LUND UNIVERSITY

Master's Thesis

Automated modelling and optimization of a ratioed logic inverter utilizing nanowire-based transistors

By

Martin Berg and Kristofer Jansson

(2009-12-07)

Department of Electrical and Information Technology
Faculty of Engineering, LTH, Lund University
SE-221 00 Lund, Sweden

“Only one who devotes himself to a cause with his whole strength and soul can be a true master. For this reason mastery demands all of a person.”

- Albert Einstein

PREFACE

Before you venture further into the realm of equations and circuit diagrams, we would like to take the opportunity to say a few words about this report. This master's thesis is the ultimate culmination of 4.5 years studies in Engineering Nanoscience at Lund University. The background of the thesis is a project work in the nanoelectronics course at LTH, of which the main findings of were used as a starting point.

One of the main purposes of the thesis is to build a theoretical foundation for nanowire device characterization, in the spirit of the work on nanowires, performed at the Nanometer Structure Consortium in Lund. Measurements on some test devices were expected, but this did not come into fruition due to immature processing technology.

At the start of the work, the idea was to design a range of different test circuits based on nanowire transistors. However, the project evolved into a more focused enterprise, encompassing modelling, simulation and optimization of a ratioed logic inverter.

During the course of the project, knowledge from multiple areas was applied in order to solve the multitude of problems encountered. In addition, different bits of code developed during the course of the project, converged into a large software suite, automating large parts of the project. Although many long nights were devoted to, the work was deemed rewarding and interesting.

The authors would like to acknowledge the support of Professor Lars-Erik Wernersson and Professor Peter Nilsson, our supervisors for this project. Thanks are also extended to the participants in the nanowire group meetings and Wireless With Wires (WWW) meetings for their input and support.

Finally, the authors would like to take the opportunity to thank their respective families and friends for their undying support during the ordeals, encountered during the writing of this thesis.

Live long and prosper,

Martin Berg & Kristofer Jansson
Lund, december 2009

ABSTRACT

Automated modelling and optimization of a ratioed logic inverter utilizing nanowire-based transistors

The continuing trend of electronic miniaturization is approaching a critical limit for conventional silicon based technology. Transistors based on nanowires are an interesting concept that might extend this trend deeper into the nano-scale domain. In this work, the usage of nanowire-based transistors is studied in the context of digital circuit design.

The authors propose a complete workflow from measurement, through transistor modelling and circuit simulation to an optimized inverter design. This workflow is realized in the form a software suite, developed for this purpose. By implementing intelligent algorithms and extraction methods, this allows the workflow to be highly automated, reducing the time needed for a complete analysis from weeks to a matter of minutes.

A modified MOSFET model is introduced based on the proper deep-submicron MOSFET model. By adding a series resistance and introducing a dependence on the nanowire quantity in addition to neglecting the channel length modulation, the Jansson-Berg model is defined. The model is implemented into the software and an acceptable correspondence to measurement data is achieved. To simulate the parasitic capacitances in the transistor design, a model consisting of 21 separate capacitances is derived.

The Jansson-Berg model is analysed in the application of simulating an inverter. Due to the unavailability of p-type transistors, a ratioed logic design is proposed, utilizing only n-type transistors. As an exact analytical solution of the circuit performance is nearly impossible, it is concluded that series solutions are necessary. Although simplifications are made, the analytical solutions achieve a good correspondence to a purely numerical analysis.

Finally, the transistor design is optimized based on the energy-delay product. This is done intelligently, by testing about 300 million transistor designs, bound by specified design rules. Based on this large data set, it is concluded that further downscaling is advisable.

Automatiserad modellering och optimering av en inverterare med nanotrådstransistorer

Trenden för miniatyrisering av mikroelektronik närmar sig gränsen för vad som är möjligt att uppnå med konventionell kiselteknologi. En teknik, med potential att pressa utveckling mot allt mindre komponenter, är nanotrådstransistorer. I detta examensarbete studeras implementering av dessa transistorer i digitala kretsar.

En fullständig analys av en krets innebär flera påföljande steg, innefattande mätning, transistormodellering, kretssimulering och slutligen optimering. Författarna föreslår ett arbetsflöde, där de olika delarna sammanställs i form en egenutvecklad programvara. Genom implementering av intelligenta algoritmer och extraktionsmetoder är det möjligt att automatisera dessa processer och därigenom minska arbetsbördan för en komplett analys, från veckor till en fråga om minuter.

En modell för karakteriseringen av nanotrådstransistorerna implementeras genom modifiering av en enkel MOSFET-modell. Modellen modifieras för att inkorporera serieresistans och försumma kanallängdsmodulation samt att ett beroende på antalet nanotrådar införs. Genom implementering i mjukvaran av en automatiserad modellering, kan en god överensstämmelse erhållas mellan mätdata och modell. För att få en simulera de parasitiska kapacitanserna i transistordesignen, tas en modell fram som består av 21 st. separata parasitiska element.

Transistormodellen används för att numeriskt och analytiskt studera en prestanda för en digital inverterare. Eftersom inga transistorer av p-typ finns att tillgå, används en kvotbaserad logik, bestående av enbart transistorer av n-typ. En exakt tidsanalytisk lösning av en sådan inverterare är i princip omöjlig och därför är serieutvecklingar nödvändiga. Även om detta är en grov förenkling, föreligger en god överensstämmelse mellan den analytiska och numeriska analysen.

Slutligen optimeras transistordesignen utifrån produkten mellan energiförbrukning och kvadrerad stegfördröjning. Detta utförs med hjälp av en intelligent algoritm som under projektets gång testat omkring trehundramiljoner olika transistordesigner. Från denna stora datamängd dras slutsatsen att ytterligare nedskalning av transistordesignen är fördelaktig.

Automatisierte Modellierung und Optimierung eines Nicht-Gatter basiert auf Nanodraht-Transistoren

In dieser Arbeit wurde die Verwendung von Nanodraht-basierten Transistoren, im Zusammenhang digitalen Schaltung, untersucht. Die Autoren vorschlagen einen kompletten Arbeitsablauf von der Messung, durch die Transistor-Modellierung und Schaltungssimulation, zu einem optimierten Wechselrichterdesign. Dieser Arbeitsablauf ist in der Form eines Computerprogramms realisiert, für diesen Zweck entwickelt. Ein einfaches MOSFET-Modell wurde verwendet, um eine analytische Lösung der Schaltung zu erwerben. Obwohl die Lösung vereinfacht ist, eine gute Entsprechung zu einer numerischen Analyse erreicht ist. Eine Optimierung erfolgt, basiert vom Energie-Laufzeit Produkt, und der Rückschluss wird gezogen, dass weitere Verkleinerung ratsam ist.

La Modélisation et l'Optimisation Automatique d'un Fonction-NON Utilisation de Nanofils à base de Transistors

Dans ce travail, l'utilisation de nanofils à base de transistors est étudiée dans le cadre de la conception de circuits numériques. Les auteurs proposent un flux de travail complet de la mesure, par la modélisation et la simulation de circuits à transistors à une conception optimisée de l'onduleur. Ce flux de travail est réalisé sous la forme d'une suite logicielle, développé à cet effet. Un modèle MOSFET simple a été utilisée en vue d'acquérir une solution analytique de la performance du circuit. Même si la solution n'est pas exacte, une bonne correspondance à une analyse chiffrée est atteinte. Le processus d'optimisation se fait sur la base de l'énergie-produit de retard et il est conclu que davantage de réduction d'échelle est souhaitable.

ABBREVIATIONS AND SYMBOLS

Abbreviations

DSMM	–	Deep SubMicron Model
EDP	–	Energy-Delay product
FET	–	Field-Effect Transistor
HC	–	Half-circle
MOSFET	–	Metal-Oxide-Semiconductor Field-Effect-Transistor
NWFET	–	NanoWire Field Effect Transistor
PDN	–	Pull-Down Network
PDP	–	Power-Delay product
PDSMM	–	Proper Deep SubMicron Model
PP	–	Parallel-plate
SW	–	Side-wall
VTC	–	Voltage Transfer Characterization
WIGFET	–	Wrapped-Insulator-Gate Field-Effect

Latin symbols

A	–	Plate capacitor area	$C_{DS,pp}$	–	Drain-Source overlap capacitance
A_n	–	Optimization parameter	C_{GD}	–	Gate-Drain capacitance
β	–	Noise factor	$C_{GD,edge,x}$	–	Gate-Drain edge capacitance (x)
C_{DS}	–	Drain-Source capacitance	$C_{GD,edge,y}$	–	Gate-Drain edge capacitance (y)
$C_{DS,edge,x}$	–	Drain-Source edge capacitance (x)	$C_{GD,fri}$	–	Gate-Drain fringe capacitance
$C_{DS,edge,y}$	–	Drain-Source edge capacitance (y)	$C_{GD,fri,x}$	–	Gate-Drain fringe capacitance (x)

$C_{GD,fri,y}$	–	Gate-Drain fringe capacitance (y)	$C_{GS,fri,x}$	–	Gate-Source fringe capacitance (x)
$C_{GD,pp}$	–	Gate-Drain overlap capacitance	$C_{GS,fri,y}$	–	Gate-Source fringe capacitance (y)
$C_{GD,wire}$	–	Gate-Drain wire capacitance	$C_{GS,pp}$	–	Gate-Source overlap capacitance
$C_{GD,wire,x}$	–	Gate-Drain wire capacitance (x)	$C_{GS,wire}$	–	Gate-Source wire capacitance
$C_{GD,wire,y}$	–	Gate-Drain wire capacitance (y)	$C_{GS,wire,x}$	–	Gate-Source wire capacitance (x)
C_{GE}	–	Gate-Electrode capacitance	$C_{GS,wire,y}$	–	Gate-Source wire capacitance (y)
$C_{GE,edge,r}$	–	Gate-Electrode edge capacitance	C_{input}	–	Input capacitance
$C_{GE,fri}$	–	Gate-Electrode fringe capacitance	C_{load}	–	Load capacitance
$C_{GE,fri,r}$	–	Gate-Electrode fringe capacitance (r)	C_N	–	Effective noise capacitance
$C_{GE,pp}$	–	Gate-Electrode overlap capacitance	C_{ox}	–	Gate-oxide capacitance
$C_{GE,wire}$	–	Gate-Electrode wire capacitance	$C_{parasitic}$	–	Parasitic capacitance
C_{GS}	–	Gate-Source capacitance	C_{pp}	–	Overlap capacitance
$C_{GS,edge,x}$	–	Gate-Source edge capacitance (x)	C_{π}	–	Half-circle fringe capacitance
$C_{GS,edge,y}$	–	Gate-Source edge capacitance (y)	C_{sw}	–	Sidewall capacitance
$C_{GS,fri}$	–	Gate-Source fringe capacitance	$C_{top,top}$	–	Top-to-top capacitance
			d	–	Plate capacitor distance

d_{DS}	–	Drain to source distance	L	–	Gate height
d_{GD}	–	Gate to drain distance	L_{eff}	–	Effective gate height
d_{GE}	–	Gate to electrode distance	l	–	Unit cell width
d_{GS}	–	Gate to source distance	l_x	–	Unit cell width (x)
E	–	Electron energy	l_y	–	Unit cell width (y)
E_C	–	Conduction band energy	M_{load}	–	Load transistor symbol
E_F	–	Fermi energy	M_{PDN}	–	PDN transistor symbol
$E_{v,n}$	–	Energy level in valley	m	–	Miller effect index
f_S	–	Simulation data	$m_{d,v}$	–	Effective mass
f_D	–	Measurement data	N_0	–	1-D density of states substitution
g_m	–	Process transconductance	N_C	–	Effective density of states
g_v	–	Valley degeneracy	N_D	–	Donor concentration
H	–	Electrode height difference	N_{load}	–	Number of nanowires in load transistor
H_E	–	Electrode height	N_{PDN}	–	Number of nanowires in pull-down transistor
h	–	Planck's constant	n	–	Electron carrier concentration
I_{load}	–	Load transistor current	O	–	Higher order terms
I_{PDN}	–	PDN transistor current	P_A	–	Average power dissipation
k	–	Gain factor	Q	–	Simulation quality
k_B	–	Boltzmann's constant	q	–	Elementary charge
k'	–	Process transconductance parameter			

R	–	Nanowire radius	V_{DD}	–	Supply voltage
R_s	–	Series resistance	V_{DS}	–	Gate-Source voltage
S	–	Space between electrodes	V_{DSAT}	–	Saturation voltage
T	–	Temperature; Thickness	V_{GS}	–	Gate-Source voltage
T'	–	Conformal mapping thickness	V_{IH}	–	Input-High voltage
$T_{GE,fri}$	–	Gate-Electrode fringe thickness	V_{in}	–	Input voltage
$T_{GE,fri,r}$	–	Gate-Electrode fringe thickness (r)	V_{IL}	–	Input-Low voltage
$T_{GE,wire}$	–	Gate-Electrode wire thickness	V_N	–	Thermal noise voltage
t	–	Time	V_{OH}	–	Output-High voltage
t_{ox}	–	Gate oxide thickness	V_{out}	–	Output voltage
v_{sat}	–	Saturation velocity	V_{OL}	–	Output-Low voltage
V_C	–	Critical voltage	V_T	–	Threshold Voltage
			W	–	Nanowire diameter
			W_E	–	Electrode width
			W_O	–	Electrode overlap

Greek symbols

α	–	WIGFET structural parameter	ε_{eff}	–	Effective permittivity
α^*	–	Voltage independent α-parameter	ε_{GE}	–	Gate-Electrode spacer relative permittivity
β	–	Noise factor	ε_{GS}	–	Gate-Source spacer relative permittivity
γ	–	WIGFET empirical parameter	ε_r	–	Insulator relative permittivity
ε_0	–	Permittivity of free space			

η	–	Empirical voltage dependent parameter	σ_S	–	Jansson-Berg saturation region substitution
λ	–	Channel length modulation	τ	–	Delay time coefficient
λ_j	–	WIGFET structural parameter	$\tau_{0\to1}$	–	Low-to-high delay
			$\tau_{1\to0}$	–	High-to-low delay
λ_j^*	–	Voltage independent λ_j -parameter	τ_p	–	Propagation delay
			φ_m	–	Metal work function
μ	–	Field-dependent mobility	φ_{ms}	–	Work function difference
μ_0	–	Mobility	φ_s	–	Semiconductor work function
ξ	–	Elliptical mapping scaling parameter	φ_{th}	–	Thermal voltage
ρ_j	–	WIGFET empirical parameter	χ	–	WIGFET empirical parameter
σ	–	Jansson-Berg region-independent substitution	χ_m	–	Metal electron affinity
			χ_s	–	Semiconductor electron affinity
σ_L	–	Jansson-Berg linear region substitution	ψ_T	–	Threshold point surface potential

CONTENTS

Part I INTRODUCTION

1.1 Background

In the last decades, an ever-increasing transistor integration density has been observed. This development has led to great achievements in electrical engineering, such as the personal computer and the cell phone. There are however clear indications that the ongoing scaling of silicon-based electronics may come to a halt.

To overcome the problems that are facing the electronic industry, new and innovative solutions are needed. Some proposals are based on changes of transistor materials, e.g. going from silicon to high mobility III/V-materials, such as indium arsenide. One other major revolution is approaching sizes where quantum mechanical effects are prominent. Nanowires are an interesting concept since it incorporates both of these two concepts and much advancement have already been done in this field.

It is important to have models of high quality, in order to characterize nanowire devices, since it provides an insight into what benefits can be expected in the future. Using conventional figures of merits, derived from the models, also allows benchmarking against competing technologies.

1.2 Workflow

During the course of the project, a highly automated workflow is devised. The workflow encompasses all steps from device manufacture, via device simulation and optimization back to the manufacturing of a new device. This workflow is illustrated in Figure 1.

Device manufacture and measurements are out of the scope of this thesis, but full automation of steps 3-6, by the development of specialized software is one of the main goals. This means that the only bottleneck for completing a full cycle is in the device manufacture and measurements. The result is that valuable person-hours of work can be allocated to more relevant tasks than routine and tedious modelling and optimization. The final step of the software suite is the output of a full modelling report and a proposal of a new optimized design, possibly to be put to the test at the next device manufacture.

Figure 1 – Proposed workflow, from device manufacture to report generation. All steps excluding 1 and 2 are implemented in the developed software and are highly automated.

1.3 Structure

The report is divided into three main parts. The first section consists of transistor characterization, in which the transistor layout is described as well as the theory behind the modelling. This is followed by methodology and results of the modelling section. In the second main part, the transistors are implemented in an inverter utilizing ratioed logic. Timing analysis is performed and the problems with parasitic capacitances and resistances are addressed. Finally, optimization of the device is done and an analysis of what can be expected from a future downscaling of the devices will be presented.

In order to provide the necessary tools for transistor modelling and circuit optimization, a computer application is developed. It features the ability to load transistor measurement data as provided from the measurement setups. By using a number of simple controls, it is possible to obtain full transistor characterization based on a range of different MOSFET models. Additionally, the software features the ability to simulate and optimize a ratioed logic inverter utilizing the previously characterized transistors.

Part II THE NWFET

2.1 Transistor structure

The transistor that is being studied in this thesis is based on the Wrapped Insulated Gate Field Effect Transistor (WIGFET). Unlike the conventional planar MOSFET, this transistor has a vertical structure utilizing a nanowire as channel.

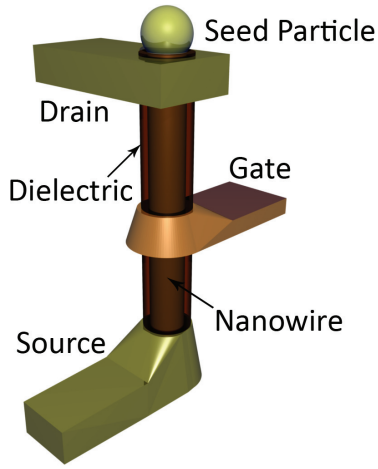


Figure 2 – A WIGFET consisting of a single nanowire.

2.2 Nanowire

The nanowire consists of a semiconductor material, often utilizing high-mobility III/V-materials such as indium arsenide (InAs). The diameter of the nanowire is in the range of tens of nanometres and thus well in the realm of nanotechnology. In reality, the nanowire has a hexagonal cross-section area, but it is often approximated as being of a cylindrical shape.

2.2.1 Growth

Nanowires are often grown using standard epitaxial growth techniques, with the help of a catalytic seed particle. The seed particle is positioned with the aid of electron beam lithography, which allows for placement with a high resolution. The size of the seed particle determines the diameter of the nanowire, while the length is controlled by the growth time.

2.2.2 Field-dependent mobility

Since the nanowire acts as a MOSFET channel, is preferred to have a high mobility. In short channel devices, the mobility is highly dependent on the electrical field along the wire. In addition, it directly influences the current levels and the onset of saturation. It should be noted that this saturation is strictly due to velocity saturation, which is described in section 2.2.3.

