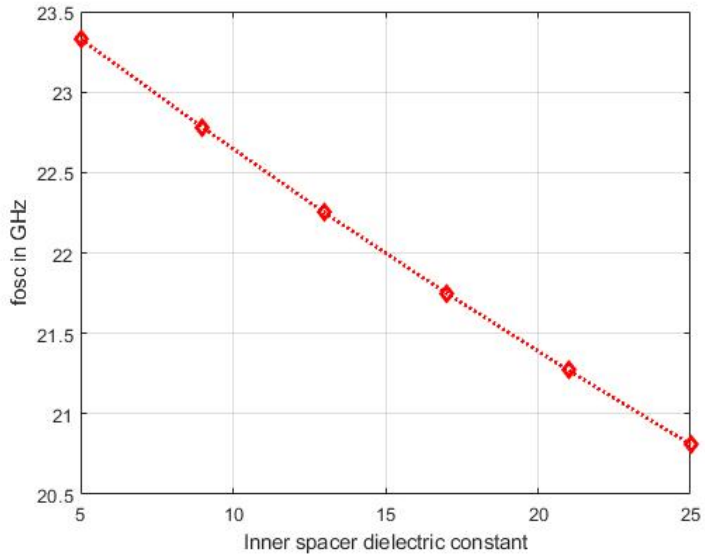


Figure 5.7: f_{osc} as a function of inner spacer dielectric constant.

Another interesting parameter to be investigated is the length of spacers/inner spacers (L_{sp}/L_{spin}).
Supposing that the etching depth of $Si_{0.7}Ge_{0.3}$ sacrificial layers is so accurate that spacers and inner spacers are perfectly aligned, their length have been varied simultaneously.
As the spacers/inner spacers become longer, the extensions length increases and so S/D epitaxies move far away from the gate. This leads to a decrease of parasitic capacitances $C_{FR,geost}$ and $C_{add,tot}$, rising f_{osc}.
An increase by 11 % is registered moving from lower to upper bound.

Table 5.3: f_{osc} variation with spacer/inner spacer length.

L_{sp}=L_{spin}	5 nm	7 nm	9 nm	11 nm	13 nm
f_{osc}	23.33 GHz	24.47 GHZ	25.13 GHz	25.56 GHz	25.85 GHz

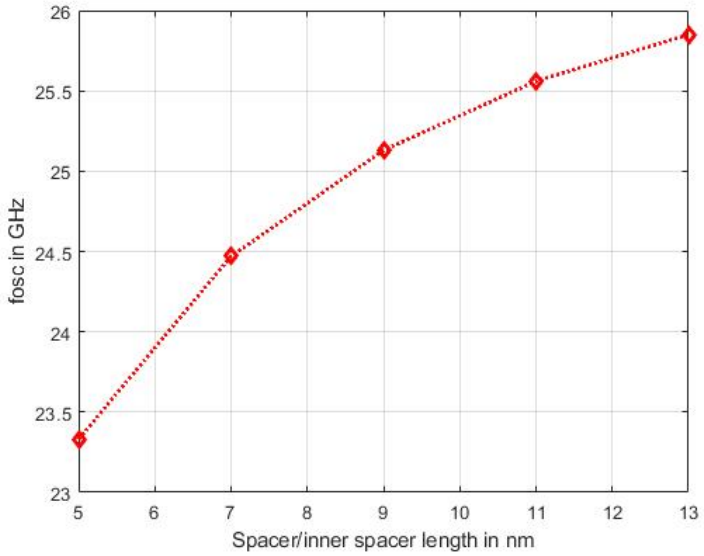


Figure 5.8: f_{osc} as a function of spacer/inner spacer length.

The length of the inner spacers is limited by the precision of the etching depth of $Si_{0.7}Ge_{0.3}$ sacrificial layers. If this process is not so precise, inner spacers can be not aligned with the spacers, resulting longer or shorter.
This set of simulations is performed just varying the inner spacer length (L_{spin}) and maintaing L_{sp} to the standard value of table 4.1.
As expected, thicker inner spacers reduce parasitic coupling between gate and S/D structures. It results in an increase of the oscillation frequency (+4% from the minimum to the maximum L_{spin} value).

Table 5.4: f_{osc} variation with inner spacer length.

L_{spin}	3 nm	5 nm	7 nm	9 nm	11 nm	13 nm
f_{osc}	22.87 GHz	23.33 GHz	23.54 GHz	23.65 GHz	23.73 GHz	23.78 GHz

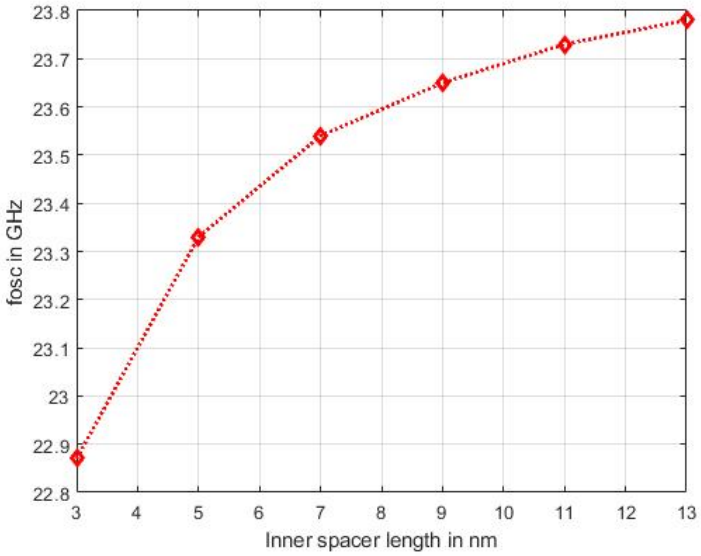


Figure 5.9: f_{osc} as a function of inner spacer length.