The field-dependent mobility is often used in physical models describing the DC-characteristics of transistors. In more simple models however, it is sufficient to use the mobility at zero bias, μ_0.

The field-dependent mobility, μ, is given by the equation below:

$$\mu = \mu_0 \left(1 + \frac{\mu_0 V_{DS}}{v_{sat}\, L}\right)^{-1} \qquad (1)$$

In the above expression, v_{sat} is the saturation velocity, L the effective channel length and V_{DS} the drain to source voltage (the end terminals of the nanowire). From now on, the effective channel length is assumed equal to the height of the gate electrode. In Figure 3, the expression above characterizing the field-dependent mobility is plotted as a function of the applied voltage along the channel.

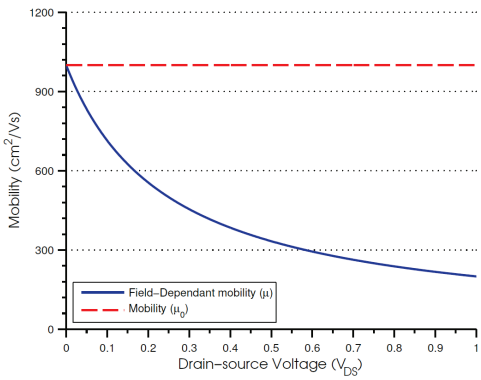


Figure 3 – Field-dependent mobility dependence on the applied voltage.

It is seen that the mobility diminishes with an increased V_{DS}. Within normal operation voltages, the mobility usually does not extend below a third of the zero-field mobility.

2.2.3 Saturation velocity

In recent microelectronic technologies, the drive current is limited by velocity saturation. Under normal operation conditions, the velocity of the carriers in the channel increases with the applied electric field. However, at some point a critical electric field is reached, where the carrier velocity saturates. At this point, any excess electron energy is released by phonon emission.

If the field-independent mobility and the critical voltage, V_c, is known, it is possible to calculate saturation velocity, v_{sat}, by the following expression:

$$v_{sat} = \frac{\mu_0 V_c}{L} \qquad (2)$$

2.3 Gate dielectric

The nanowire is surrounded by a gate electrode, which is isolated by a dielectric material. In order to maximize the gate capacitance and minimize the leakage current between gate and channel, a high-k dielectric, such as HfO_2, is coating the nanowire. The gate dielectric is often deposited by Atomic Layer Deposition (ALD), which is a technique highly suitable for vertical structures [**1**].

2.3.1 Gate capacitance

The gate capacitance is modelled as an oxide capacitance in series with a depletion capacitance. The oxide and depletion capacitances are shown in the band diagram of Figure 4 and are described in more detail below.

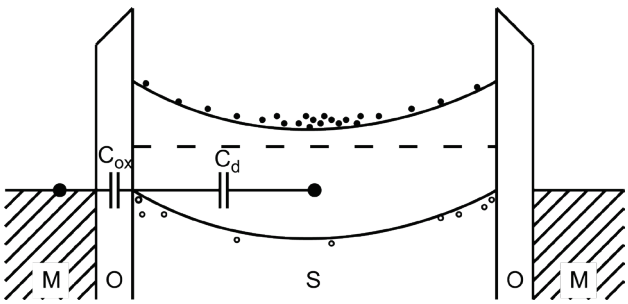


Figure 4 – Band diagram along the nanowire cross section.

Oxide capacitance

The gate oxide capacitance, C_{ox}, is the conventional parameter to model the gate capacitance in MOSFETs. The size of the oxide capacitance parameter determines the number of carriers in the channel for a given gate to source voltage, V_{GS}. For a cylindrical geometry, which applies to a nanowire, the oxide capacitance per unit gate area, C_{ox}, can be expressed as [**2**], [**3**]:

$$C_{ox} = \frac{\varepsilon_0 \varepsilon_r}{R} \ln^{-1}\left(1 + \frac{t_{ox}}{R}\right) \qquad (3)$$

Here ε_0 is the permittivity of free space, ε_r the relative permittivity of the gate oxide, t_{ox} the gate oxide thickness and R the radius of the cylindrical wire. It should be noted that the relative permittivity of the gate dielectric could vary a lot depending on the material choice and processing imperfections.

Depletion capacitance

However, the oxide capacitance is not the only capacitance affecting the channel of the nanowire. A capacitance is formed in series with the oxide capacitance. This is due to the depletion of carriers along the cross section of the wire. The depletion capacitance C_d, is given by[**4**]:

$$C_d = \frac{4\pi\varepsilon_0\varepsilon_s}{R} \tag{4}$$

Here ε_s is the relative permittivity of the semiconductor material. From Figure 4, it is seen that the complete gate capacitance, C_G, is given as C_{ox} and C_d in series:

$$C_G = \frac{C_{ox}\,C_d}{C_{ox}+C_d} \tag{5}$$

Scaling dependence

In Figure 5, the scaling dependence of the oxide capacitance on wire diameter is shown for two different scaling behaviours. In one case, the thickness of the dielectric scales with the nanowire diameter. Furthermore, the gate height is scaled to remain equal to the wire diameter. In the other case, the oxide thickness is held constant.

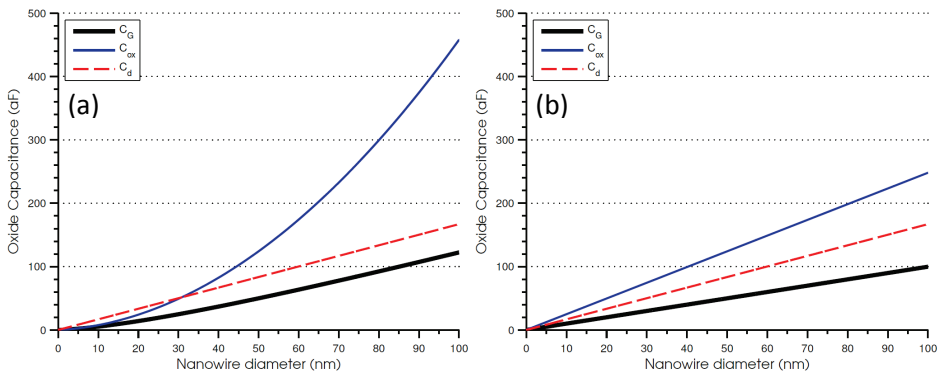


Figure 5 – Oxide capacitance with (a) non-scaling and (b) scaling oxide thickness as the nanowire geometries are scaled.

As can be easily realized, the capacitance decreases linearly with evenly scaled geometries. A more complex scaling relationship is seen when the thickness is held constant, with a behaviour that has both linear and logarithmic parts. It is also not surprising that the diameter dependence of the gate capacitance

mainly follows the smallest component. The both figures share common points at 50 nm diameter. This is expected due to the scaling having the origin at that point.

2.4 Contacts

In order to maximize the gate area, providing best possible electrostatic control, the gate is wrapped around the nanowire. This is an advantage compared to planar structures, since it results in a more effective depletion of the semiconductor. Apart from the gate electrode, metal electrodes also contacts the source and drain regions. The drain is contacted at the top of the nanowire, while the source contact can be of different configurations.

In early developments of these transistors, the substrate was used as contact. This is possible thanks to the high inherent mobility of indium arsenide. More recently, however, there is a trend of actively contacting the source side of the wire. A single nanowire transistor structure, with contacted source side, is illustrated in Figure 2.

The ungated portions of the nanowire contribute heavily to the parasitic resistances. Another contribution is from the metal contacts, which exhibits Schottky barriers . The series resistance affects the operation of the devices by reducing the drive current. Also, as shown in Figure 25, the transistor characteristic shows a more resistive behaviour, which is often far from that of an ideal transistor. Both of these negative effects can be reduced by doping the nanowire and a lot of work is being done in order to make this possible.

2.5 Array configuration

A complete nanowire transistor can consist of either a single nanowire or an array of multiple wires. By increasing the number of nanowires in the transistor, the drive current is increased, while at the same time reducing the series resistance. On the other hand, the increased size of the array, results in higher parasitic capacitances.

An array of multiple nanowires can be described by the unit cell, presented in Figure 6. The geometrics defined in the unit cell layout are used for the modelling of the parasitic capacitances and resistances. A full transistor array is acquired by placing multiple unit cells next to each other.

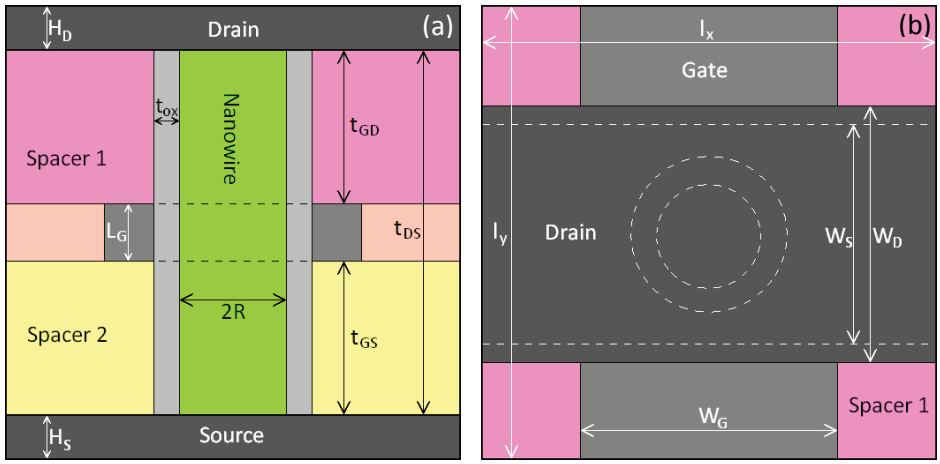


Figure 6 – Transistor unit cell (a) from the side and (b) from above.

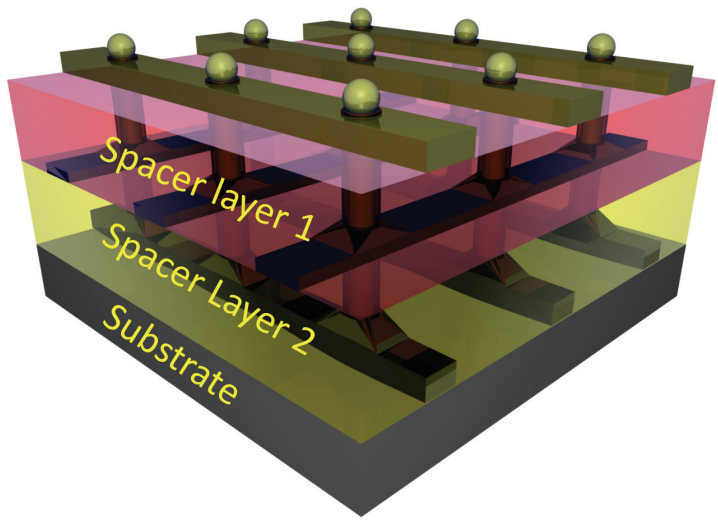


Figure 7 – A WIGFET network array consisting of multiple nanowires. The source contact is at the bottom of the array, while the drain is at the top. The gate is in a networked structure and wrapped around the individual nanowires. The transparent parts are the drain and source spacer layers respectively.

The contacts are arranged in a network formation so that each subsequent layer is perpendicular to the previous. By using a networked electrode structure, the parasitic overlap capacitances are kept low. This allows for a decrease in the thickness of the spacer layers, in turn reducing the parasitic series resistance. In order to simplify the device manufacture, the width of the

contacts may simply be set to the width of the unit cell. However, large contacts are ideally avoided due to the large parasitic capacitances inherent to such a design.

The spacer layers are used to separate the different contacts. They ideally consist of a high-stability, low-k material and most commonly SiO_2 or photo resists are used to construct these layers. A typical 3x3 array is illustrated in Figure 7.

2.6 Band structure

The dimensions of nanowires are on the length scale, where physics based on quantum mechanics is prominent. It is therefore expected that the band structure of the semiconductor material play a large role in description of the carrier transport characteristics. In two of the three spatial directions, the semiconductor is subject to quantum confinement and subbands are formed. These subbands, which may vary in shape, have an energy displacement in respect to each other. The local energy minima of such bands are called valleys and a subband can consist of multiple valleys. An example of a general band structure is presented in Figure 8(a).

Multiple bands contribute to the current through the nanowires. The carriers occupy states in the different subbands and the occupation shows a strong dependence on the energy. Low energy subbands will have a large contribution whereas higher subbands have a smaller one.

It is possible to derive a structural parameter, φ, which is used to scale the current through the transistor by adding up the contribution of all the subbands [**5**]:

$$\varphi = \sum_n \sum_v \left(g_v \sqrt{m_{d,v}}\, \mathrm{e}^{-\frac{E_{v,n} - q\psi_T}{q\phi_{th}}} \right) \tag{6}$$

In the equation above, $E_{v,n}$ is the local energy minimum v of the nth subband. The parameter $m_{d,v}$ is the density of states effective mass of valley v and can be acquired by determining the curvature of a parabolic fitting of the affected valley. The valley degeneracy, g_v, represents the number of different contributions along k-vectors with the same relationship to the quantum

confinement. The relationship between these parameters is defined in Figure 8(b). Furthermore, φ_{th} is the thermal voltage, which is defined as:

$$\varphi_{th} = \frac{k_B T}{q} \tag{7}$$

Here k_B is the Boltzmann constant, T the temperature and q the elementary charge.

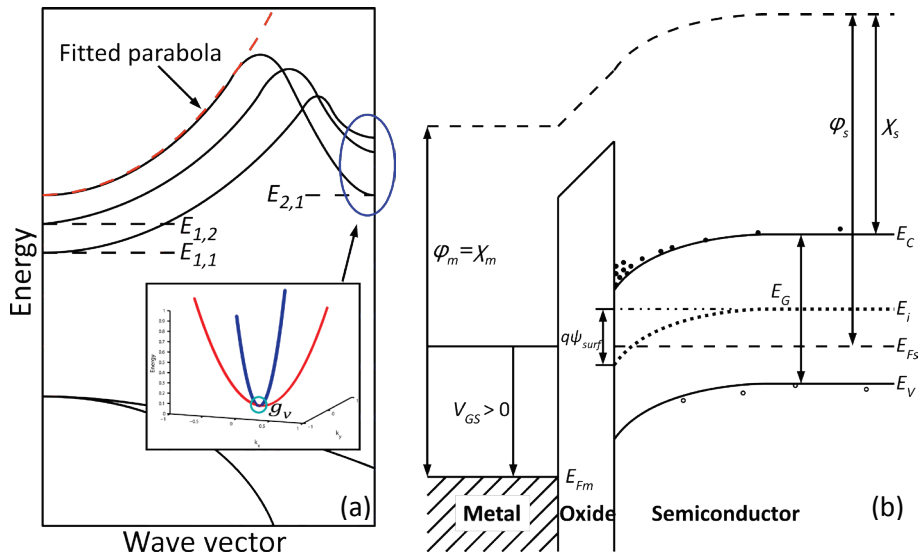


Figure 8 – (a) Band structure with quantum confinement, (b) Energy band diagram of a standard MOSFET.

2.6.1 Surface potential at threshold point

The surface potential is defined as the band bending due to the Fermi level and the applied gate voltage. The surface potential at threshold point, ψ_T, is the band bending when the applied gate voltage is equal to the threshold voltage. Ideally, this potential should be a function only of the nanowire doping concentration [**6**]. However, in this case, it is also dependent on charges in the insulator interface and thus on the processing steps in the production of the transistor.

2.6.2 Work function difference

The work function of a material, is the energy required to move an electron from a state at the Fermi level, E_F, to the vacuum level. The work function difference, φ_{ms}, between the metal (m) and the semiconductor (s), plays a

large role in the scaling of the current levels. Thus, it is often used in physical transistor models.

The work function of a metal can be acquired from tables. The work function of the semiconductor however, is highly dependent on the number of carriers in the semiconductor. The number of carriers is, in turn, dependent on the doping levels in the nanowire. Since not much is known about these levels, no calculation of the work function difference is possible. The definitions of both the relevant work functions are illustrated in Figure 8(b).

However, if the doping levels are quantified in the future, the Fermi level, and in turn the work function difference, can be calculated [**7**]. Two cases need to be considered: n-doped and p-doped materials. For the n-doped material, φ_{ms} is obtained by:

$$E_C - E_F = -kT \ln\left(\frac{N_D}{N_C}\right) \tag{8}$$

$$\varphi_{ms} = \chi_m - \left(\chi_s + (E_C - E_F)\right) \tag{9}$$

In the equations above, χ denotes the electron affinities for the metal and semiconductor, N_D is the donor doping concentration, N_C the conduction band effective density of states (DOS), E_C the energy at the bottom of the conduction band and m_0 the electron mass.

Similarly, for a p-doped material, the work function difference is given by the following set of equations.

$$E_F - E_V = -kT \ln\left(\frac{N_A}{N_V}\right) \tag{10}$$

$$\varphi_{ms} = \chi_m - \left(\chi_s + E_G - (E_F - E_V)\right) \tag{11}$$

Here, the parameter E_V is the valence band energy, N_A the acceptor dopant concentration, N_V the valence band effective DOS and E_G the bandgap energy. The one-dimensional effective density of states, N_{1D}, of both the conduction band and the valence band is calculated by the same expression:

$$\left[N_{1D} = \sqrt{\frac{m_d m_0 kT}{2\pi\hbar^2}}\right] \tag{12}$$

To acquire the density of states, the density of states effective mass is substituted by either the effective electron or hole mass, for conduction or valence band DOS respectively.

In Figure 9, the work function difference dependence on the doping concentration is shown. It is observed that the work function difference is approximately constant at -3.5 eV for n-type doping, while it is around -6.5 eV for p-type doping.

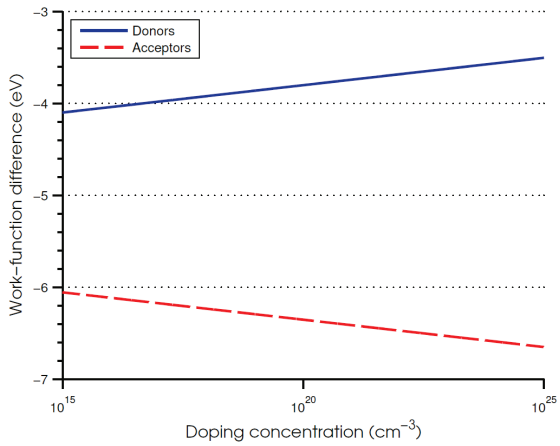


Figure 9 – Work function difference dependence on the doping concentration.

Part III PARASITIC CAPACITANCES

There are many different capacitances associated with the transistor array configuration. These are each contributing to the total parasitic capacitance and they are all modelled and discussed in this section.

3.1 Capacitance types

The capacitances can be divided into four main groups, based on the geometry of their respective electric fields. These groups are overlap capacitances, wire capacitances, fringing capacitances and edge capacitances. They are here presented with their respective subscripts: o, w, f and e. The device structure that forms the basis for the modelling is presented in Figure 10, demonstrating all the modelled parasitic capacitances. The numbers in the figure are explained in Table 1.