5.3 Influence of nanosheets size on f_{osc}

An interesting aspect to be explored is how the nanosheets width influences the oscillation frequency of the five stage ring oscillator made of 3-NSGAAFET.
To this purpose, f_{osc} is evaluated for different values of W_{sh}, maintaining the other parameters of section 4.1 unchanged for both n-type and p-type.
For each n-type width W_{sh}, p-type is sized so that effective resistance of inverters pull-up and pull-down networks are almost equivalent (subsection 5.1.1).

Table 5.5: f_{osc} variation with nanosheets width.

W_{sh}	5 nm	10 nm	15 nm	20 nm	25 nm	30 nm	35 nm
f_{osc}	7.19 GHz	14.05 GHz	19.03 GHz	21.60 GHz	22.86 GHz	23.33 GHz	23.15 GHz

As nanosheets width increases from 5nm to 30nm the oscillation frequency rises. Indeed, even if capacitive contributions enhance for wider channels, the increase in drive current manages to overcome the raise of parasitic capacitances.
This is valid till a maximum value, reached at $W_{sh} = 30nm$; then capacitive components start to dominate and so f_{osc} slightly lowers.

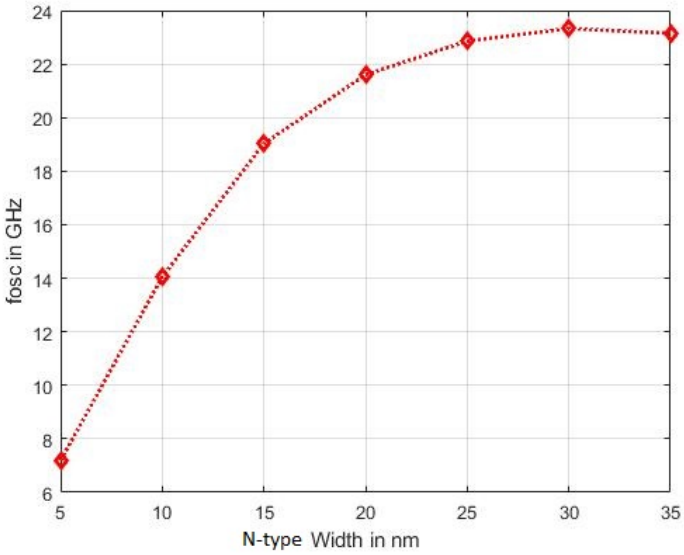


Figure 5.10: Oscillation frequency of the five-stage ring oscillator as a function of W_{sh}.

Chapter 6

Basic cells realization and analysis

The aim of this chapter is to realize a sort of cell library, based on NSGAAFET technology. It should contain some basic building blocks that can be used in more complex digital system design.
To this purpose, an inverter, a two-input NAND and a two-input NOR are implemented and analysed in terms of speed.
From several simulations, an optimum nanosheet width value is found, for each of the cells, representing a good trade-off between high performance and occupation area.

6.1 Settings for delay evaluation

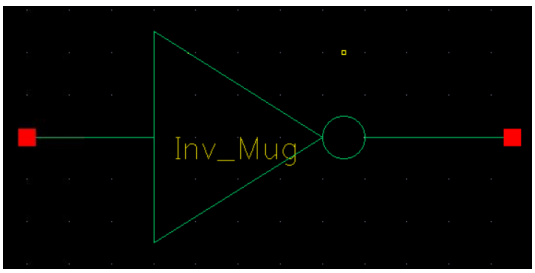
The new modified model for multi-stacked NSGAAFET is now used to implement some simple cells, using which more complex digital circuits can be realized.
These basic gates are:

- Inverter;
- Two-input NAND;
- Two-input NOR.

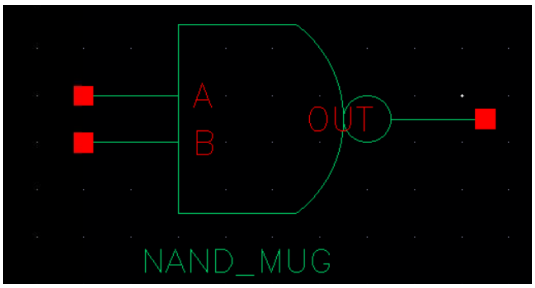
They have been characterized in terms of delay which is the average between the low-high ($t_{pd,LH}$) and high-low ($t_{pd,HL}$) propagation times.

$$\tau = \frac{t_{pd,LH} + t_{pd,HL}}{2} \tag{6.1}$$

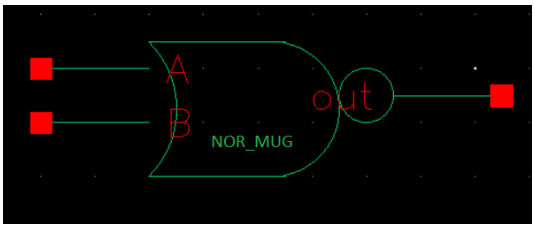
$t_{pd,LH}$ ($t_{pd,HL}$) is defined as the period of time that occurs between a 50% variation of the input from high (low) to low (high) value and a 50% variation of the output from low (high) to high (low) value.



(a) Inverter cell.



(b) Nand cell.



(c) Nor cell.

Figure 6.1: Basic gates based on NSGAAFET technology.

Gates size is varied in order to evaluate how these two delays change with nanosheets width. The other parameters are maintained equal to the ones of table 4.1.
Input signals of frequency 1GHz and a fan-out of four (FO4) minimal inverters have been set. Each load is considered made of NSGAAFETs with just one single nanosheet, having width 10 nm for the n-type and 8 nm for the ptype.

All the transient simulations present in the following sections are performed with Cadence Virtuoso, using Spectre as simulator. Both $t_{pd,LH}$ and $t_{pd,HL}$ are computed using the delay function of Virtuoso Calculator.

6.2 Inverter

The computation of the delay is made for different gate sizes.
Nanosheets width of the two n-type NSGAAFETs is varied between 5 nm to 30 nm by steps of 5 nm. As a consequence, the p-type transistor is sized so that effective resistance of pull-up and pull-down network are almost equivalent.
To satisfy this requirement, the p-type nanosheets width should be β times the n-type one, with $\beta \simeq 0.85$ already defined in section 5.1.1.

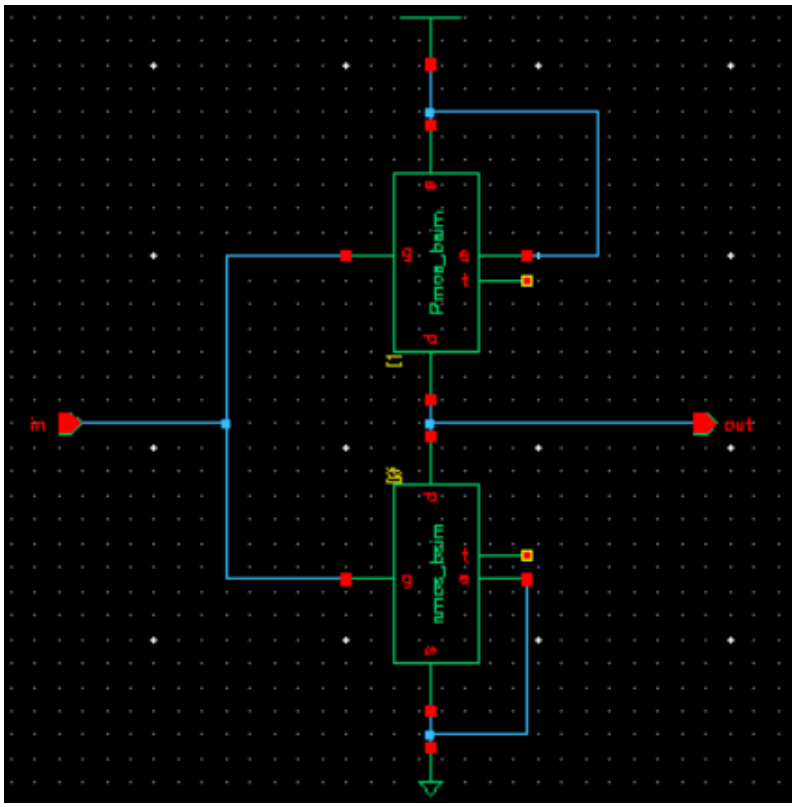


Figure 6.2: Schematic of inverter gate.

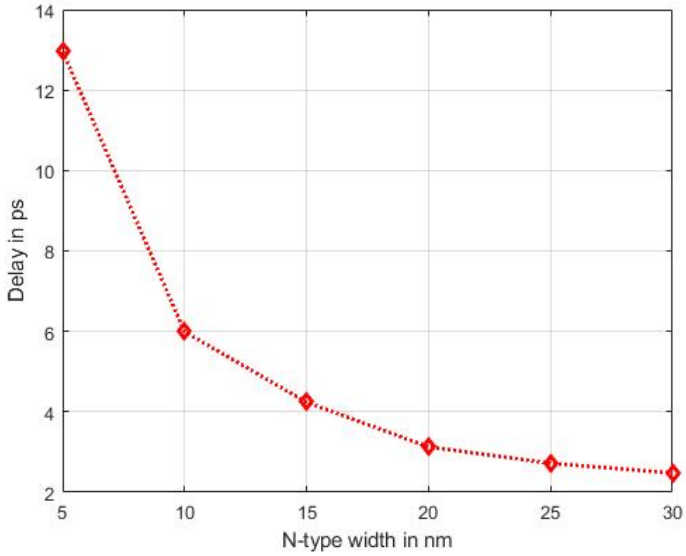


Figure 6.3: Inverter delay as a function of nanosheets width.

Simulation results for the inverter, consisting of 3-NSGAAFETs with parameters of table 4.1, are reported in fig. 6.3.
It shows a decrease of the delay by almost 81 % as the nanosheets width increases from lower to upper bound.
For n-type W_{sh} = 30 nm, the delay reaches a value of τ = 2.48 ps.

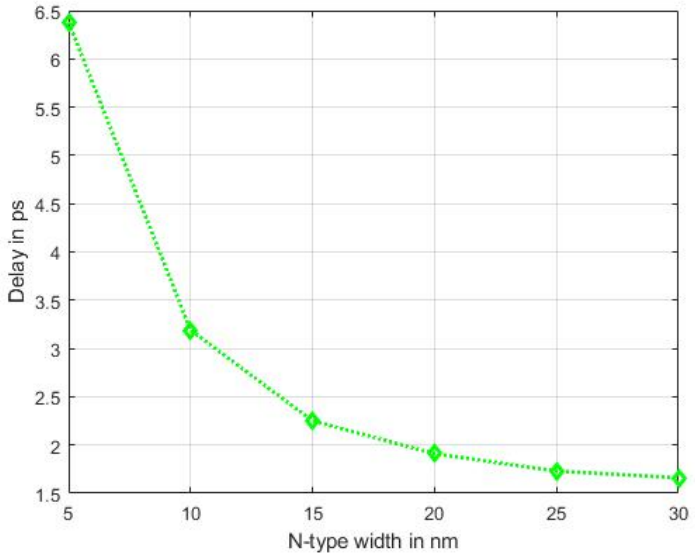


Figure 6.4: Delay vs nanosheets width for an inverter composed of 3-NSGAAFETs, with two stacks of three nanosheets for each one.

Until now, each transistor of the gate has been supposed made of just one stack of three nanosheets.
The addition of another pile of three nanosheets, for each 3-NSGAAFET, has been investigated and the results, in terms of inverter delay, are illustrated in fig. 6.4.
For W_{sh}=30 nm, $\tau \simeq$ 1.66 ps with a reduction by 33% with respect to the single pile case.

6.3 NAND gate

For a NAND gate, the value of each of the two propagation times, involved in the delay computation, depends on the input pattern.
In order to get a worst-case analysis, the two input patterns which bring to the highest $t_{pd,LH}$ and $t_{pd,HL}$ values have been taken into account.