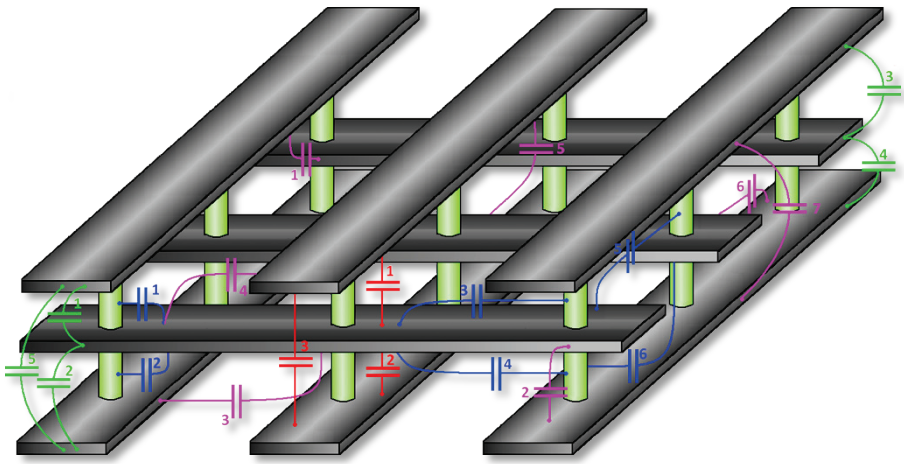


Figure 10 –The different parasitic capacitances in a WIGFET array. The numbers and colours are explained in Table 1.

	1	2	3	4	5	6	7
Red	$C_{GD,o}$	$C_{GS,o}$	$C_{DS,o}$	-	-	-	-
Green	$C_{GD,e,x}$	$C_{GS,e,x}$	$C_{GD,e,y}$	$C_{GS,e,y}$	$C_{DS,e,x}$	-	-
Blue	$C_{GD,w}$	$C_{GS,w}$	$C_{GD,w,y}$	$C_{GS,w,y}$	$C_{GD,w,x}$	$C_{GS,w,x}$	-
Magenta	$C_{GD,f}$	$C_{GS,f}$	$C_{GS,f,y}$	$C_{GD,f,y}$	$C_{GD,f,x}$	$C_{GS,f,x}$	$C_{DS,f,y}$

TABLE 1 – THE DIFFERENT PARASITIC CAPACITANCES IN A WIGFET ARRAY.

3.1.1 General assumptions

In the following sections, each capacitance type is modelled, but for the sake of simplicity, only the general expressions are presented. The full expressions can instead be found in the appendix.

Some simplifications are made in the analysis, in order to get the general expressions. The electrode widths, lengths and thicknesses are all assumed equal. This means that, when referring to the electrode E, either the source or drain can be substituted. Some capacitances are subscripted with an r, which means that the capacitances are affecting adjacent unit cells. For these capacitances, the r can be replaced by either a x or y direction.

Test parameter set

To produce the plots presented in this part, the parameter set in Table 2 is used. This parameter set is consistent with the transistor layout used in the modelling section. The spacing between nanowires, l, are held equal for both directions and all metal electrodes are set at the same thickness, W_E, and width, H_E. The spacer layer thicknesses, t, are also presented with their corresponding relative permittivity.

	l (nm)	W_E (nm)	H_E (nm)	t_{GD} (nm)	t_{GS} (nm)	t_{DS} (nm)
Test parameter set	200	100	50	200	100	350
	t_{ox} (nm)	R (nm)	ε_{GD}	ε_{GS}	ε_{ox}	
Test parameter set	10	50	4.5	2.5	15	

TABLE 2 – PARAMETER SET FOR CAPACITANCE COMPARISON ANALYSIS.

However, the actual spacer layer thicknesses are not used in the equations describing the parasitic capacitances below. Since multiple materials may occupy the space between electrodes, an effective spacer layer thickness, d, is calculated. This is done by using the different relative permittivities of the intermediate materials. As an example, d_{DS} is determined by scaling the drain side spacer with the relative permittivity of the source side.

3.1.2 Scaling rules

It is important to study how the parasitic capacitances change during device scaling, in order to predict what could be gained in terms of performance. Scaling by two different methods is presented for all capacitances in their corresponding sections. In the first scaling method, only the parameters of the nanowire, diameter and length, are changed. However, this is a quite unrealistic case, since it is expected that a better solution would be to scale down all relevant dimensions in the same manner. In the second scaling regime, dimensions such as the electrode widths and height, are scaled linearly with the nanowire diameter.

3.2 Overlap capacitances

The main idea behind the use of the networked layout in Figure 10, is that it is possible to reduce some of the overlap parasitics. There are three different overlap capacitances, acting between pairs of the three metal electrodes: source, drain and gate. To model the magnitude of these capacitances, a standard parallel plate (PP) capacitor model is used.

$$C_{PP} = \frac{\varepsilon_0 \varepsilon_r A}{d} \tag{13}$$

In the equation above, A is the overlap area, ε_0 the permittivity of free space, ε_r the relative permittivity of the intermediate material and d the distance between the parallel plates. By subtracting the cross-sectional area of the nanowire from the overlap of the corresponding electrodes, it is found that the capacitances can be described by the following set of equations.

$$\boxed{C_{GE,o} = \frac{\varepsilon_0 \varepsilon_{GE} \left(W_E^2 - \pi (R + t_{ox})^2\right)}{d_{GE}}} \tag{14}$$

$$\boxed{C_{DS,o} = \frac{\varepsilon_0 \varepsilon_{GS} \left(W_E l - \pi (R + t_{ox})^2\right)}{d_{DS}}} \tag{15}$$

In order for the entire spacer layer to have the same relative permittivity in the drain to source overlap capacitance, the distance between gate and drain is recalculated to obtain an effective distance, d_{DS}.

3.2.1 Scaling

Figure 11 shows how the overlap capacitance is affected by the two different scaling methods. It is seen that the overlap capacitance is reduced with shrinking nanowires and electrodes.

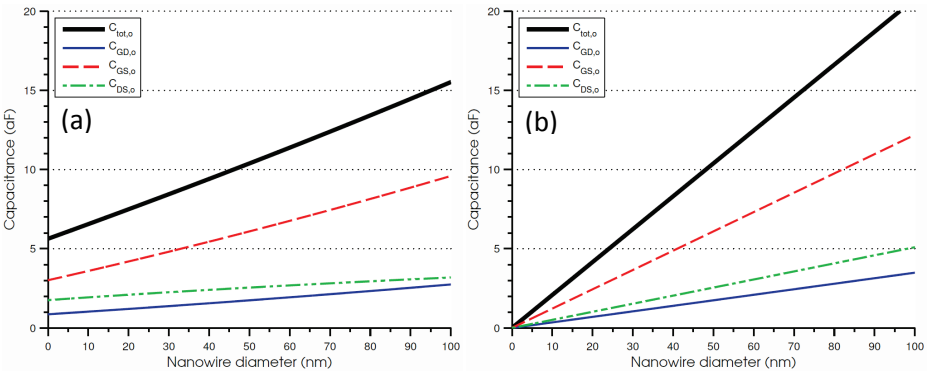


Figure 11 – Overlap capacitances as a function of the nanowire diameter, without device scaling in (a) and with scaling in (b).

3.3 Wire capacitances

A lot of the parasitic capacitance can be attributed to the capacitance between the nanowire and the surrounding gate electrode. There are six such capacitances, as shown in Figure 10: two within the unit cell and four between adjacent unit cells. To model these capacitances, the procedure presented by [**8**] is used.

The capacitances between the nanowires and the gate electrode is modelled with the electric field lines stretching from the sidewall (SW) of one electrode to the top of another electrode. The formula used for all the wire capacitance expressions is presented in (16) and the substitution of the effective electrode height is shown in (17):

$$C_{SW} = \frac{2}{\pi}\varepsilon_0\varepsilon_r \ln\left(\frac{H+T^{'}+\sqrt{S^2+T^{'2}+2HT^{'}}}{S+H}\right) \tag{16}$$

$$\left[T^{'} = Te^{\frac{W+S-\sqrt{S^2+T^2+2HT}}{\tau W}}\right] \tag{17}$$

Here τ is an empirical parameter equal to about 3.7. All the other variables are simply geometrical parameters, which are presented in Figure 12. It is

important to note that (16) gives the capacitance per unit depth. In order to obtain the absolute capacitance, scaling with the smallest of the affected depths has to be done.

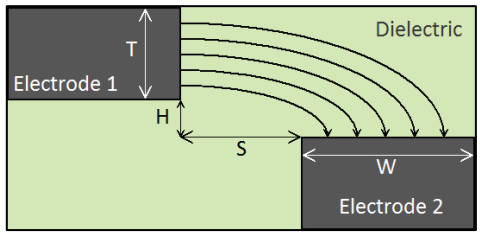


Figure 12 – Geometrical parameters for capacitance formulas.

3.3.1 Geometrical mapping

In order to model the wire capacitances, the gate electrode is split into two subsections: one part circumventing the wire and one part connecting to the next nanowire. The geometrical parameters are simply extracted from the unit cell, but great care has to be taken in order to ensure that the depth scaling is correctly modelled.

The capacitance part that is surrounding the wire is simply scaled with the circumference of the nanowire. However, in order to scale the section that is contacting to the next wire, it is important to realize that not all parts of the wire are affected equally, because of the curved shape. It is proposed that this is solved by first attributing relative contributions by the use of sinusoidal weighting. Thus, the contribution from the middle portion, as seen from the electrode, is one and the contributions from the outer edges are zero. These contributions are added up and scaled with the diameter of the nanowire. This kind of modelling should give a result closer to the real value, than using only the diameter as scaling depth.

3.3.2 The expressions

The two wire capacitances formed within the unit cell is given by (18), with the corresponding substitutions in (19). It should be noted that the depth-scaling factor, due to sinusoidal mapping, is characterized by a scaling factor of π^{-1} in respect to the circumference scaling.

$$C_{GE,w} = 4R\varepsilon_0\varepsilon_{GE}\left(\ln\left(\frac{a_1}{\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}}\right)+\frac{1}{\pi}\ln\left(\frac{a_2}{\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}+\frac{W_E}{2}-R-t_{ox}}\right)\right) \tag{18}$$

$$\begin{bmatrix} a_1 = T_{GE,w,1}+\sqrt{\left(\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}\right)^2+T_{GE,w,1}^2} \\ a_2 = T_{GE,w,2}+\sqrt{\left(\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}+\frac{W_E}{2}-R-t_{ox}\right)^2+T_{GE,w,2}^2} \\ T_{GE,w,1} = d_{GE}\,\mathrm{e}^{\frac{\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}+\frac{W_E}{2}-R-t_{ox}-\sqrt{\left(\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}\right)^2+d_{GE}^2}}{\left(\frac{W_E}{2}-R-t_{ox}\right)\tau}} \\ T_{GE,w,2} = d_{GE}\,\mathrm{e}^{\frac{\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}+\frac{l}{2}-R-t_{ox}-\sqrt{\left(\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}+\frac{W_E}{2}-R-t_{ox}\right)^2+d_{GE}^2}}{\left(\frac{l-W_E}{2}\right)\tau}} \end{bmatrix} \tag{19}$$

There are four different wire capacitances between unit cells, as seen in Figure 10. These have approximately the same geometries as the ones within a unit cell, with the exception that the extra distance has to be accounted for. After derivations, it is shown that these capacitances can be modelled by:

$$C_{GE,w,r} = \frac{4R}{\pi}\varepsilon_0\varepsilon_{GE}\ln\left(\frac{T_{GE,w,r}+\sqrt{\left(\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}+\frac{l-W_E}{2}\right)^2+T_{GE,w,r}^2}}{\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}+\frac{l-W_E}{2}}\right) \tag{20}$$

$$\begin{bmatrix} T_{GE,w,r} = d_{GE}\,\mathrm{e}^{\frac{\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}+\xi+\epsilon\sqrt{\left(\frac{\varepsilon_{GE}}{\varepsilon_{ox}}t_{ox}+\frac{l-W_E}{2}\right)^2+d_{GE}^2}}{\xi\tau}} \\ \xi = \begin{cases}\frac{l-W_E}{2} & \text{if } r=y \\ \frac{W_E}{2}-R-t_{ox} & \text{if } r=x\end{cases} \\ \epsilon = \begin{cases}\frac{l-W_E}{2} & \text{if } r=y \\ l-\frac{W_E}{2} & \text{if } r=x\end{cases} \end{bmatrix} \tag{21}$$

There are no modelled capacitances between the source and drain contact to their respective nanowire section, since they are approximated to be at the same potential.

Scaling

According to the scaling graphs in Figure 13, the wire capacitances are independent on the electrode scaling. This is because the nanowire dimensions are greatly dominating the modelling up to the point where there is no electrode dependence. In reality however, the scaling of the wire capacitances is expected to be even more pronounced. This is because the nanowire starts to shield the electric field lines, when there is a small spacing between them.

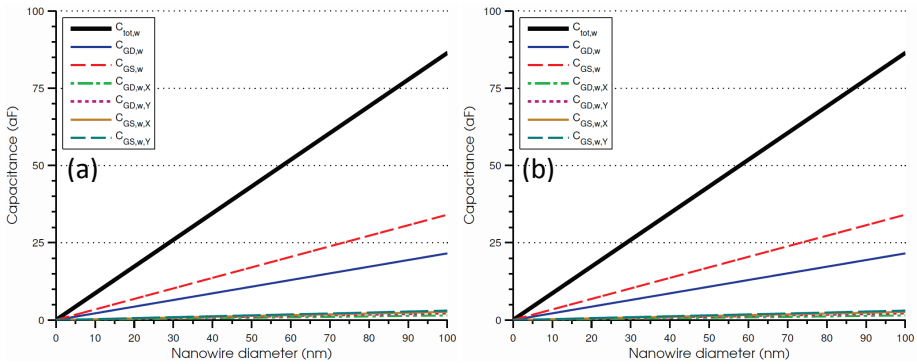


Figure 13 – Wire capacitances as a function of the nanowire diameter, without device scaling in (a) and with scaling in (b).

3.4 Fringe capacitances

There are three different fringe capacitances inside every unit cell and a total of four between neighbouring unit cells. All but one of the fringe capacitances are due to the fringing fields between the edge of one electrode and the flat side of another. These capacitances can be described by the sidewall formula (17). The one that does not conform to the sidewall formula is $C_{DS,f}$, which acts between the long side edges of the drain and source electrodes. These capacitances are described by[**8**]:

$$C_{HC} = \frac{1}{\pi}\varepsilon_0\varepsilon_{di}\left(1+\frac{2W}{S}\right) \tag{22}$$

The distance, S, is realized to be d_{DS}, while W is the thickness of the electrodes. Scaling with the length of the unit cell along the long-sides of the source and drain electrodes, reveals that the capacitance is described by the following expression:

$$C_{DS,f} = \frac{2(l-W_E)}{\pi}\varepsilon_0\varepsilon_{GS}\ln\left(1+\frac{2H_E}{d_{DS}}\right) \tag{23}$$

The other fringe capacitances within the unit cell are found by obtaining the necessary metrics from the unit cell layout and applying them to (16). The two capacitances are presented below:(25)

$$C_{GE,f} = \frac{8W_E}{\pi}\varepsilon_0\varepsilon_{GE}\ln\left(\frac{T_{GE,f}+\sqrt{d_{GE}^2+T_{GE,f}^2}}{d_{GE}}\right) \tag{24}$$

$$T_{GE,f} = \frac{l-W_E}{2}\mathrm{e}^{\frac{H_E+d_{GE}-\sqrt{d_{GE}^2+\left(\frac{l-W_E}{2}\right)^2}}{H_E\tau}} \tag{25}$$

The capacitances between adjacent unit cells are derived by the same principle. However, there are some constraints for the formulas to work, due to limitations in the original derivation. The inter unit cell capacitances are presented in the following equation:

$$C_{GE,f,r} = \frac{4W_E}{\pi}\varepsilon_0\varepsilon_{GE}\ln\left(\frac{c}{d_{GE}+\frac{l-W_E}{2}}\right) \quad \text{If } d_{GE} > \frac{l-W_E}{2} \tag{26}$$

$$\begin{aligned} c &= \frac{l-W_E}{2} + T_{GE,f,r} + \sqrt{d_{GE}^2 + T_{GE,fri,r}^2 + (l-W_E)T_{GE,f,r}} \\ T_{GE,f,r} &= \left(\frac{l}{2}-R-t_{ox}\right)\mathrm{e}^{\frac{H_E+d_{GE}-\sqrt{d_{GE}^2+\left(\frac{l}{2}-R-t_{ox}\right)^2+\left(\frac{l}{2}-R-t_{ox}\right)(l-W_E)}}{H_E\tau}} \end{aligned} \tag{27}$$

Scaling

In Figure 14, the scaling of the fringe capacitance is presented. It is clear that with shrinking wire dimensions, it is better to scale the electrodes as well , instead of keeping them constant. This is a similar observation as that for the overlap capacitances.

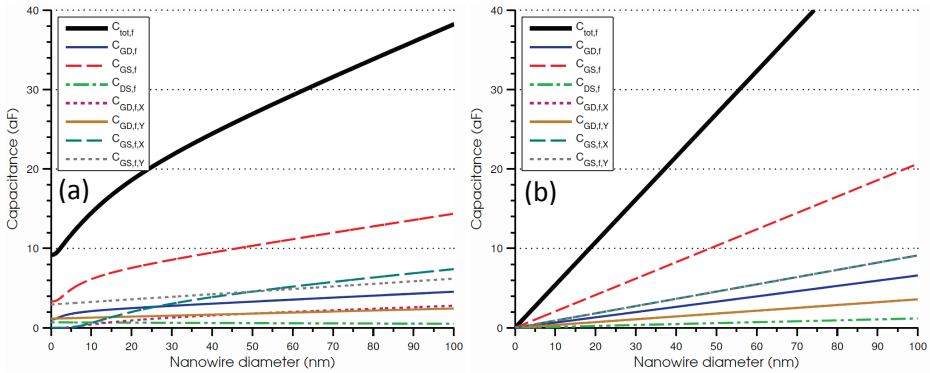


Figure 14 – Fringe capacitances as a function of the nanowire diameter, without device scaling in (a) and with scaling in (b).

3.5 Edge capacitances

In a quadratic array, as shown in Figure 10, there are also some capacitances that only affect the outer unit cells. There are five such capacitances and these can be described by two different formulas.

The drain to source capacitance, at the array edge, is described by the half circle formula. By applying the corresponding geometrics, it is realized that the capacitance is described by:

$$C_{DS,e,y} = \frac{W_E}{\pi}\varepsilon_{GS}\ln\left(1+\frac{2H_E}{d_{DS}}\right) \tag{28}$$

To calculate the four remaining edge capacitances, a different approach has to be applied. This is due to the edge of the gate electrode not being directly below or above the drain or source edges. To solve this, (29) can be used instead [**8**].

$$C_{top,top} = \frac{\varepsilon_0\varepsilon_{di}W}{\frac{H+T}{\alpha}+\pi W\left(\ln\left(1+\frac{2W}{S}\right)+e^{-\frac{S+T}{3S}}\right)^{-1}} \tag{29}$$

$$\left[\alpha = e^{-\frac{H+T}{S+W}}\right] \tag{30}$$

By applying the appropriate geometrics, the following expression for the remaining edge capacitances is derived:

$$C_{GE,e,r} = \frac{W_E \varepsilon_0 \varepsilon_{GE}}{\pi} \frac{2H_E}{\frac{l-W_E}{\alpha_{GE,e,r}} + 2H_E \left(\ln\left(1+\frac{2H_E}{d_{GE}}\right) + e^{-\frac{l-W_E}{6d_{GE}}\frac{1}{3}} \right)^{-1}}$$

$$\left[\alpha_{GE,e,r} = e^{-\frac{l-W_E}{2(d_{GE}+H_E)}} \right] \tag{31}$$

Scaling

Figure 15 depicts the two scaling behaviours of the edge capacitances. As before, shrinking dimensions leads to smaller parasitic capacitances. It can thus be concluded that it is advisable, from a performance point of view, to minimize the geometrical dimensions of the nanowires and the surrounding structure.

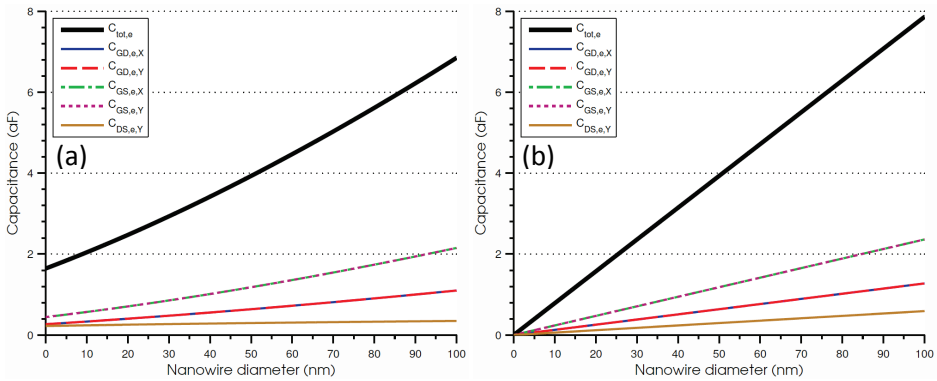


Figure 15 – Edge capacitances as a function of the nanowire diameter, without device scaling in (a) and with scaling in (b).

3.6 Array size dependence

The 21 different capacitances each show different dependences on the number of nanowires, N, in a quadratic array. All the capacitances that operate within the unit cell have a simple linear characteristic. However, the wire dependence for the capacitances affecting adjacent unit cells, is found to be $\left(N-\sqrt{N}\right)$. The edge capacitances, as the name implies, only affect the edge unit cells and therefore has a $2\sqrt{N}$ relationship.

By combining the nanowire quantity scaling behaviour for the different capacitances, the two expressions, (32) and (33) are obtained.

$$C_{GE} = N\left(C_{GE,o} + C_{GE,w} + C_{GE,f}\right) + 2\sqrt{N}\left(C_{GE,e,x} + C_{GE,e,y}\right) + \left(N - \sqrt{N}\right)\left(C_{GE,w,x} + C_{GE,w,x} + C_{GE,f,x} + C_{GE,f,y}\right) \quad (32)$$

$$C_{DS} = N\left(C_{DS,o} + C_{DS,f,y}\right) + 2\sqrt{N}\left(C_{DS,e,x}\right) \quad (33)$$

In Figure 16, the wire dependence is presented for the different capacitance groups and their respective contributions are evaluated.

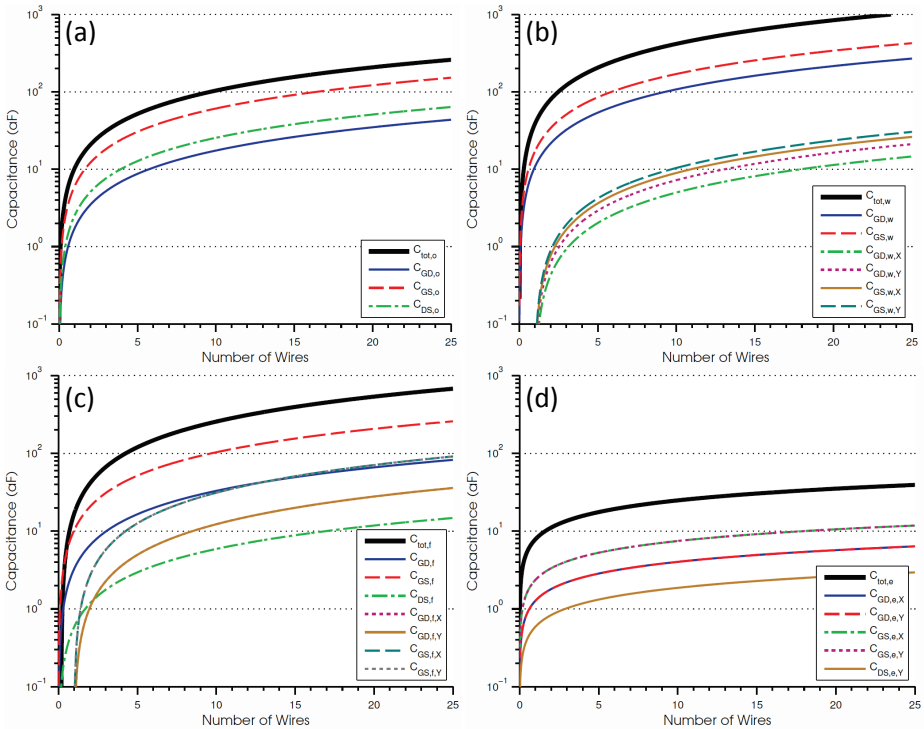


Figure 16 – Comparison between the magnitudes of the different capacitances. (a) Overlap capacitances, (b) Wire capacitances, (c) Fringe capacitances and (d) Edge capacitances.

From Figure 16(a), it is apparent that the gate to source capacitance is the largest of the overlap capacitances. This is followed by the drain to source capacitance, while the gate to drain capacitance is the smallest of the three. The reason for the larger gate to source capacitance is that the spacer layer between those connectors has a higher relative permittivity.

Figure 16(b) shows the wire capacitances and it is clearly observed that the capacitances operating within a unit cell are dominating the ones between

adjacent cells. This is due to the much smaller distances involved for the intra-cell capacitances. The shape of the curves is also differing slightly, which is due to the wire dependence in (32).

The same behaviour that describes the wire capacitances, also describes the shape of the fringing capacitances (Figure 16(c)). This is to be expected, since the same type of expression determines the magnitude of these capacitances.

Lastly, the edge capacitances are shown in Figure 16(d). These capacitances do not vary as much with *N* as the other capacitances. The values are also quite low, which is to be expected due to the small surfaces at the edges.

3.7 Simplifications

By studying Figure 16, it is realized that many simplifications is possible for the total parasitic capacitance expression. There are essentially two parasitic capacitances dominating all the other 19. These are the two wire capacitances operating within each unit cell: $C_{GD,w}$ and $C_{GS,w}$.

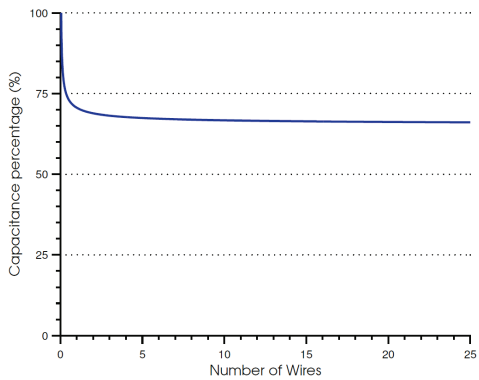


Figure 17 – The quality of the simple analytical capacitance expressions in relation to the full capacitance model.

In Figure 17, the sum of the two capacitances and the intrinsic gate capacitance is compared to the sum of all the others. The figure shows the quality of the simplification as a function of the number of nanowires in the array. Since there is an approximately 70 % correlation between the two, it is deemed a valid approximation to only account for the two wire capacitances in the case of the test transistor. This approximation should still be valid when all transistor dimensions scale equally, since it has been shown that all capacitances scale linearly in this regime.

Part IV TRANSISTOR MODELLING

4.1 MOSFET types

A MOSFET can be operated in different modes, depending on the characteristics of the transistor. Transistors are generally divided into four different main types, depending on their respective charge carrier and operation characteristics.

4.1.1 Majority charge carrier

A first distinction can be made between p-type and n-type devices. This distinction is made based on which charge carrier, i.e. electrons or holes, dominates the conducting channel. Effectively, this means that an n-type transistor becomes more conducting with increasing positive gate voltage, while the conductance of a p-type device increases with a more negative gate voltage.

4.1.2 Enhancement and depletion mode

A further distinction is made between devices that are either conducting or isolating, under zero gate voltage. An enhancement mode device needs an applied gate voltage to form a conducting channel. This is opposed to a depletion mode device, which is conducting at zero applied gate voltage. All the transistors studied in the project will be enhancement mode, n-type devices, because of limitations in the processing technology at this time.

4.2 Regions

The main characteristic of all the simple MOSFET models is quite similar, so the same arguments about their voltage dependences are true for all of them. All these simple models are based on a set of expressions, which govern the current for different regions. Usually three regions are identified and the region that applies depends on the voltages over the transistor.

The first is the cut-off region, in which the gate to source voltage is below the threshold voltage. The threshold voltage is defined as the gate to source voltage, above which the transistor is said to be “on” [**9**]. In this case, the channel is non-conducting and the current is negligible compared to the magnitude of the current in the other regions. Usually, this current is assumed to be zero, but it is possible to use a more accurate model. One example of such a model for the sub-threshold current is presented in the appendix.

However, for the remainder of this report, the sub-threshold current is disregarded.

The second region shows a linear behaviour and applies to the case when V_{DS} is the main controller of the current. This linearity is due to the quadratic term being much smaller than the linear term for small voltages. Essentially, the transistor acts as a simple resistance in this region.

The third and last region is the saturation region. In this region, an increasing drain-to-source voltage, results in only a slight current increase. The saturation may be either in the form of pinch-off saturation or velocity saturation. This depends on which of the corresponding voltages, $V_{GS} - V_T$ or V_{DSAT}, is the smallest [**10**].

4.3 MOSFET Models

In some cases, it is important that the model used to describe the transistor is not too complicated. To get a simple physical picture of the transistor, a model with as few parameters as possible is desired. Additionally, this simplifies the manual, analytical circuit analysis that is done in section 8.3. As a starting point, some traditional MOSFET models, with increasing complexity, will be investigated.

4.3.1 Shockley-Sah

The Shockley-Sah model is one of the earliest and simplest model to describe the drain current, I_D, in a field-effect transistor. The model is defined by the following equations [**11**][**12**][**13**]:

$$I_D = \begin{cases} 0 & \text{if } V_{GS} \leq V_T \\ k\left(V_{GS} - V_T - \frac{V_{DS}}{2}\right)V_{DS} & \text{if } V_{DS} \leq (V_{GS} - V_T) \\ k\frac{(V_{GS} - V_T)^2}{2} & \text{if } V_{DS} > (V_{GS} - V_T) \end{cases} \tag{34}$$

In the above expression, V_T is the threshold voltage. The transistor gain factor, k, is related to process transconductance, k', in the following way:

$$\left[k = k' \frac{W}{L}\right] \tag{35}$$

Here W and L is the channel width and the length respectively. The process transconductance coefficient is simply the product of the oxide capacitance per unit area and the channel mobility:

$$[k^{'} = C_G \mu] \tag{36}$$

The Shockley-Sah model was developed for long-channel devices, which means that it only considers pinch-off saturation. Velocity saturation and other short-channel effects are completely left out. The effectiveness of using this outdated model to describe a nanowire transistor is questionable.

4.3.2 Shichman-Hodges model

A more sophisticated model is the Shichman-Hodges model, defined by the following set of equations [**13**][**14**]:

$$I_D = \begin{cases} 0 & \text{if } V_{GS} \leq V_T \\ k\left(V_{SAT} - \frac{V_{DS}}{2}\right) V_{DS}(1 + \lambda V_{DS}) & \text{if } V_{DS} \leq V_{SAT} \\ k\frac{V_{SAT}^2}{2}(1 + \lambda V_{DS}) & \text{if } V_{DS} > V_{SAT} \end{cases} \tag{37}$$

This model introduces λ as the channel-length modulation and V_{SAT} as the drain-to-source voltage, above which the current is limited by saturation current. This saturation is in the form of pinch-off saturation and is given by:

$$V_{sat} = V_{GS} - V_T \tag{38}$$

The essential point is that this model takes into account the non-ideality of the transistor resistivity in the saturation region. This effect is caused by a growing depletion region on the drain side as the source-drain voltage is increased [**15**]. The magnitude of this effect is modelled by the channel length modulation.

Velocity saturation extension

In the original Shichman-Hodges model, only pinch-off saturation is taken into account. However, the model may be modified to incorporate velocity saturation as well [**10**].

The voltage V_{SAT} in this case is defined as:

$$V_{SAT} = \min(V_{GS} - V_T, V_{DSAT}) \tag{39}$$

In this case, V_{DSAT} is the drain-to-source voltage, for which the critical voltage of the channel is reached. The velocity saturation can be described either by a constant voltage or as a voltage dependent expression [**16**].

4.3.3 Deep submicron MOSFET model

For newer technologies, more and more non-ideal effects emerge, which influences the transistor characteristics. In order to consider these, some further modifications are done to the Shichman-Hodges model. The result is the Deep Submicron MOSFET Model (DSMM), defined as [**13**]:

$$I_D = \begin{cases} 0 & \text{if } V_{GS} \leq V_T \\ k\left(V_{GS} - V_T - \frac{V_{DS}}{2}\right)\frac{V_{DS}(1+\lambda V_{DS})}{1+\frac{V_{DS}}{V_C}} & \text{if } V_{DS} \leq V_{DSAT} \\ k\frac{V_{DSAT}^2}{2}(1+\lambda V_{DS}) & \text{if } V_{DS} > V_{DSAT} \end{cases} \tag{40}$$

Here V_C is the critical voltage, above which the mobility of the channel is degraded. It should be noted that the denominator $(1 + V_{DS}/V_C)$, is a compensating factor for the field dependence of the mobility. The velocity saturation in this model is described by a voltage dependent expression:

$$\left[V_{DSAT}(V_{GS}) = V_C\left(\sqrt{1 + 2\frac{V_{GS}-V_T}{V_C}} - 1\right)\right] \tag{41}$$

In the DSMM, the pinch-off saturation is left out because velocity saturation occur well before the pinch-off saturation condition is met.

The dependence of the saturation voltage on the gate-to-source voltage is shown in Figure 18. It is clearly observed that the saturation voltage follows a square root dependence, with an increasing saturation voltage for higher gate-to-source voltages. As expected, the expression does not apply in the cut-off region.

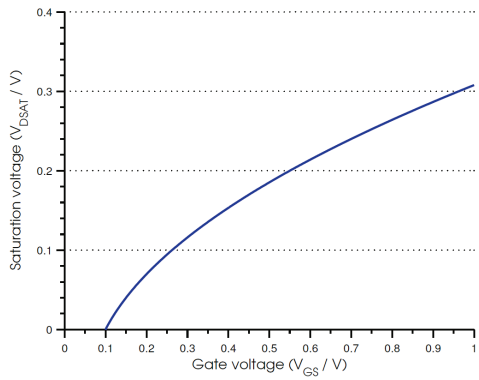


Figure 18 – Saturation voltage dependence on the gate voltage. The critical voltage is set to 0.08 V and the threshold voltage to 0.1 V.

4.3.4 Proper deep submicron MOSFET model

The DSMM suffers from the fact that it cannot be solved analytically with an exact solution, even for relatively simple differential equations. However, by introducing some small modifications it is possible to make it solvable. One approach is using the proper deep submicron MOSFET model (PDSMM):

$$I_D = \begin{cases} 0 & \text{if } V_{GS} \leq V_T \\ k\left(V_{DSAT} - \frac{V_{DS}}{2}\right)V_{DS}(1 + \lambda V_{DSAT}) & \text{if } V_{DS} \leq V_{SAT} \\ k\frac{V_{DSAT}^2}{2}(1 + \lambda V_{DS}) & \text{if } V_{DS} > V_{SAT} \end{cases} \quad (42)$$

It can be seen that the saturation region is the same as in the DSMM. On the other hand, the linear region more resembles the Shockley-Sah model, in that there is no V_{DS} dependence on the channel length modulation [**13**].

4.3.5 Jansson–Berg model

To account for the high resistance in the nanowires, the PDSMM is modified to include a series resistance, R_s. Additionally, the series resistance emulates the same curve form changes as the channel length modulation. This means that the channel length modulation, can be neglected and it is therefore excluded. As the WIGFET transistor in many cases consists of an array of nanowires, the number of nanowires, N, is included as well. Since the model should describe a vertical cylindrical geometry, the gain factor is modified to take the inputs of the nanowire radius, R, and the gate height, L:

$$\left[k = k^{'}\pi\frac{2R}{L}\right] \tag{43}$$

The factor 2π in this substitution is due to the conversion from the nanowire radius to a circumference. The circumference should correspond to the gate width in a conventional, planar MOSFET.

In order to simplify the analytical analysis and circuit calculations, the Jansson–Berg model is formulated by two effective conductance substitutions, $\sigma_L(V_{GS}, V_{DS})$, in the linear region and $\sigma_S(V_{GS}, V_{DS})$ in the saturation region:

$$\left[\sigma_L(V_{GS}, V_{DS}) = \left(\frac{1}{k\left(V_{sat}(V_{GS}) - \frac{V_{DS}}{2}\right)} + R_s\right)^{-1}\right] \tag{44}$$

$$\left[\sigma_S(V_{GS}, V_{DS}) = \left(\frac{2V_{DS}}{kV_{sat}^2(V_{GS})} + R_s\right)^{-1}\right] \tag{45}$$

It will be shown in section 8.3 that these two substitutions converge into a simple σ, dependent only on the supply voltage. With these substitutions made, the model has a remarkably simple expression as shown below:

$$I_D = NV_{DS}\begin{cases}0 & \text{if } V_{GS} \le V_T \\ \sigma_L(V_{GS}, V_{DS}) & \text{if } V_{DS} \le V_{SAT} \\ \sigma_S(V_{GS}, V_{DS}) & \text{if } V_{DS} > V_{SAT}\end{cases} \tag{46}$$

4.3.6 Summary of MOSFET models

In Figure 19, a summary of the different MOSFET models is presented, with an increasing accuracy and complexity.