For what concerns $t_{pd,LH}$, the greatest value ($t_{pd,LHmax}$) is obtained when just the p-type guided by input B (see fig. 6.5) moves from OFF to ON state, while the

other p-type remains OFF.
For what concerns $t_{pd,HL}$, the highest value ($t_{pd,HLmax}$) is obtained when both n-type transistors move from OFF to ON state.

The two $t_{pd,LHmax}$ and $t_{pd,HLmax}$ lead to the maximum delay for the NAND gate.

$$\tau_{max} = \frac{t_{pd,LHmax} + t_{pd,HLmax}}{2} \tag{6.2}$$

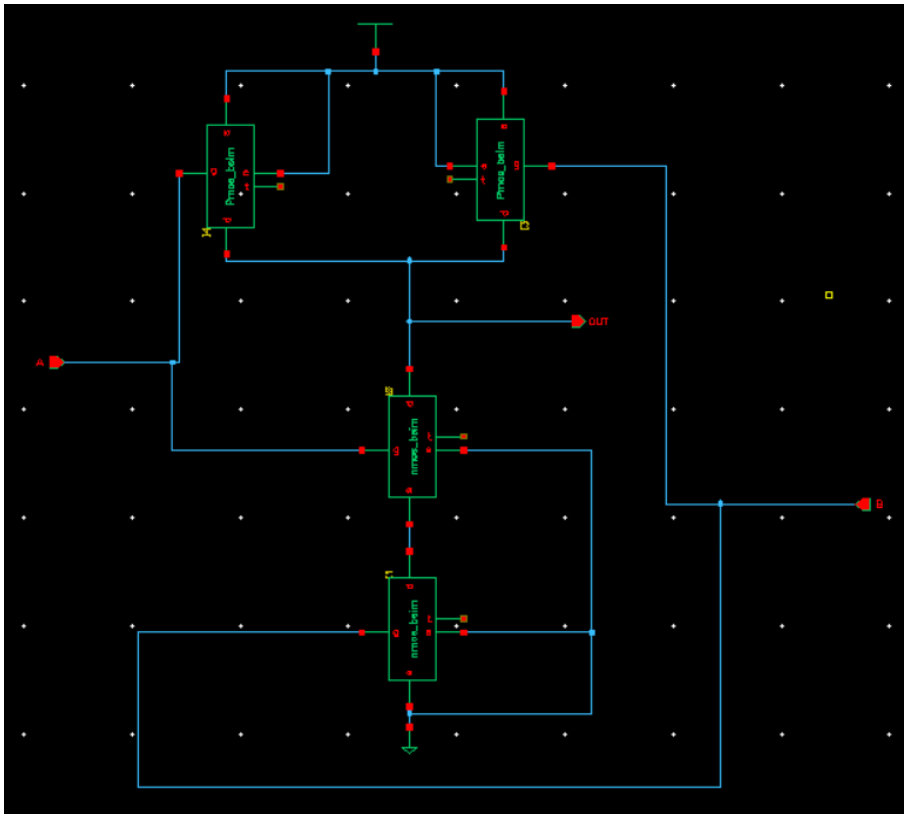


Figure 6.5: Schematic of NAND gate.

The evaluation of the delay is made for different gate sizes.
Nanosheets width of the two n-type NSGAAFETs is varied inside a phisically acceptable range.
As a consequence, the two p-type transistors are sized so that:

- the effective resistance $R_{eq,up}$ of the pull-up network, in the worst case (the one with the highest $R_{eq,up}$, so when just one p-type is ON), is equal to the effective resistance $R_{eq,down}$ of the pull-down network, in the worst case (the one with the highest $R_{eq,down}$, so when both the n-types are ON).

To satisfy this requirement, $W_{sh,p} \simeq \frac{\beta}{2} \cdot W_{sh,n}$, where $W_{sh,p}$ and $W_{sh,n}$ are the nanosheets width of p- and n-type NSGAAFET, respectively.

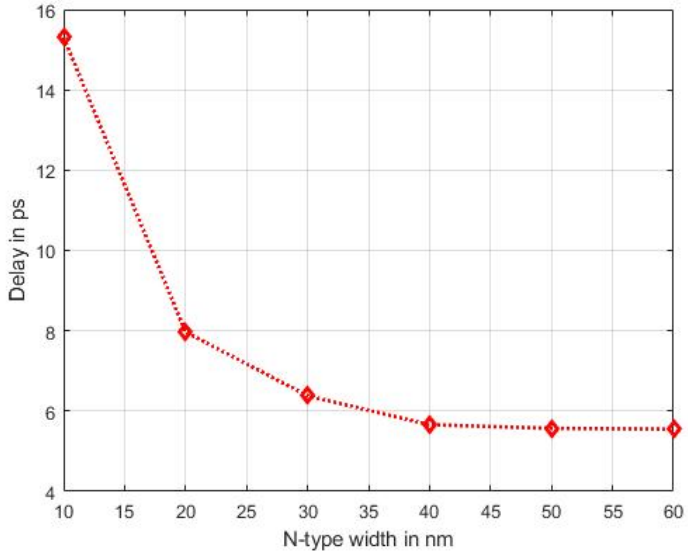


Figure 6.6: NAND gate delay as a function of nanosheets width.

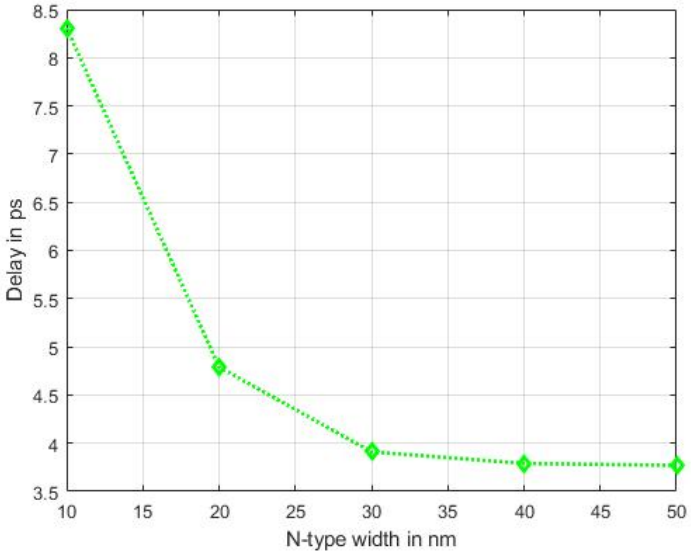


Figure 6.7: Delay vs nanosheets width for a NAND made of 3-NSGAAFETs, with two stacks of three nanosheets for each one.

Simulation results for the NAND gate, composed of 3-NSGAAFETs with parameters of table 4.1, are reported in fig. 6.6.
As expected, the delay decreases with nanosheets width increase, reaching a saturation value of $\tau_{max} \simeq 5.57$ ps at around $W_{sh,n} = 50$ nm.
Then, the addition of another pile of three nanosheets, for each 3-NSGAAFET, has been investigated and the results, in terms of NAND delay, are illustrated in fig. 6.7.
For $W_{sh,n}$=50 nm, $\tau_{max} \simeq 3.77$ ps with a reduction by 32% with respect to the single pile case.

6.4 NOR gate

For the NOR gate, the value of each of the two propagation times, involved in delay computation, depends on the input pattern.
In order to get a worst-case analysis, the two input patterns which bring to the highest $t_{pd,LH}$ and $t_{pd,HL}$ values have been taken into account.

For what concerns $t_{pd,LH}$, the greatest value ($t_{pd,LHmax}$) is reached when both p-type transistors move from OFF to ON state.
For what concerns $t_{pd,HL}$, the highest value ($t_{pd,HLmax}$) is obtained when just the n-type guided by input A moves from OFF to ON state, while the other n-type remains OFF.

The two $t_{pd,LHmax}$ and $t_{pd,HLmax}$ lead to the maximum delay for the NOR gate.

$$\tau_{max} = \frac{t_{pd,LHmax} + t_{pd,HLmax}}{2} \tag{6.3}$$

The evaluation of the delay is made for different gate sizes.
Nanosheets width of the two n-type NSGAAFETs is varied between 10 nm to 30 nm by steps of 5 nm. As a consequence, the two p-type transistors are sized so that:

- the effective resistance $R_{eq,up}$ of the pull-up network, in the worst case (the one with the highest $R_{eq,up}$, so when both the p-types are ON), is equal to the effective resistance $R_{eq,down}$ of the pull-down network, in the worst case (the one with the highest $R_{eq,down}$, so when just one n-type is ON).

To accomplish it, $W_{sh,p} \simeq 2 \cdot \beta \cdot W_{sh,n}$.

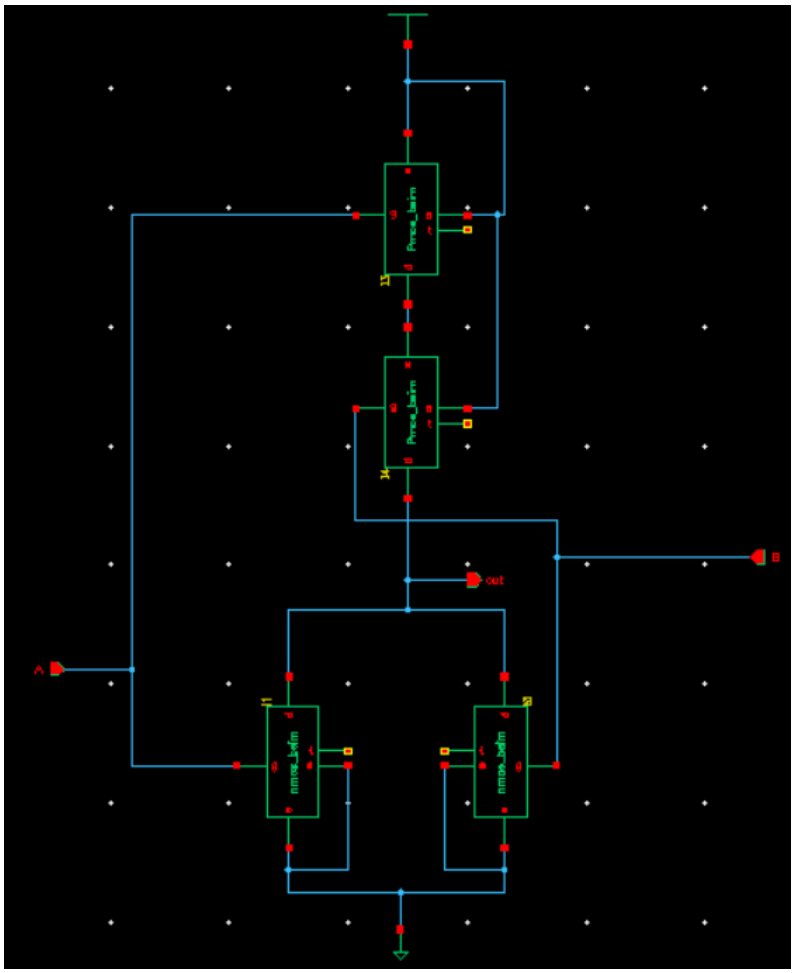


Figure 6.8: Schematic of NOR gate.

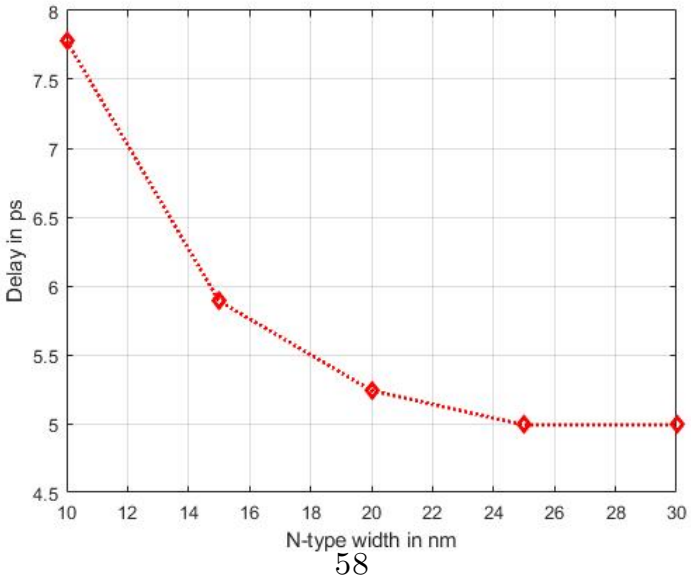


Figure 6.9: NOR gate delay as a function of nanosheets width.