4.4 Semi-analytical WIGFET model

One of the drawbacks with the simple MOSFET models is that they are somewhat lacking in the physical description of the transport mechanics in the nanowire. Continued scaling of the device dimensions introduces large contributions of quantum effects to the electrical properties. Especially the diameter of the wire is important since it directly affects the band structure, due to quantum confinement [**17**]. With very short gate lengths, there are also many short-channel effects that some of the simplest models do not take into account.

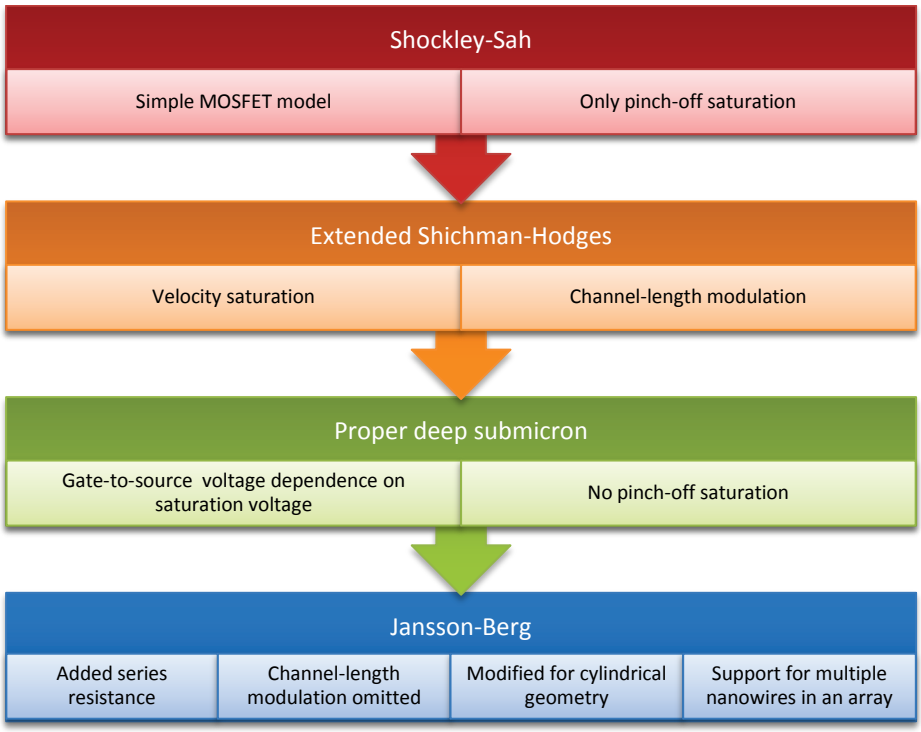


Figure 19 – Flowchart of the different MOSFET models with an increasing accuracy and complexity.

One promising model has been reported by [**5**], which has been deemed interesting for further analysis. Although much more complex than a simple MOSFET model, it has a relatively simple physical description.

4.4.1 Modelling

The model is based on drift-diffusion, instead of the more popularly modelled ballistic transport. The motivation is that device dimensions have yet to reach the limits, in which ballistic transport properties are dominating. The derivation of the model is done by calculating the channel charge, while considering the one-dimensional density of states and Fermi-Dirac statistics.

Some simplifications are made, in order to obtain a closed-form expression. However, these simplifications also reduce the physical description into a more empirical one. In the derivation, the calculations are split up into two regions: the “on”-state ($V_{GS} > V_T$) and the “off”-state ($V_{GS} \leq V_T$). As for the case with the simple MOSFET models, the cut-off solution is discarded and

assumed to be zero. However, a more complete sub-threshold expression is presented in the appendix.

The drain current is obtained by integrating the channel charge with respect to the drain-to-source voltage. Simplifications are done in order to obtain the charge and current, resulting in a few empirical parameters. The equations below describe the closed form expression of the current:

$$I_{DS} = Nk(V_{GS} - \psi_T - \varphi_{ms} + \vartheta(V_{GS})) \cdot \left(V_{DS} + \eta\varphi_{th}\left(1 - e^{-\frac{V_{DS}}{\eta\varphi_{th}}}\right)\right) \tag{47}$$

It should be remarked that the mobility in the process transconductance coefficient is field-dependent, as described in section 2.2.2.

The η-parameter represents a part of the gate-to-source voltage dependence:

$$\eta = 1 + \chi(V_{GS} - V_T)^{\gamma} \tag{48}$$

Here χ and γ are empirical parameters. The parameter $\vartheta(V_{GS})$ is highly dependent on the band structure of the nanowire and is defined as:

$$\left[\vartheta(V_{GS}) = \frac{\lambda_1 + \pi W C_G}{2\lambda_2} - \frac{\sqrt{(\lambda_1 + \pi W C_G)^2 - 4\lambda_2(\lambda_0 - \pi W C_G(V_{GS} - \varphi_{ms} - \psi_T))}}{2\lambda_2}\right] \tag{49}$$

4.4.2 Structural parameters

It is by the λ_j-parameters that the band structure directly affects the device characteristics. These parameters are ideally voltage dependent, but in order to allow for a simple integration, they are assumed voltage-independent. The expression to calculate the voltage-independent λ_j, is given by:

$$\lambda_j = -\frac{\sqrt{2\pi}}{h} q^2 \frac{\varphi_{th}^{1-j}}{j!} \mathrm{Li}_{\frac{1}{2}-j}(-1)\rho_j\varphi \tag{50}$$

An advanced function, known as a polylogarithm, $\mathrm{Li}_s(z)$, appears. This is defined as:

$$\mathrm{Li}_s(z) = \sum_{k=1}^{\infty} \left(\frac{z^k}{k^s}\right) \tag{51}$$

In order to get an accurate description of the channel charge, an infinite number of λ_j have to be applied to the model. In order to acquire a simple expression, only three such parameters are used in combination with correction factors that account for the misalignment. These correction factors, ρ_1 and ρ_2, are empirical and the they should have a value near unity.

Pre-calculation

Many of the λ_j-parameters can be pre-calculated and an internal relation between them can be established. This means that only the band structure parameter, φ, defined in section 2.6, has to be found. The resulting relations are presented below:

$$\begin{cases} \lambda_0 \approx 5.87 \cdot 10^{-5} \cdot \varphi_{th}\varphi \\ \lambda_1 \approx 3.69 \cdot 10^{-5} \cdot \rho_1\varphi \\ \lambda_2 \approx 0.58 \cdot 10^{-5} \cdot \rho_2 \frac{1}{\varphi_{th}}\varphi \end{cases} \tag{52}$$

Part V MODELLING METHODS

5.1 Software

During the course of the project, an in-house MatLab based software suite, providing tools for modelling and circuit design, has been developed. By providing a direct import interface, with support for the xls-files generated by the probe-stations, it is aimed at complete integration and automation of the measurement-simulation workflow. A screenshot of the modelling environment is shown in Figure 20.

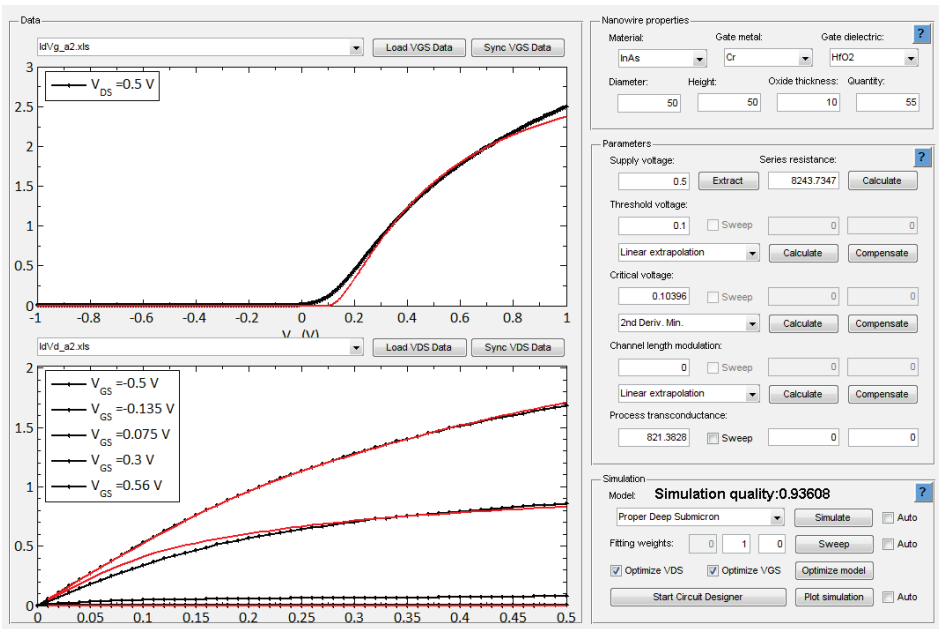


Figure 20 – A screenshot depicting the graphical user interface of the modelling software.

5.1.1 Data correction

As have thoroughly been experienced by the authors, measurement data is often not of an adequately high quality, to be reproducibly modelled. Often, data is deviating, even if it originated from the same sample. This might be caused by the utilization of different measurement setups, use of varying measurement techniques or that measurements are performed during different times. To circumvent this problem, a data correction algorithm is implemented, which provides a method to compensate for these deviations.

Specifically, the data correction is used to match the V_{GS} curve to the V_{DS} curves or vice versa. In this case, it is up to the discretion of the user, to decide which of the curves that is acting as control data. The data correction algorithm searches for values to be corrected in the control data and updates the appropriate parameters of the other data set accordingly. This method is highly effective in providing the fitting methods with unambiguous data, at the expense of a slightly reduced realism and increased uncertainty.

5.1.2 Quality evaluation

For automated software to be able to evaluate the quality of a proposed solution, it needs to be provided with precisely defined metrics of an unambiguous quality. This is especially important when programming intelligent algorithms, as will be discussed further below. Having a quality metric with a direct feedback is also an important tool when doing manual curve fitting. Another important property is that it increases the objectivity of the modelling.

The quality, Q, in this case is simply defined as the mean percentage difference between the simulated, f_S, and measured data, f_D, for all data points, $0 \leq x \leq N$:

$$Q = \frac{1}{N}\sum_{x=0}^{N} \begin{Bmatrix} \frac{f_S(x)}{f_D(x)} & \text{if } f_D(x) \geq f_S(x) \\ \frac{f_D(x)}{f_S(x)} & \text{if } f_D(x) < f_S(x) \end{Bmatrix} \tag{53}$$

5.1.3 Parameter determination

In order to determine the parameters in the transistor models, a range of different methods can be applied. In Figure 21, an overview of a handful of available approaches is shown. The chart is roughly divided into one theoretical and one practical part. The difference is that the theoretical methods can be applied before measurements are made. In contrast, the practical methods are dependent on measurement data for the extraction.

Manual testing

The most easily implemented and straightforward optimization method, to quantify unknown parameters, is using manual identification. Although generally considered very time-consuming, an experienced user can fit model

parameters to measured data in a matter of minutes. This, however, requires an extensive knowledge about the inner workings of the model. Particularly, how different parameters influence, not only the curve shape, but also other parameters. Time consumption is a crucially prohibiting factor when large batches of measurement data need to be quantified and properly modelled. Batch processing of different measurement series is generally a requirement, in order to increase the model confidence and being able to determine the behaviour of process-independent variables.

Parameter calculation

Many model parameters are possible to calculate theoretically, or alternatively being extracted from tables. It is convenient if as many parameters as possible can be pre-calculated, as this limits the need of processing power. Additionally, pre-calculation consolidates the connection of the model to underlying physics. An example of a model relying heavily on parameter calculation is the Compact Circuit NWFET model described in section 4.3.6.

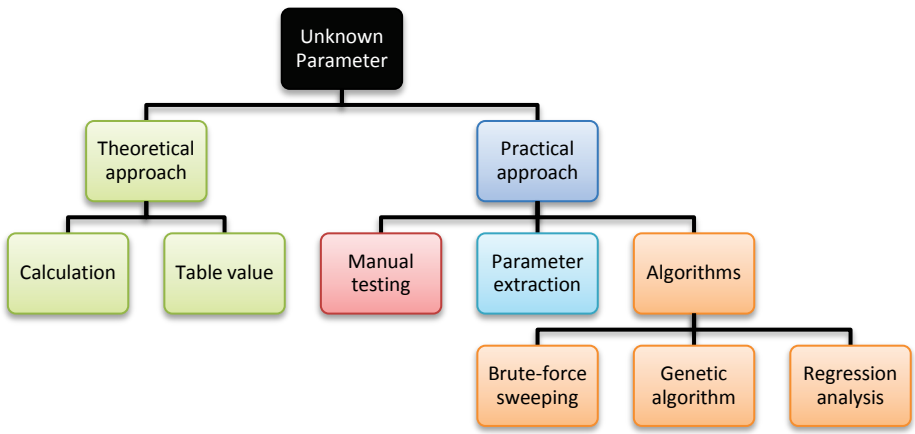


Figure 21 – Hierarchical chart showing a summary of the different applied parameter determination methods.

Parameter extraction

Other models are constructed around parameter extraction directly from the measurement data curves. An example of a model relying heavily on parameter extraction is the well-known Shichman-Hodges model, utilizing conventional parameters such as threshold voltage and channel length modulation. For a detailed walkthrough of different extraction methods for device parameters, see section 5.2.

Brute-force sweeping

For parameters that are neither extractable from measurement data, nor possible to calculate, manual testing is an alternative. However, this method is very difficult and time consuming and when processing large batches of data, manual testing quickly becomes an unmanageable task. The key to successful batch processing is automation. Automation ensures minimal user interaction and maximum reproducibility, as it eliminates user partiality and preconceptions.

The simplest and most apparent method to automate a parameter fitting process, is sweeping the different parameters within plausible limits, followed by an evaluation of the simulation quality. The problem with this approach is that the execution time of such a sweep grows exponentially with the number of parameters. This means that it rapidly becomes impractical, if not entirely unusable, for more than a few variables. In order to decrease the number of parameter sweeps, a combination with other extraction methods is recommended for as many parameters as possible.

Intelligent algorithms

In order to decrease the number of tested parameter sets, more refined algorithms need to be developed. In this work an algorithm is proposed, utilizing random steps within an intelligently defined search area. The basic flow during execution begins with a random step of all swept parameters:

$$A_{n+1} = A_n \cdot \left(1 + \mathrm{rand} \cdot (1 - Q_n)\right) \tag{54}$$

Here A is the parameter value, *n* the recursive index, rand a random number and *Q* the curve quality in comparison with the measurement curve. The random number is in the simple case, just a rectangular distribution between -1 and 1.

After each of the iterations, a quality value is calculated. If this new quality is higher than the last, the value is saved and the iterations continue. However, if the quality is worse, the recursion is rolled back one step and new random numbers are generated. This results in a narrower search interval for an increasing curve quality, allowing further quality refinement.

By using this algorithm, it is possible to achieve a near-perfect fitting for four parameters in fewer than 100 iterations. One important advantage of this algorithm is scalability; it could easily be adopted to sweep tens of parameters with realistically achievable calculation times.

Any multi-variable system show increasingly complicated behaviour with the number of variables involved in the problem. A complex system often has a large quantity of local minima and maxima. This constitutes a problem in any stochastic optimization process, when the global optimum is wanted. For a random number based solution, the optimization variable is most likely to approach the nearest minima or maxima as schematically shown in Figure 22. However, there is no guarantee that this is the best possible optimization.

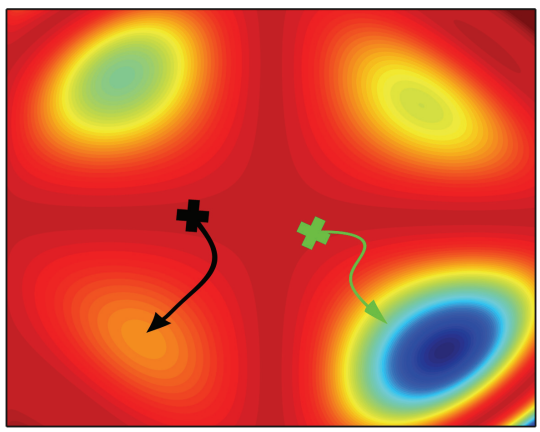


Figure 22 – Randomly placed points tend to approach the nearest minima. In the case of the black point it may get stuck at a local minimum. The green point however, finds the global minimum.

By changing the distribution of the random number generator from rectangular to another distribution, the algorithm may be further refined. A commonly used distribution is a normal distribution with an expectation value of zero. This allows for a more focused stepping in addition to a small probability of steps up to infinity. This means that the time for refining an optimum is decreased, while at the same time allowing a finite probability to skip any local minima.

Regression analysis

A powerful method for fitting arbitrary, theoretical models to measurement data is the built-in function in MatLab, called lsqcurvefit. This works by using regression-analysis to fit a predefined function to data, such as a transistor model, by using the least-square method. However, this tool is not included in the standard MatLab package and in order to avoid potential licensing problems, it is not used in the standard version of the modelling suite. For users that have access to this function, the modelling can be speeded up considerably.

5.2 Parameter extraction methods

Essentially, there are only a small number of parameters that there is a need to estimate in the fitting process. Many of the parameters may be extracted directly from data, including threshold voltage, saturation voltage, channel-length modulation and series resistance. However, care should be taken, as the extraction may not always result in the best possible parameters for the model fitting.

5.2.1 Threshold voltage

The threshold voltage is one of the most commonly used parameters in the characterization of a transistor. There are many methods to determine this parameter [**18**], [**19**], but the following has been decided due to its simplicity and robustness. It is proposed that the threshold voltage may be extracted by finding the maximum transconductance from the $I_D - V_{GS}$ characteristics and extrapolating the linear slope around that point down to the $V_{GS} - \text{axis}$ [**20**], as illustrated in Figure 23(a).

The transconductance is defined as:

$$g_m = \frac{\partial I_D}{\partial V_{GS}} \tag{55}$$

5.2.2 Saturation voltage

The saturation voltage might be qualitatively decided directly from the $I_D - V_{DS}$ characteristics. However, this becomes increasingly difficult for large channel-length modulations or non-ideal curve shapes. Drain currents above and below the saturation point exhibit, a more or less linear dependence on the drain-to-source voltage. This makes it possible to distinguish these regions from the saturation point by taking the second derivative of the $I_D - V_{DS}$ characteristic and extracting the minimum point. An example of one of these extractions can be seen in Figure 23(b). By using this method, no distinction is made for the type of saturation, i.e. velocity saturation or pinch-off saturation.