Simulations results for the NOR gate, made of 3-NSGAAFETs with parameters of table 4.1, are reported in fig. 6.9.
As expected, the delay decreases with nanosheet width increase reaching a saturation value of $\tau_{max} \simeq 4.99$ ps at around $W_{sh,n} = 25$ nm.

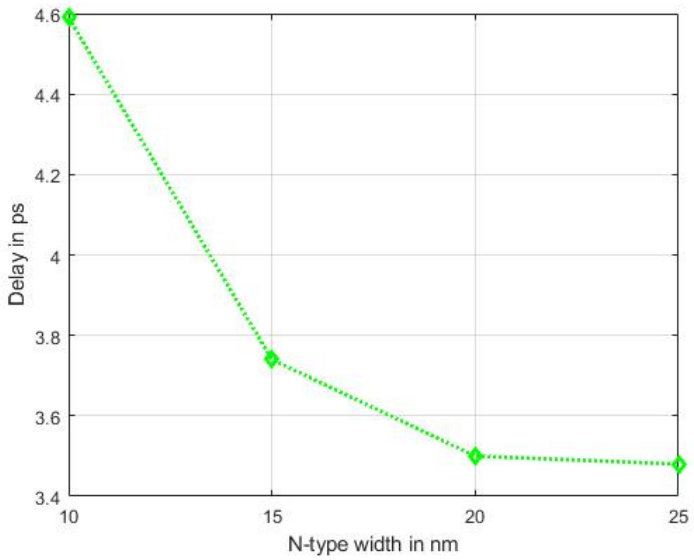


Figure 6.10: Delay vs nanosheets width for a NOR made of 3-NSGAAFETs, with two stacks of three nanosheets for each one.

Then, the addition of another pile of three nanosheets, for each 3-NSGAAFET, has been investigated and the results, in terms of NOR delay, are illustrated in fig. 6.10.
For $W_{sh,n} = 25$ nm, $\tau_{max} \simeq 3.48$ ps with a reduction by 30% with respect to the single pile case.

Chapter 7

Conclusions and future perspectives

This entire work was mainly focused on adapting the already existing multi-gate model BSIM-CMG to the advanced technology of multi-layered NSGAAFET.
The starting point (chapter 1) was to explore the fabrication process, in order to capture the peculiarities of this novel device compared to the FinFET one. Four customized steps have been recognized in the process flow, leading to the final structure explained in section 1.3.
Once the NSGAAFET has been examinated, the attention moved to analyse BSIM-CMG model. It is valid for multi-gate devices of different shapes, including a quadruple-gate, which is a very basic version of NSGAAFET.
In chapter 2, an overview of BSIM-CMG core part was provided. This was done with the aim to inspect its capability to capture the behaviour of different multi-gate structures. Then, a detailed analysis of parasitic source/drain resistances and capacitances modeling was reported.
Chapter 3 was aimed at modifying the original BSIM-CMG model in order to extend it to a multi-stacked nanosheet gate-all-around field effect transistor. To this purpose, new model parameters were introduced and parasitics modeling was carefully improved.
In chapter 4, DC performance evaluation of a HP 3-NSGAAFET was provided. Main device parameters, such as I_{dsat}, I_{sub}, SS, $V_{th,sat}$ and $DIBL$, were extracted using Cadence Virtuoso. The obtained values showed excellent drive current, as expected for HP devices, and optimal channel electrostatic control. This key aspect was investigated for rising nanosheets widths, showing an increase of $DIBL$.
Chapter 5 was mainly focused on AC performance assessment of the aforementioned 3-NSGAAFET. A five-stage ring oscillator was simulated and its oscillation frequency extracted. Then, the impact of the process variations and of the nanosheets size on f_{osc} was investigated. Simulations underlined an increase of f_{osc} with the raise of nanosheets width, till a saturation value that was chosen as the

optimal one.
Finally, chapter 6 was devoted to the realization of some basic blocks, using 3-NSGAAFETs. These cells were analysed in terms of delay, varying nanosheets width. For each gate, an optimum size, representing a good trade-off between high speed and occupation area, was found.
This was a short summary of the work done. This one could be improved by acting on two sides: modeling and applications.
As regards the first field, some fitting parameters of the extended model could be properly adjusted, if a TCAD physical model for a 3-NSGAAFET was available. This would allow the model to have better compatibility with the real device.
From the applications point of view, it could be interesting to apply the developed cells, based on 3-NSGAAFET technology, in more complex digital system design.

Appendix

Appendix A

Modified Verilog-A modules

In this section, just two of the main modules modified for multi-stacked NS-GAAFET are reported. They concern parasitic capacitances and resistances computation.

A.1 S/D Resistance module

```
case(RDSMOD)
1: begin
Rdsi      =  0.0;
Dr        =  1.0;
T2        =  vgs_noswap - vfbsd;
T3        =  sqrt(T2 * T2 + 1.0e-1);
vgs_eff   =  0.5 * (T2 + T3);
T4        =  1.0 + PRWGS_i * vgs_eff;
T1        =  1.0 / T4;
T0        =  0.5 * (T1 + sqrt(T1 * T1 + 0.01));
T5        =  RSW_i * (1.0 + RSDR_a * lexp(0.5 * PRSDR *
lln(V(si,s) * V(si,s) + 1.0E-6)));
i_s = 1;
Rext_source = RSWMIN_i + T5 * T0;
Rext_tot_source = Rext_source;
if (GEOMOD == 2)begin
if (Nsh > 1) begin
while (i_s < Nsh) begin
Rext_tot_source =
(Rext_tot_source * Rext_source)/
(Rext_tot_source + Rext_source);
i_s = i_s + 1;
```

```
end
end
end
Rsource  =  rdstemp *
(RSourceGeo + (Rext_tot_source) * WeffWRFactor);
T2       =  vgd_noswap - vfbsd;
T3       =  sqrt(T2 * T2 + 1.0e-1);
vgd_eff  =  0.5 * (T2 + T3);
T4       =  1.0 + PRWGD_i * vgd_eff;
T1       =  1.0 / T4;
T0       =  0.5 * (T1 + sqrt(T1 * T1 + 0.01));
T5       =  RDW_i * (1.0 + RDDR_a * lexp(0.5 *
PRDDR * lln(V(di,d) * V(di,d) + 1.0E-6)));
i_d = 1;
Rext_drain = RDWMIN_i + T5 * T0;
Rext_tot_drain = Rext_drain;
if (GEOMOD == 2)begin
if (Nsh > 1) begin
while (i_d < Nsh) begin
Rext_tot_drain =
(Rext_tot_drain * Rext_drain)/(Rext_tot_drain + Rext_drain);
i_d = i_d + 1;
end
end
end
Rdrain   =  rdstemp *
(RDrainGeo + (Rext_tot_drain) * WeffWRFactor);
end

0: begin
Rsource  =  RSourceGeo;
Rdrain   =  RDrainGeo;
i=1;
T4       =  1.0 + PRWGS_i * qia;
T1       =  1.0 / T4;
T0       =  0.5 * (T1 + sqrt(T1 * T1 + 0.01));
Rext = RDSWMIN_i + RDSW_i * T0;
Rext_tot = Rext;
if (GEOMOD == 2)begin
if (Nsh>1) begin
while (i< Nsh) begin
Rext_tot = (Rext_tot * Rext)/(Rext_tot + Rext);
```

```
i=i+1;
end
end
end
Rdsi      =  rdstemp * (Rext_tot) * WeffWRFactor;
Dr        =  1.0 +
(NFINtotal) * beta * ids0_ov_dqi / (Dmob * Dvsat) * Rdsi;
end

2: begin
i=1;
T4        =  1.0 + PRWGS_i * qia;
T1        =  1.0 / T4;
T0        =  0.5 * (T1 + sqrt(T1 * T1 + 0.01));
Rext = RDSWMIN_i + RDSW_i * T0;
Rext_tot = Rext;
if (GEOMOD == 2) begin
if (Nsh>1) begin
while (i< Nsh) begin
Rext_tot = (Rext_tot * Rext)/(Rext_tot + Rext);
i=i+1;
end
end
end
Rdsi      =  rdstemp * (RSourceGeo + RDrainGeo + Rext_tot) *
WeffWRFactor;
Dr        =  1.0 +
(NFINtotal) * beta * ids0_ov_dqi / (Dmob * Dvsat) * Rdsi;
Rsource   =  0.0;
Rdrain    =  0.0;
end
endcase
```

A.2 Fringe capacitance module

```
`define Cfringe_2d(block_name, Hg, Hc, Lext, Wfin,
 Lc, Lg, Tox, Cf1, Cgg, Nsh, Hstack, flag)
begin : block_name
real Hr, Lr, Hgdelta, Lmax, y, x;
real CcgSat, Cnon, TT1, Ccg1, r1cf, Rcf, Ccg2;
real Ccg, C1, C2, C3, Cfglog, dcf, TT0;
```

```
real TT2, Cfgsat, delta, Cfg;
Hr = 2.3 + 0.2 * ((Hg) + (Tox)) / (Hc);
Lr = 1.05;
Hgdelta = abs((Hg) + (Tox) - (Hc));
Lmax = (Lext) * Lr;
y = min((Hc), (Hg) + (Tox));
x = (Lext) / (Hr + 1.0);
Cnon = 1.7e12;
CcgSat = epssp * (y - x) / (Lext);
TT1 = Cnon * CcgSat;
if(TT1 > `EXPL_THRESHOLD)
Ccg1 = CcgSat;
else
Ccg1 = 1.0 / Cnon * ln(1.0 + lexp(TT1));
r1cf = 0.5 *
min((Hc) / ((Hg) + (Tox)), ((Hg) + (Tox)) / (Hc));
Rcf = Hgdelta * r1cf;
Ccg2 = epssp * 2 / `M_PI *
ln(((Lext) + 0.5 * `M_PI * Rcf) / (Lext));
if( GEOMOD==2 && flag==1 ) begin
Ccg = (Hstack) * (Ccg1 + Ccg2);
end else begin
Ccg = (Wfin) * (Ccg1 + Ccg2) ;
end
x = Lmax / (Hg);
C1 = 4.0 / (sqrt(2.0 * (x + 1)) * `M_PI);
C2 = sqrt((Tox) * (Tox) + 2.0 * (Hg) * (Tox) +
(Hg) * (Hg) * (x + 1)) * sqrt(x + 1) + (Tox) +
(Hg) * x + (Hg);
C3 = (Tox) * sqrt((x + 1) * (x + 4)) + Tox * (x + 2);
Cfglog = epssp * (C1 * ln(C2 / C3) + 12.27);
dcf = Hr * Lr;
TT0 = sqrt(dcf * dcf + 1.0);
TT1 = sqrt((dcf * dcf + 1) * ((dcf * (Tox)) * (dcf * (Tox)) +
2 * dcf * Lmax * (Tox) + (dcf * dcf + 1) * Lmax * Lmax))
+ dcf * (Tox) + dcf * dcf * Lmax + Lmax;
TT2 = (TT0 + 1.0) * (dcf * (Tox));
Cfgsat = 2.0 * epssp * sqrt(2) / `M_PI * (Cf1) * dcf
/ TT0 * ln(TT1 / TT2);
delta = 1.2e-12;
TT1 = Cfgsat - Cfglog - delta;
if(GEOMOD==2 && flag==1) begin
```

```
Cfg     =  (Nsh * Wfin) * (Cfgsat - 0.5 * (TT1 +
sqrt(TT1 * TT1 + 4 * delta * Cfgsat)));
end else begin
Cfg     =  (Wfin) * (Cfgsat - 0.5 * (TT1 +
sqrt(TT1 * TT1 + 4 * delta * Cfgsat)));
end
Cgg     =  Ccg + Cfg;
end
```

Bibliography

[1] ITRS experts, "International Technology Roadmap for Semiconductors", 2020 report, 2020.