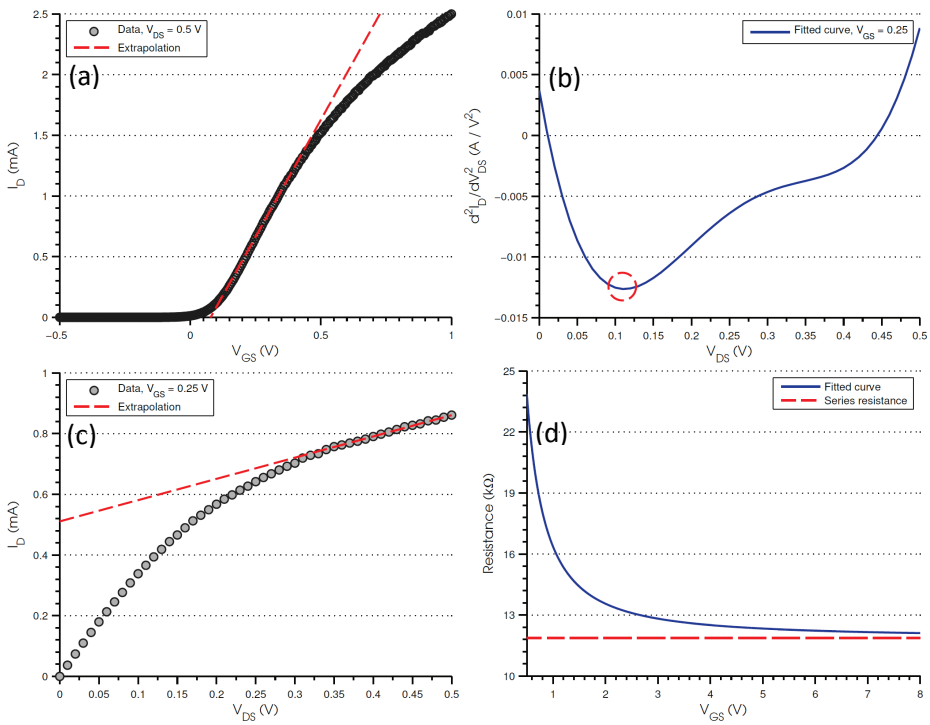


Figure 23 – Method used to extract the threshold voltage (a), the saturation voltage (b), the channel-length modulation (c) and the series resistance (d). It should be noted that in order to acquire a noise immune derivative curve, a polynomial fitting is used on the measurement data.

Critical voltage

For devices with sub-100 nm gate lengths, it can be assumed that the saturation is caused by short-channel effects. In the case of the PDSMM, the saturation can be described by (41). This means that the critical voltage, V_C, can be calculated if the saturation voltage is known. Isolating V_C in this expression results in the following equation:

$$V_C = \frac{-V_{DSAT}^2}{2(V_{DSAT} - V_{GS} + V_T)} \tag{56}$$

5.2.3 Channel-length modulation

The channel-length modulation is also extracted by linear extrapolation, but in this case from the saturation region. The extraction of the channel-length modulation parameter, λ, is done by determining the slope of the extrapolated curve, as seen in the following equation:

$$I_D = I_D^{'} + \lambda I_D^{'} V_{DS} \tag{57}$$

An example of this linear extrapolation is presented in Figure 23(c).

5.2.4 Series resistance

Extraction of the series resistance is done by deriving the equivalent resistance of the transistor by dividing the measured drain-to-source voltage with the transistor current. This resistance is then plotted against the gate voltage and is extrapolated towards infinity utilizing the least square method to do a power regression analysis. The least square method is a powerful statistical method to find approximate solutions when dealing with over-determined systems. In the regression analysis, the values R, a and b are determined:

$$\frac{V_{DS}}{I_D} = R_s + aV_{GS}^b \tag{58}$$

It can be seen that the resistance, R, should be equal to the series resistance, R_s, if the exponent, b, is negative and the gate voltage approaches infinity. The resistance for the transistor is assumed negligible when it is completely on. An example of this extrapolation is shown in Figure 23(d).

Compensation

When modelling the transistor behaviour, it is important that the different parameters are not dependent on each other. As the series resistance influences the curve form as well as the nature of the parameter values, it is necessary to compensate the parameter extraction methods for this effect. To achieve intrinsic parameters, the series resistance is first extracted, followed by its subtraction from the measurement data. From this new curve, it should be possible to calculate parameters, not dependant on the series resistance. However, this is not always the case as the curve form may show unexpected behaviour due to artefacts from the series resistance subtraction. The series resistance compensation in the software should thus be used, only with uttermost caution.

Part VI NWFET PARAMETERS

6.1 Quality of MOSFET-models

In this part of the report, the results of the model fitting to the NWFET data are presented. Before a more complete analysis can be done for the transistors, it is evaluated which one of the simple models that best represent the transistor. In Figure 24, the qualities of the different MOSFET models are evaluated. It can be seen that the Shockley-Sah model overestimates the current, probably due to exclusion of velocity saturation. The model is therefore deemed unusable. Acceptable results are achieved with both the Shichman-Hodges and Proper Deep Submicron models, with correspondence of about 89% respectively 92%. As the Proper Deep Submicron model shows a significant increase in simulation quality, as well as providing better scalability, it is decided to be the primary model for the analytical analysis. It should be noted that all these models are modified with the Jansson-Berg modifications.

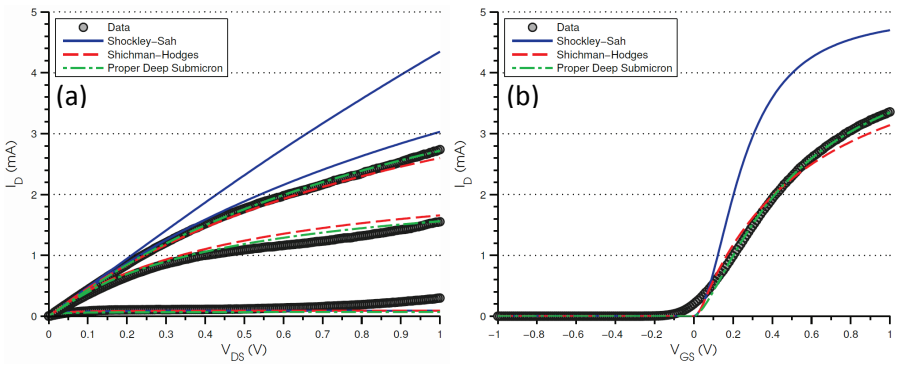


Figure 24 – (a) and (b) shows a comparison between the different simple MOSFET models for the V_{DS} and V_{GS} characteristics respectively.

6.1.1 Importance of series resistance

The impact of the series resistance is clearly illustrated in Figure 25, showing the PDSMM with and without modification for series resistance. It is apparent that no significant results can be achieved by ignoring this resistance. This result indicates that if the series resistance could be minimized, a large increase in drive current is to be expected.

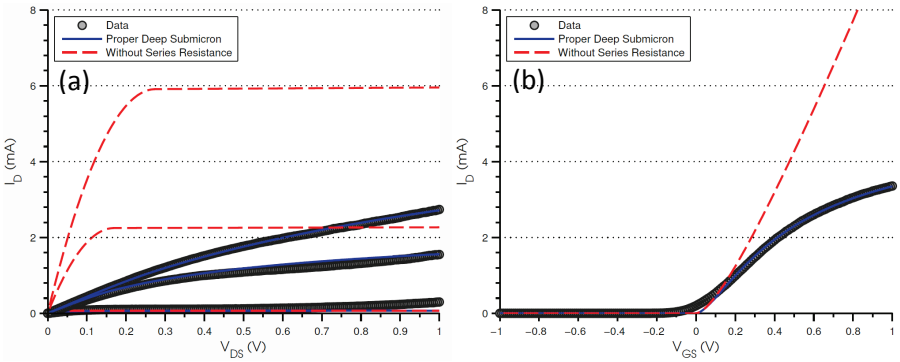


Figure 25 – (a) and (b) depicts the impact of added series resistance.

6.2 PDSMM Characterization

By using the intelligent algorithm to optimize an extended range of modelling samples, a complete characterization of the transistor behaviour is acquired. A summary of the derived parameters are presented as a box plot in Figure 26, illustrating the data range and quartiles as well as the median values marked with dots. Figures that are more precise are presented in Table 3. The quality of the fittings, for the corresponding data sets, ranges between about 87% and 96%.

6.2.1 Analysis

It can be seen that the deviation is reasonably low, considering the unrefined state of the processing techniques. Especially, the material dependent parameters, V_C and k, have a relatively low deviation in respect to the mean values. The data set is polluted by the presence of some samples at extreme values, which is probably due to failures in the processing or measurements. The largest problem, from a circuit design perspective, is the limited control of the threshold voltage, which ranges from negative to positive values.

Another limiting factor is the high series resistance, which later will be shown to be severely degrade the circuit performance. The channel-length modulation factor, λ, is found to be very low for most transistor. This indicates that it is a valid approximation to neglect this effect. The high channel-length modulation for a few devices is probably due to failures in the series resistance extraction or modelling.

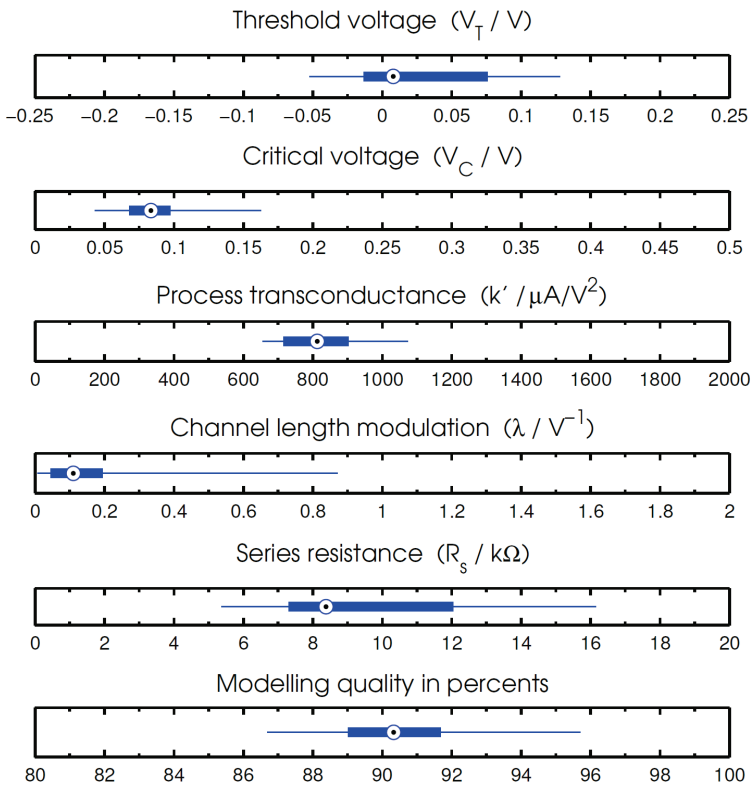


Figure 26 – Box plot illustrating the distribution of parameter values.

	V_T (V)	V_C (V)	k' (μA/V^2)	λ (V^{-1})	R_s (kΩ)
Mean	0.028	0.088	830	0.20	9.8
Median	0.008	0.083	810	0.11	8.4
Min	-0.052	0.043	660	0.01	5.4
Max	0.128	0.163	1070	0.87	16.2

TABLE 3 – SUMMARY OF PARAMETER VALUES FOR THE FITTED CURVES.

6.2.2 The test transistor

To progress with the circuit design, a benchmarking set of parameters is determined, inspired by the parameters in Table 3. As the Jansson-Berg model predicted, the channel-length modulation can be negligible in favour to the series resistance. The exclusion of the channel-length modulation from the

parameter set, simplifies the following analytical analysis. The other values are rounded to improve remembrance and are from now on denoted as the test transistor parameters, as presented in Table 4.

	V_T (V)	V_C (V)	k' (μA/V^2)	λ (V^{-1})	R_s (kΩ)
Value	0.1	0.08	800	0	8.0

TABLE 4 – THE TEST TRANSISTOR PARAMETER SET.

6.2.3 Characteristics

By using the test transistor parameters from Table 4, a more detailed transistor characteristic is constructed. This is illustrated in Figure 27.

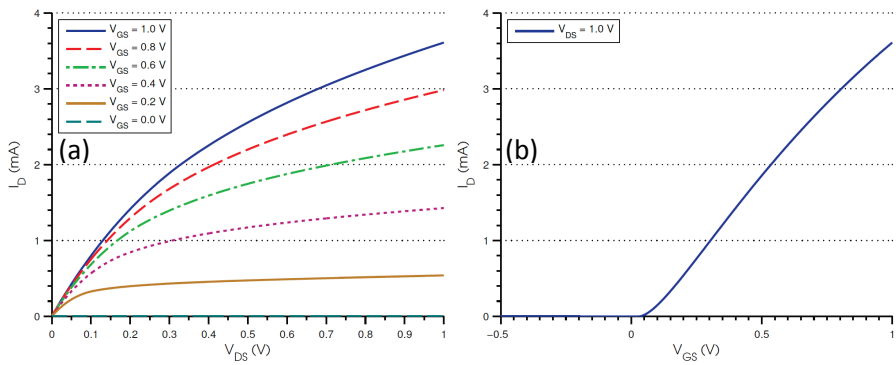


Figure 27 – Test transistor DC characteristic plotted against (a) drain-source voltage and (b) gate-source voltage.

6.3 WIGFET Model

Since there are only four different empirical parameters to account for in the model, fitting to the data should be relatively unproblematic. All other physical parameters can be either measured or calculated. However, due to the complexity of the calculations and uncertainties of physical data, not all of these are solved conventionally. A more insightful analysis is complex enough to be suitable for an independent project and is out of the scope of this thesis. The primary focus is to evaluate the suitability of the model to be used in future WIGFET characterizations.

6.3.1 Band structure

The band structure for the nanowires with a diameter of 50 nm is not calculated. However, calculations have been presented in e.g. [**17**], for nanowires up to 25 nm in diameter. As the band structure is not calculated, the structural parameters, λ_j, has to be empirically fitted, in order to characterize the transistor.

6.3.2 Doping concentrations

To calculate the work function difference, the doping concentration has to be extracted from measurements. This data cannot be produced, as of the moment of writing, but it is estimated that the needed measurements can be performed sometime in the near future. Until then, the work function difference also has to be tested during the fitting process. The calculation of the surface potential at threshold point, suffers from the same problem. The necessary data simply does not exist yet, implying that this parameter as well has to be empirically fitted.

6.3.3 Model fitting

The total number of unknown parameters needed to model the transistor behaviour thus adds up to nine. However, with the derived relation between the different λ_j-parameters established, this number is reduced. In addition, the work function difference and surface potential at threshold point can be combined into one parameter. These two simplifications imply that six parameters have to be empirically decided.

It is easily realized that such as large number of arbitrary parameters results in a huge amount of possible solutions to the problem. Even so, some good qualified guesses can be made. Initial analysis shows that the curvature of the model can be made to match the measurement data closely, with a mobility and saturation velocity around 550 $\mathrm{cm}^2/\mathrm{Vs}$ and 200 $\mathrm{km/s}$, respectively. These values are physically realistic and close to the values extracted from the Jansson-Berg model.

It is found that a complementary, simple model, such as the Jansson-Berg model, simplifies the model fitting considerably. This is due to the possibility to extract e.g. the mobility and saturation velocity from this model, thereby reducing the number of unknown parameters. This is a functionality that

should be further investigated in the future followed by a more intricate implementation into the automated workflow.

The important parameters used in the final modelling with the NWFET model are presented in Table 5:

	V_T (V)	μ_0 ($cm^2V^{-1}s^{-1}$)	v_{sat} (m/s)	C_G (aF)	R_S (kΩ)
Value	0.09	1300	$2.0 \cdot 10^5$	6.4	8.0
	$\psi_T + \varphi_{ms}$ (V)	φ	χ (V^{-1})	γ	ρ_1	ρ_2
Value	0.18	$7.1 \cdot 10^{-5}$	0.5	1	1	1

TABLE 5 – PARAMETERS USED IN THE NWFET MODEL FITTING.

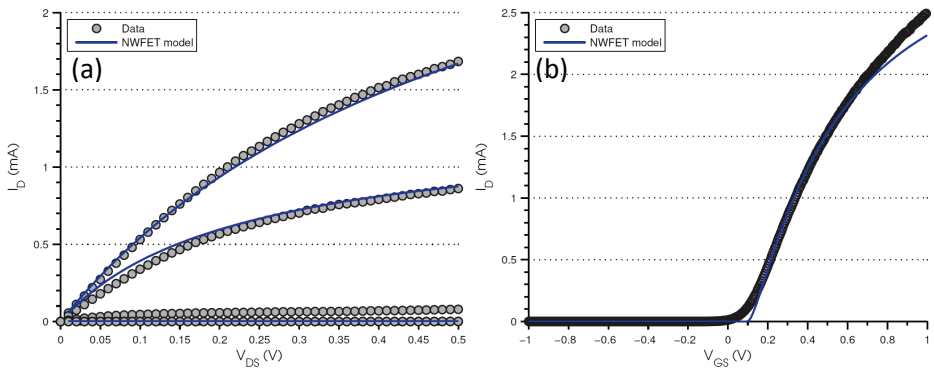


Figure 28 – NWFET model characteristics created form manual fitting to measurement data

6.3.4 Quality

Figure 28 shows an example of a data fitting to the model. It is deemed extremely time consuming to estimate parameters, resulting in an adequate fitting quality. However, there is no reason why this model should not be able to characterize WIGFET transistors properly. The modelling process will be immensely improved if some additional measurement data is obtained and a calculation of the band structure is done.

6.3.5 Characteristics

Because of time constraints, the analysis of the NWFET-model is not as detailed as that for the PDSMM. This is mainly due to the many unknown

parameters that need to be empirically fitted. The model is not represented in the modelling software in a complete form but the framework is there to incorporate it easily. The same is true for most other model that could be of interest.

The analysis of the measured nanowire transistor data with the Analytical Compact Circuit model for Nanowire FETs has resulted in the output characteristics, shown in Figure 29. It should be noted that the current shows more saturation for an increasing gate-to-source voltage, than in the PDSMM. This is most likely due to the voltage dependence on the mobility in the NWFET model. It is probable that the behaviour of the NWFET model is the more realistic of the two.

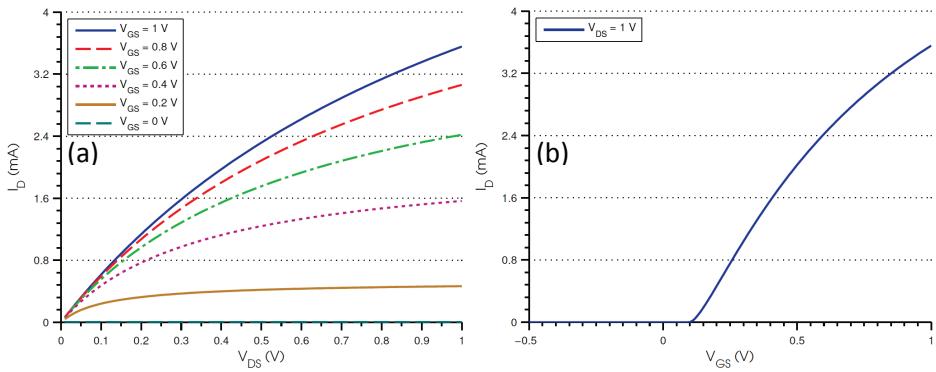


Figure 29 – The full NWFET model characteristics from this analysis.

Part VII CIRCUIT DESIGN

7.1 Logic families

Digital circuits can be based on a number of different logic families, offering different switching performance, power consumption and integration complexity.

7.1.1 CMOS

The conventional logic family for digital circuits is CMOS, which offers both high switching speeds and low power dissipation. A requirement for CMOS is the availability of both n- and p-type transistors. As the nanowire transistors studied during the course of this project has no p-type counterparts, other logic families need to be considered. The most simple and apparent choice for this task is ratioed logic, consisting of only n-type transistors.

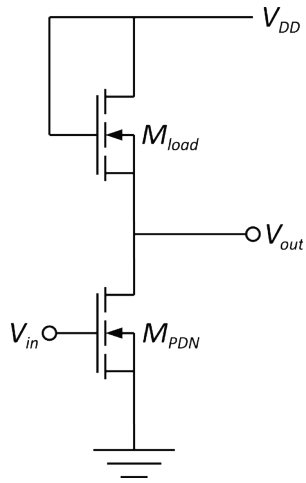


Figure 30 – A ratioed logic inverter in n-type, enhancement mode

7.1.2 Ratioed logic

Ratioed logic is based on the principle of using a load transistor instead of a pull-up network. The load transistor can be either n- or p-type and of either enhancement or depletion mode.

One of the drawbacks with ratioed logic is an increase in static power dissipation in comparison to CMOS. Since there is no p-type device that can block the current path when the pull-down network is conducting, a direct path is formed from the supply voltage to ground. One advantage, aside from not needing p-type transistors, is that ratioed logic offers a lower total number of transistors in comparison with e.g. CMOS logic [**10**].

7.2 The inverter

From now on, a ratioed logic inverter is studied, using a load transistor of n-type in enhancement mode. The ideal circuit diagram for such an inverter is presented in Figure 30. In this case, the gate of the load transistor is connected directly to V_{DD} at all times, as seen in Figure 30. In theory, the high output level should be equal to $V_{DD} - V_T$, since an n-type transistor cannot

pull an output voltage up to the supply voltage. The low output level does not reach zero because of the voltage division between the load transistor and the transistors in the pull-down network. The input signal controls the current flow of the transistor in the pull-down network and it is easily verified that the output node is inverted compared to the input voltage.

7.3 Figures of merit

While benchmarking devices, it is important to study well-defined and conventional figures of merit. This allows easy comparison between different studies and publications. In this project, the relevant figures of merit are given below.

7.3.1 Propagation delay

An important timing metric, considering the switching speed performance of a logic gate, is the propagation delay, τ_p. It is usually defined as the time between 50 % of the input signal to 50 % output signal strength [**10**]. There may be many different propagation delays for different transitions of a logic gate. In that case, a mean value between these diverse values is used to represent the overall propagation delay. In the case of a simple inverter, the propagation delay is separated into a low-to-high delay time, $\tau_{0\to1}$, and a high-to-low delay time, $\tau_{1\to0}$. Thus, the propagation delay for an inverter is given by the following expression:

$$\tau_p = \frac{\tau_{0\to1}+\tau_{1\to0}}{2} \tag{59}$$

7.3.2 Power dissipation

In the microelectronics of today, the power dissipation is an increasingly important performance metric. This is mainly due to the scaling of CMOS devices, dictating that the power consumption per device should be held constant while shrinking the device size. Consequently, an ever-increasing integration leads to higher power dissipation per chip area. The power consumption is usually divided into static and dynamic power consumption for the switching and stationary state respectively.

Average power consumption

In the case of a ratioed logic design, the static power consumption is very high. This is due to the formation of a direct current path from the supply voltage to ground, for a high input signal. In the case of a CMOS inverter, the dynamic power should be considered. However, in the case of a ratioed logic design, the static power consumption dominates and the contribution from the switching may safely be omitted.

To quantify the power consumption of the ratioed logic inverter, the time average, static power consumption is determined for a random input signal. The power consumption of a transistor is conventionally defined as the current multiplied by the source-to-drain voltage.

$$P_A = V_{DS} I_D \tag{60}$$

Peak power consumption

In some cases, it might be more relevant to study how large the power consumption is at its peak. It might be that interconnects and other circuits must be dimensioned to handle large power peaks. However, this is not investigated further, as the peak power consumption does not have a significant impact on the ideal device performance.

7.3.3 Power-delay and energy-delay product

In order to incorporate both the performance and the power consumption into a combined figure of merit, the power-delay product (PDP) can be used. It represents the energy it takes for a switching to occur. The power-delay product is defined as the product between the propagation delay and the average power consumption:

$$\mathrm{PDP} = \tau_p P_A \tag{61}$$

For a better metric in deciding the overall performance of the inverter, in terms of both energy dissipation and switching speed, the energy-delay product (EDP) is used. The energy-delay product assigns additional weight to the propagation delay, by weighting it with an exponent of 2. [**10**]:

$$\mathrm{EDP} = \tau_p^2 P_A \tag{62}$$

This quantity is used in this project as the ultimate performance metric. The energy-delay product is used to find the parameters, which gives the best possible performance in terms of switching frequency, while at the same maintaining a relatively low power consumption.

7.3.4 Regeneracy

Regeneracy is an important property for logic gates, as it allows the circuit to operate without amplification for arbitrarily large designs. The principle is that the threshold voltage should be above V_{IL} and below V_{IH}, in order to avoid an opening of the next stage, causing unwanted switching. [**10**].

One method to determine regeneracy is to study a ring oscillator. A ring oscillator should only oscillate for devices, which have a regenerative property. This is the method used in the numerical solution, although the ring oscillator is approximated by a linear chain of inverters. Alternatively, the regeneracy can be derived through the static output voltage levels. This is the method of choice in the analytical solution and will be demonstrated in section 8.8.

Finally, the regeneracy can be determined by studying the voltage transfer characteristic (VTC). The VTC should have two stable operating points: above and below the V_{IH} and V_{IL} respectively. Additionally it should be unstable between these values. By differentiation of the VTC, the gain characteristic relative to the input voltage can be found. It is thus easy to control if the absolute slope is below or above unity gain for the different regions [**10**]. This is a good method to use in circuit simulation environments such as Cadence.

Part VIII CIRCUIT ANALYSIS

8.1 Problem formulation

The inverter design has already been described in section 7.2, but to make a full analytical analysis possible, parasitic capacitances and resistances must be considered. This, more complex, circuit diagram is depicted in Figure 31(a). It should be noted that the series resistances are already incorporated in the Jansson-Berg transistor model used for this analysis.

In the analytical expressions describing the inverter properties, all capacitances affecting the output node are lumped together in a single capacitance, C_{load}. This load capacitance greatly influences the performance of the inverter, in terms of both power consumption and propagation delay.

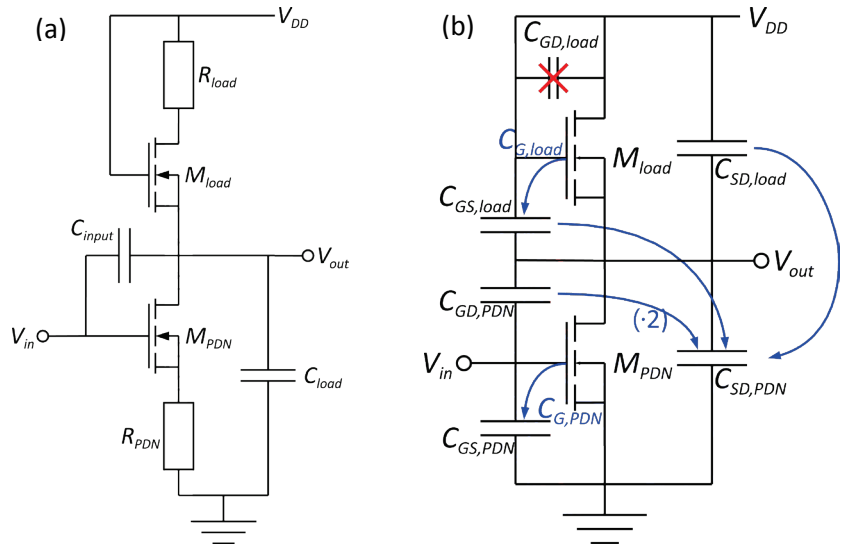


Figure 31 – (a) Extended circuit diagram of the ratioed logic inverter. (b) Merging of the capacitances.

8.1.1 Nodal analysis

To derive the output of the circuit, the task is to extract the voltage of the output node, V_{out}, as a function of the input voltage, V_{in}. One way to define this problem is to use nodal analysis. Kirchhoff's first rule, also known as the current law, states that sum of all incoming and outgoing currents from a node is zero. By using this rule, a differential equation is produced, which states how the output is related to the input potential:

$$\frac{dV_{out}\,(t)}{dt} = \frac{C_{input}\,\frac{dV_{in}\,(t)}{dt} + I_{load}\,(t) - I_{PDN}\,(t)}{C_{load} + C_{input}} \tag{63}$$

Here I_{load} and I_{PDN} are the currents flowing through the transistors M_{load} and M_{PDN} respectively. It is important to note that the currents I_{load} and I_{PDN} are functions of the gate-to-source and drain-to-source voltages, which severely complicates the solution of the equation.

In the analytical solution, the miller effect is applied to the input capacitance, including it in the load capacitance. This should be a valid approximation for derivation of the device metrics, but it eliminates potential under- and overshoots of the output signal.

$$\frac{dV_{out}\,(t)}{dt} = \frac{I_{load}\,(t) - I_{PDN}\,(t)}{C_{load}} \tag{64}$$

To provide a complete analysis of the device metrics of the inverter, both numerical and analytical solutions are studied.

8.1.2 Capacitance merging

Each of the two transistors has four different capacitances, adding up to a total of eight, and they can be depicted in Figure 31(b). These can be combined into one fan-in and one fan-out capacitance. The fan-out capacitance is the contribution to the total load capacitance from the next step of a circuit tree. It scales linearly with the number of parallel inverters in the following stage.

It is easily realized that both the source to drain capacitances can be grouped together. Any voltage variation on the output node will affect both the capacitances in the same way, effectively meaning that they are in parallel. Additionally, both the gate capacitances can be added to their respective gate-to-source parasitic capacitance. This is because, in an NMOS, the channel potential is the same as the source potential.

The parasitic $C_{GS,load}$ is coupled between the same nodes as $C_{SD,load}$, meaning that it can be folded and lumped together with the others. In order to fold the input capacitance $C_{GD,PDN}$, the Miller effect has to be considered, effectively meaning that it is multiplied by a factor of 2. This is due to a voltage variation on the input node, resulting in a double variation at the output node,

in turn affecting the resulting capacitance. Finally, it is clear that $C_{GD,load}$ can be neglected, since it obviously is short-circuited.

All of the capacitances have now been lumped together into two capacitances as given by (65) and (66):

$$C_{fanout} = C_{G,PDN} + C_{GS,PDN} \tag{65}$$

$$C_{fanin} = C_{SD,PDN} + C_{SD,load} + C_{G,load} + C_{GS,load} + 2 \cdot C_{GD,PDN} \tag{66}$$

8.1.3 Magnitude of parasitic components

In Figure 32, the relative contribution of the different transistor parasitics are studied. This is done for a standard transistor unit cell, utilizing the test parameters presented in Table 2,

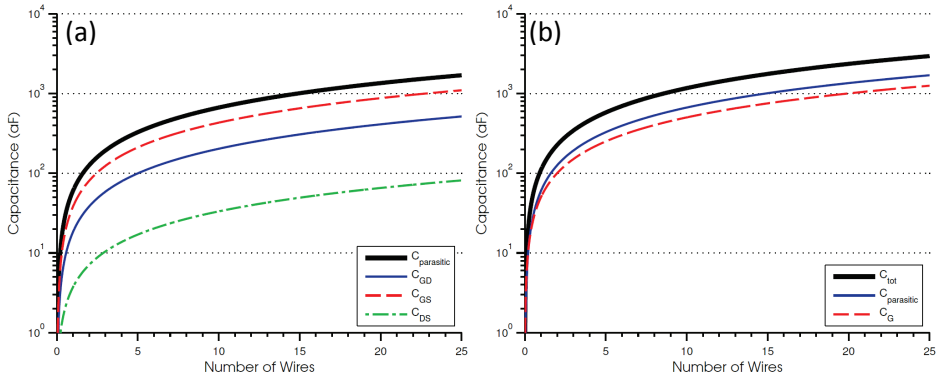


Figure 32 – (a) Comparison of parasitic capacitances between different electrodes. (b) Comparison between intrinsic and parasitic capacitances.

Figure 32(a) shows the magnitude for the three capacitance components as a function of the number of nanowires in the transistor. It is observed that the drain-to-source capacitance is significantly lower than the other two and it should be safe to neglect.

The total parasitic components are also compared to the intrinsic gate capacitance, as shown in Figure 32(b). It is clearly shown that the components are of comparable magnitude, indicating that the parasitic capacitance is very large.

8.1.4 Load capacitance

The load capacitances, affecting the performance of the inverter, is the sum of C_{fanout} and C_{fanin}. Both of these capacitances are multiplied with an integer factor, based on the number of fan-in and fan-out devices that are connected to the inverter.

The load capacitance for the inverter is shown in Figure 33(a), as a function of the number of nanowires in the load and PDN transistor, with a fan-out of one. As expected, the load capacitance increases with the number of wires in both transistors. It is therefore expected that, in order to increase the performance of the inverter, the number of nanowires should be minimized.

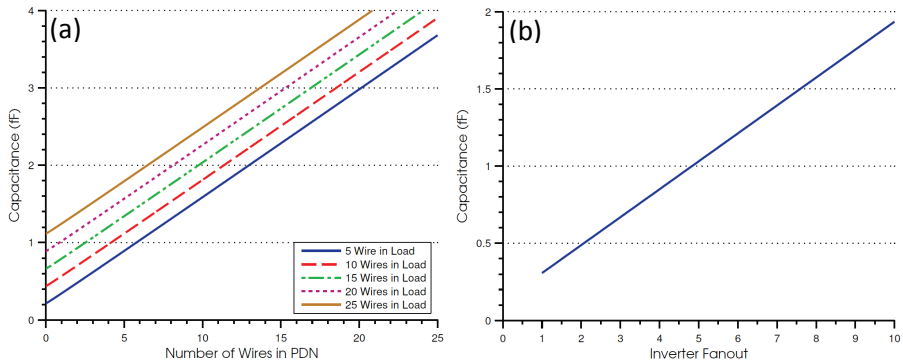


Figure 33 – (a) The dependence of the load capacitance on the number of nanowires in the load and PDN transistors dependence and (b) dependence of the load capacitance on the inverter fan-out.

8.1.5 Fan-out dependence

The gate-to-source capacitance does not directly influence the performance of a single inverter. However, an inverter is usually connected to other inverters of logic gates. The fan-out is a measure on how many inverters is connected to the output of the studied inverter.

As the gate-to-source capacitance of the fan-out inverters is connected directly to the previous output node, it is easily realized that an increasing number of fan-out inverters increases the fan-out capacitance. The fan out dependence is shown as a plot in Figure 33(b), where the load capacitance is plotted against the number of fan-out inverters. It is noted that the load capacitance increases greatly with the fan-out and quickly dominates the fan-in capacitance.

8.2 Numerical solution

By using a module of the NWFET modelling suite called the circuit designer, it is possible to analyse the numerical solution of common benchmarking devices, e.g. inverters and ring oscillators. This is done by numerically solving the complex form of the differential equation (63), by using the stiff ordinary differential equation solver *ode15s* in MatLab. From this solution, performance metrics are extracted, which may be used for benchmarking against competing technologies. A screenshot of the circuit simulation and optimization module in the software is shown in Figure 34.

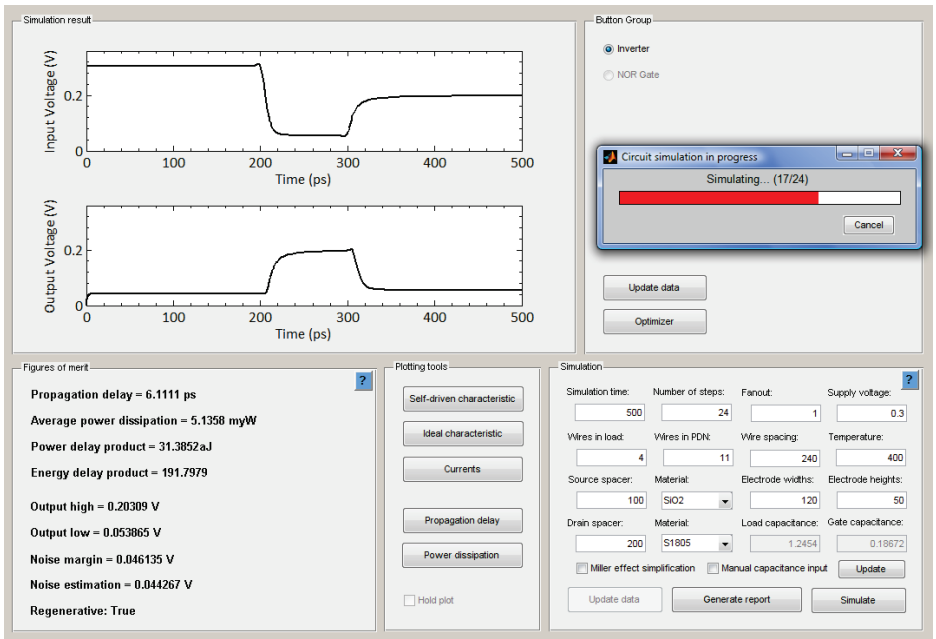


Figure 34 – A screenshot depicting the graphical user interface of the circuit designer.

8.3 Analytical solution

Analytical solutions are derived by solving the simple form of the differential equation (64), using conventional mathematics assisted by Maple. The analytical solution provides the user with a relatively simple set of equations, which may be of educational help in understanding the behaviour of the circuit. Numerical methods on the other hand, are extremely hard, if not impossible, to use for manual calculations. Another advantage of the analytical expressions is that they require much less computer power. This

makes them useful for large-scale optimizations, as is further investigated in section 10.1.

8.4 Delay times

By solving the differential equation (63), it is possible to derive the delay times analytically. Calculations of the delay times are divided into one solution for the low-to-high transition and one for the high-to-low transition.