[2] J. Ch. Barbe et al.,"Stacked nanowires/nanosheets GAA MOSFET from technology to design enablement", CEA, LETI, Grenoble, 2017.

[3] J.P. Duarte et al., "A Universal Core Model for multiple-Gate Field-Effect Transistors", IEEE Transactions on Electron Devices, Feb 2013.

[4] J.P. Duarte, "Mathematical Compact Models of Advanced Transistors for Numerical Simulation and Hardware Design", Technical report, May 2018.

[5] Samar K. Saha, "Finfet Devices for VLSI Circuits and Systems", 2020.

[6] J.P. Duarte et al., "Multi-Gate MOSFET Compact model", Technical Manual, 2015.

[7] J.P. Duarte, "Unified Compact Model for Gate All Around FETs-Nanosheets, Nanowires, Multi Bridge Channel MOSFETs".

[8] Y. Choi et al., "FinFET Process Refinements for Improved Mobility and Gate Work Function Engineering", Berkeley ,2002.

[9] Takagi et al., "On the Universality of Inversion Layer Mobility in Si MOSFET's", IEEE Transactions on Electron Devices, December 1994.

[10] Chenming-Hu, "Motion and Recombination of Electrons and Holes", 2009.

[11] N. Loubet et al., “Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET”, in Proc. Symp. VLSI Technol., Jun. 2017.

[12] C. W. Yeung et al., “Channel geometry impact and narrow sheet effect of stacked nanosheet”, IEDM Tech. Dig., Dec. 2018.

[13] A. Razavieh, P. Zeitzoff, and E. J. Nowak, “Challenges and limitations of CMOS scaling for FinFET and beyond architectures”, IEEE Trans. Nanotechnol., Sep. 2019.

[14] D. Jang et al., “Device exploration of nanosheet transistors for sub-7-nm technology node”, IEEE Trans. Electron Devices, Jun. 2017.

[15] S. Barraud et al., “Performance and design considerations for gate-allaround stacked-nanowires FETs”, in IEDM Tech. Dig., San Francisco, CA, USA, Dec. 2017.

[16] E. Y. Jeong et al., “Investigation of RC parasitics considering middleof-the-line in Si-bulk FinFETs for sub-14-nm node logic applications”, IEEE Trans. Electron Devices, Oct. 2015.

[17] Gaurav Saini, Ashwani K Rana , “Physical Scaling Limits of FinFET Structure: A Simulation Study” ,International Journal of VLSI design and Communication Systems (VLSICS), Mar. 2011.

[18] Peide Ye, Thomas Ernst, Mukesh V. Khare, "The last silicon transistor", August 2019.

[19] D. Yakimets et al., “Vertical GAAFETs for the Ultimate CMOS Scaling”, IEEE Trans. Electron Devices, May 2015.

[20] A. Fallis, "Nanoscale Transistors", Kluwer Academic Publishers, 2006.

[21] J.-P. Colinge, "FinFETs and Other Multi-Gate Transistors", Boston, MA: Springer US, 2008.

[22] A. Asenov et al., “Nanowire transistor solutions for 5nm and beyond,” in 2016 17th International Symposium on Quality Electronic Design (ISQED), 2016.

[23] F. Adamu-Lema et al., “Modelling and Simulation of Advanced Semiconductor Devices,” ECS Trans., Aug. 2017.

[24] H. R. Huff and D. C. Gilmer, "High Dielectric Constant Materials", Berlin/Heidelberg: Springer-Verlag, 2005.

[25] C. H. Lin et al., “Channel doping impact on FinFETs for 22nm and beyond,” in 2012 Symposium on VLSI Technology (VLSIT).

[26] T. Al-Ameri et al., “Performance of vertically stacked horizontal Si nanowires transistors: A 3D Monte Carlo/2D Poisson Schrodinger simulation study,” in 2016 IEEE Nanotechnology Materials and Devices Conference (NMDC), 2016.

[27] A. Ghetti and D. Rideau, “3D Monte Carlo Device Simulation of NanoWire MOSFETs including Quantum Mechanical and Strain Effects,” in 2006 International Conference on Simulation of Semiconductor Processes and Devices, 2006.

[28] U. K. Das et al., “Limitations on Lateral Nanowire Scaling Beyond 7-nm Node,” IEEE Electron Device Lett., vol. 38, 2017.

[29] B. Hoeneisen and C. A. Mead, “Fundamental limitations in microelectronics-MOS technology,” Solid State Electron., vol. 15, 1972.

[30] Gaben, L., S. Barraud, M. Jaud, S. Martinie, O. Rozeau, J. Lacord, G. Hiblot et al. , "Stacked-Nanowire and FinFET Transistors: Guidelines for the 7nm node ", 2015.

[31] Mark Lundstrom and Jing Guo, "Nanoscale Transistors-Devices Physics, Modeling and simulation", New York:Sringer, 2006.

[32] Triyoso, D.H., Mulfinger, G.R., Hempel, K., et al.: "Characterization of atomic layer deposited low-k spacer for FDSOI high-k metal gate transistor", IEEE Int. Conf. on IC Design and Technology, Austin, TX, USA, May 2017.

[33] S.-D. Kim, M. Guillorn, I. Lauer, P. Oldiges, T. Hook, and M.-H. Na, “Performance trade-offs in FinFET and gate-all-around device architectures for 7 nm-node and beyond”, Oct. 2015.