8.4.1 Low-to-high

Firstly, the low-to-high transition, $\tau_{0\to1}$, is considered. In this case the output of the inverter is initially at V_{OL}, here simply assumed to be 0 V. The input of the inverter in this transition is changing from V_{DD} to 0 V. From these initial conditions, it can be concluded that the load transistor should be in the saturation region:

$$I_{load}(t) = N_{load}\left(V_{DD} - V_{out}(t)\right) \cdot \sigma_S\left(V_{DD} - V_{out}(t)\right) \tag{67}$$

Likewise, the PDN transistor is assumed to be in the linear region:

$$I_{PDN}(t) = N_{PDN} V_{out}(t) \sigma_L\left(V_{DD} - V_{in}(t), V_{out}(t)\right) \tag{68}$$

The inverter input signal, $V_{in}(t)$, which is at the gate of the PDN transistor, is here simply defined as a linear time-dependent function with the fall-time coefficient, τ.

$$V_{in}(t) = \tau t \tag{69}$$

This allows for the solution of the differential equation, with the initial condition $V_{out}(0) = 0$. As an exact solution of this differential equation is extremely difficult, a series expansion of order 2 is used instead. This results in the following expression for the output voltage:

$$V_{out}(t) = N_{load} V_{DD} \frac{t}{C_{load}\,\sigma} + O(t) \tag{70}$$

Note that a σ, dependent on the supply voltage and threshold voltage is falling out of this solution, which will be the case for many following expressions. This σ is closely related to the Jansson-Berg substitutions described in section 4.3.5.

$$\left[\sigma = \left(\frac{2V_{DD}}{kV_{sat}^2\,(V_{DD} - V_T)} + R_s\right)^{-1}\right] \tag{71}$$

Ignoring the higher order terms, $O(t)$, and assuming that 50% of the transition time is at $V_{DD}/2$, the following expression for the low-to-high delay time is derived:

$$V_{out}\,(\tau_{0\to1}) = \frac{V_{DD}}{2} \Rightarrow \boxed{\tau_{0\to1} = \frac{C_{load}}{2N_{load}\,\sigma}} \tag{72}$$

8.4.2 High-to-low

Secondly, the high-to-low transition is considered as well. Now the inverter has an initial output of V_{OH}, assumed here to simply be V_{DD}. In this case, the input is changing from 0 V to V_{DD}. From this information, it can be concluded that the load transistor should be in the linear region:

$$I_{load}\,(t) = N_{load}\left(V_{DD} - V_{out}\,(t)\right) \cdot \sigma_L\left(V_{DD} - V_{out}\,(t)\right) \tag{73}$$

The PDN transistor, on the other hand, is assumed to be in the saturation region:

$$I_{PDN}\,(t) = N_{PDN} V_{out}\,(t)\sigma_S\left(t\tau, V_{out}\,(t)\right) \tag{74}$$

Solution of the differential equation is done with the initial condition that V_{out} should be V_{DD} when the input signal (in this case simply $\tau_{0\to1}$) is at V_T:

$$\left\{V_{out}\left(\frac{V_T}{\tau_{0\to1}}\right) = V_{DD}\right\} \tag{75}$$

As this differential equation is even more complex, a series order 4 is needed:

$$V_{out}\,(t) = V_{DD} - N_{PDN}\frac{k}{C_{loa\,d}\tau_{0\to1}^2}(t - V_T\tau_{0\to1})^3 + O(t) \tag{76}$$

Ignoring the higher order terms and assuming that the 50% time of the transition is at $V_{DD}/2$, the following expression for the high-to-low delay time is derived:

$$V_{out}\,(\tau_{1\to0}) = \frac{V_{DD}}{2} \Rightarrow \tau_{1\to0} = V_T\tau_{0\to1} + \sqrt[3]{\frac{3V_{DD}\,C_{load}\,\tau_{0\to1}^2}{4kN_{PDN}}} \tag{77}$$

Insertion of the low-to-high delay time into this expression results in the following equation:

$$\boxed{\tau_{1\to0} = C_{load}\left(\frac{V_T}{2N_{load}\,\sigma} + \sqrt[3]{\frac{3V_{DD}}{4kN_{load}^2\,N_{PDN}\,\sigma^2}}\right)} \tag{78}$$

8.4.3 Propagation delay

By using the equations above, a complete expression for the propagation delay is derived:

$$\tau_p = \frac{\tau_{0\to1}+\tau_{1\to0}}{2} \Rightarrow \boxed{\tau_p = \frac{C_{load}}{2}\left(\frac{V_T+1}{2N_{load}\,\sigma} + \sqrt[3]{\frac{3V_{DD}}{4kN_{load}^2\,N_{PDN}\,\sigma^2}}\right)} \tag{79}$$

As the expression for the high-to-low transition is somewhat complicated, with a dependence on the low-to-high transition as well as including a cubic root, it is desirable to have an easier expression. In this case, the propagation delay can be estimated by assuming that the high-low transition is shorter than the low-high transition, resulting in the following expression:

$$\tau_p = \frac{\tau_{0\to1}}{\alpha} \Rightarrow \boxed{\tau_p = \frac{C_{load}}{2N_{load}\,\alpha\sigma}} \tag{80}$$

As numerical simulations have shown, the high-to low transition is generally shorter than the low-to-high transition, resulting in $1 \leq \alpha \leq 2$. In the following analytical expressions, both the transitions are assumed to be of the same length, i.e. $\alpha = 1$.

8.5 Average power dissipation

The power consumption in a ratioed logic design is dominated by the saturation current through the load transistor, while the inverter input is V_{DD}. Assuming an ideal inverter, with an output-low voltage of 0 V, the power dissipation is simply the saturation current multiplied by the supply voltage. In this case, the PDN transistor has no power dissipation, as the voltage over the transistor is 0 V.

When the input to the transistor is 0 V, the power consumption is negligible in comparison to when the input is V_{DD}. Assuming the inverter spends half of its time in state 1 and the rest in state 0, the resulting average power dissipation

is equal to half the power consumption for a high input. This results in the following expression:

$$P_A = \frac{(N_{load}\, V_{DD}^2\, \sigma + 0)}{2} \Rightarrow \boxed{P_A = \frac{N_{load}\, V_{DD}^2\, \sigma}{2}} \tag{81}$$

A more correct expression, which is valid and accurate over greater parameter ranges, can be derived by considering the static output levels:

$$\boxed{\begin{aligned} P_A = \frac{1}{2}\big(&N_{load}\,(V_{OH} - V_{OL})^2 \sigma_S(V_{OH} - V_{OL}) \\ &+ N_{PDN}\,(V_{OL})^2 \sigma_L(V_{OH}, V_{OL})\big) \end{aligned}} \tag{82}$$

Expressions for V_{OL} and V_{OH} is derived in section 8.8.

8.6 PDP and EDP

By using the simple analytical expressions, it is now possible to derive the key figures of merit, the power-delay product, PDP and energy-delay product, EDP:

$$\mathrm{PDP} = \tau_p P_A = \frac{C_{load}}{2N_{load}\,\sigma} \cdot \frac{N_{load}\, V_{DD}^2\, \sigma}{2} \Rightarrow \boxed{\mathrm{PDP} = \frac{C_{load}\, V_{DD}^2}{4}} \tag{83}$$

$$\mathrm{EDP} = \tau_p^2 P_A = \frac{C_{load}}{2N_{load}\,\sigma} \cdot \frac{C_{load}\, V_{DD}^2}{4} \Rightarrow \boxed{\mathrm{EDP} = \frac{C_{load}^2\, V_{DD}^2\, N_{load}}{8\sigma}} \tag{84}$$

8.7 Influence of series resistance

The series resistance is severely increasing the complexity of the expressions and it would be interesting to investigate in which cases it can be disregarded. By identifying terms in the expressions above, it is concluded that it can only be disregarded if it is very small in comparison to the equivalent resistance of the transistor:

$$\boxed{R_S \ll \frac{2V_{DD}}{k V_{sat}^2\,(V_{DD})}} \tag{85}$$

This relation, with the parameters of the standard test transistor, is plotted against the supply voltage in Figure 35. As the series resistance in most cases is around 10 kΩ, it is concluded that the series resistance can only be safely disregarded for very low supply voltages.

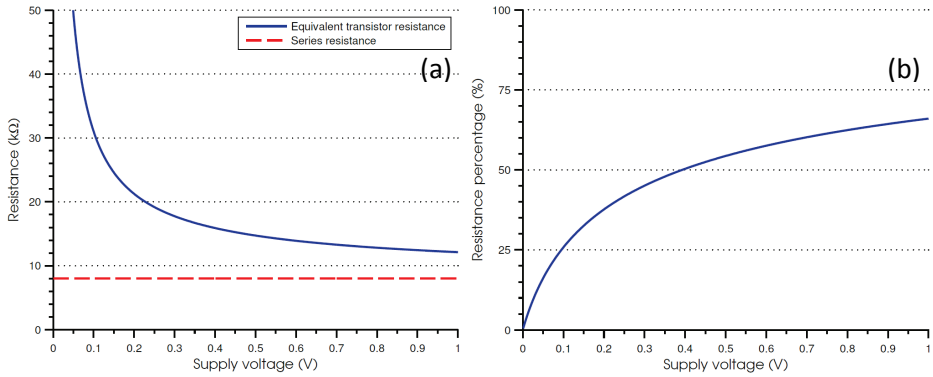


Figure 35 – The contribution of the series resistance to the overall resistance.

8.8 Static voltage levels

To be able to study if an inverter is regenerative or not, the static voltage levels have to be calculated. As a starting point, the nodal analysis from section 8.1.1 is used. Static levels are nondependent of time and thus the time derivative is equal to zero. As a result, the currents may be assumed equal, i.e. $I_{load} = I_{PDN}$.

8.8.1 Output-high

When studying the output-high value, the input is set below the threshold voltage of the PDN transistor, causing it to operate in the cut-off region ($I_{PDN} = 0$). The load transistor should have a far less voltage drop over it than the PDN transistor. Since the drain-to-source voltage is equal to the gate-to-source voltage for the load transistor, it is realized that this transistor as well must be in the cut-off region ($I_{load} = 0$).

This only applies if the output voltage approaches the output-high voltage from a lower voltage. The voltage at which the load transistor enters the cut-off region is at $V_{DD} - V_T$, where the condition is met that both currents are zero and the voltage cannot increase anymore. It is thus concluded that:

$$\boxed{V_{OH} = V_{DD} - V_T} \tag{86}$$

8.8.2 Output-low

To find the output-low voltage level, the input voltage is set to $V_{DD} - V_T$. As the output level should be low, the load transistor can be in either the linear

or the saturation region, depending on e.g. the nanowire ratio. After using the Jansson-Berg substitutions, the resulting equation can be formulated as:

$$N_{load}\,(V_{DD}-V_{OL})\sigma_L(V_{DD}-V_{OL}) = \\ = N_{PDN}V_{OL}\sigma_{L,S}(V_{DD}-V_T,V_{OL}) \tag{87}$$

This equation is invalid when the output voltage is higher than $V_{DD}-V_T$, as the load transistor approaches the cut-off region. These equations are very complicated to solve with an exact solution. One way to go around this problem is to search for series solutions. However, the series solutions become very complex, as a high order expansion is necessary.

Simplifications

In order to not being limited to series solutions, further simplification of the original equation is needed. For the linear region, it is assumed that the output voltage is sufficiently low to be neglected in comparison to the supply voltage. In that case, the equation is simplified to a quadratic equation:

$$N_{load}\,V_{DD}\sigma_L(V_{DD}-V_{OL}) = N_{PDN}V_{OL}(t)\sigma_L(V_{DD}-V_T,V_{OL}) \tag{88}$$

Advanced expression

For the advanced expression, the saturation region is maintained without simplifications, while the linear region is simplified as specified above. This means that both these expressions are of the second grade, excluding the output-low voltage dependence on the saturation voltage of the load transistor. The solution of these equations are complicated but are reproduced in the appendix.

As the dependence of the output-low voltage on the load saturation voltage has not yet been resolved, it may be compensated for by an iterative method, as proposed by e.g. [**21**]. This is done by calculating the quadratic output-low and identifying the correct root, followed by an insertion in the saturation voltage and a recalculation of the output-low voltage. However, this method does not always converge at extreme values or at the boundary between the regions. This means that an alternative method must be developed.

The solution is to plot the solution of the output-low voltage expression for the two regions versus the output-low voltage used in the saturation voltage

calculation, as illustrated in Figure 36. It is now just a matter of identifying where the output V_{OL} is equal to the input V_{OL}. This should correspond to the value approached by the iterative method. In order to identify these points, a help line consisting of the V_{OL} in the saturation voltage calculation is used. This reduces the problem to calculate, numerically or graphically, the intersection between the lines.

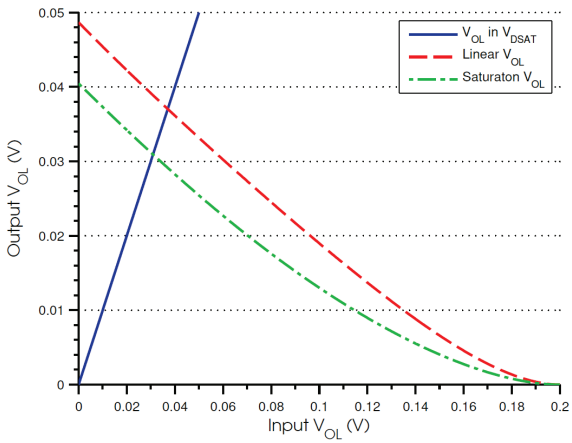


Figure 36 – The output-low voltage is calculated by identifying the intersection between the blue curve and the curves corresponding to the linear and saturation region of the PDN transistor.

The solution results in two intersections, one for the linear region and one for the saturation region of the PDN transistor. Which of these values that are correct, is determined by identifying the regions of the intersection points in the PDN transistor. This method is tested against the numerical calculations and is determined to be more or less exact.

Simple expression

The above method is well suited to use in the developed software, but are quite unwieldy for manual calculations. In this case, further simplifications are necessary and it is assumed that the PDN transistor constantly is in the linear region. Additionally, it is assumed that V_{OL} can be neglected compared to $2V_{sat}$ in $\sigma_{L,S}$. By doing this, a linear equation is left:

$$N_{load}\, V_{DD}\sigma = N_{PDN} V_{OL}(t)\sigma_L(V_{DD} - V_T, 0) \qquad (89)$$

Isolation of V_{OL} results in the following simple expression for the output-low voltage:

$$V_{OL} = \frac{N_{load}\, V_{DD}}{N_{PDN}} \sigma \left(\frac{1}{k V_{Sat}\, (V_{DD} - V_T)} + R_s \right)^{-1} \tag{90}$$

This solution corresponds much worse to the numerical calculations, but it should be noted that it imposes more severe restrictions than the advanced expression. This means that all parameter combinations determined to be regenerative by this formula also will be regenerative in the advanced expression.

8.9 Summary of circuit models

A summary of the different derived models, with decreasing complexity is presented in Figure 37.

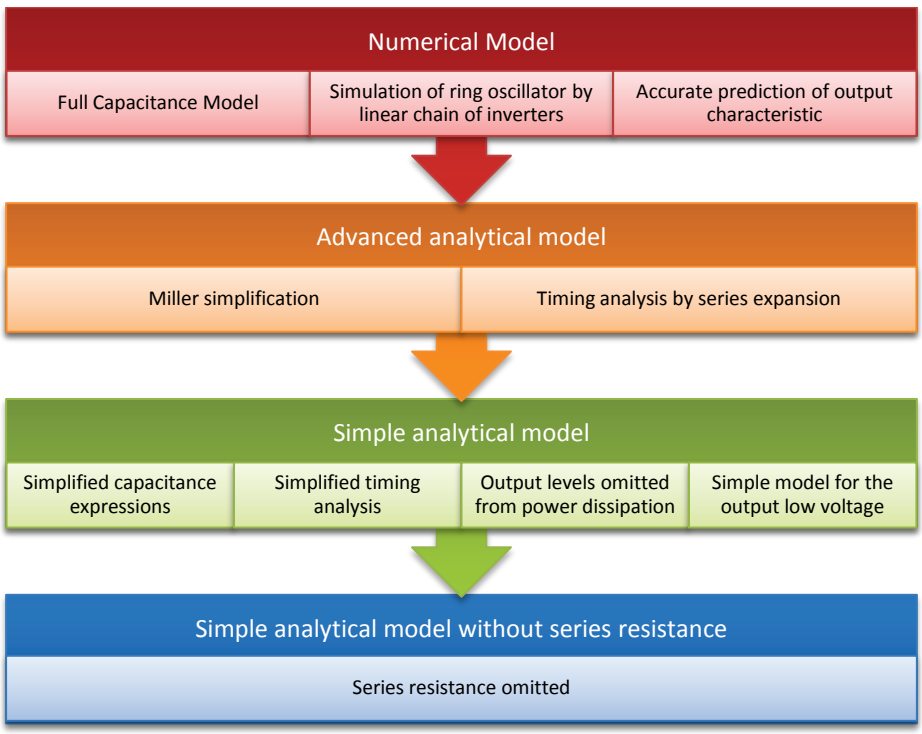


Figure 37 – Summary of the different circuit models used for simulations and calculations, with decreasing accuracy and complexity.

The numerical solution is the most exact, implementing correct propagation delay and power dissipation simulations as well as the full capacitance model. The advanced analytical model replaces the numerical methods for the figures of merit, but keeps the full capacitance model. The simple analytical model, customized for hand calculations, uses simplified expressions for the figures of merit, as well as using a reduced capacitance model. The simplest model also disregards series resistance.

The numerical model is used in the circuit simulation software, while the advanced analytical model is used in the optimization module. The two simple models are designed mainly for manual calculation.

Part IX CIRCUIT SIMULATION RESULTS

9.1 Output characteristic

By numerical calculations in the software, the output characteristic of the inverter is derived. The transistor parameters used during the simulations are the test transistor parameters presented in Table 3. For the simulation, the Jansson-Berg model is used with the number of nanowires in the load and PDN transistors set to three and seven respectively.

9.1.1 Ideal

In the case of an ideal input signal, the resulting output characteristic is shown in Figure 38.

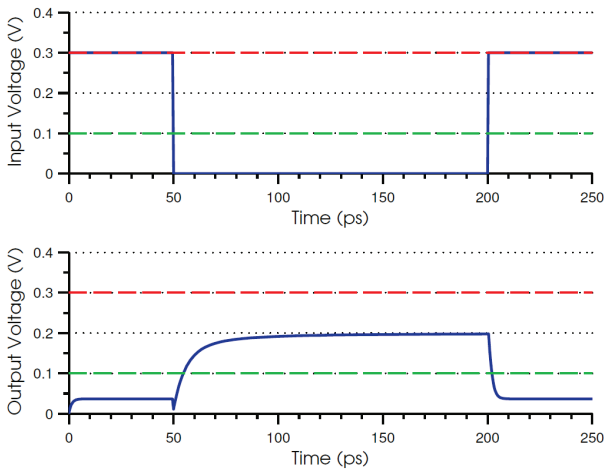


Figure 38 – Ideal output characteristic of the inverter.

It is clearly observed that the output characteristic has a non-ideal form as predicted by the use of ratioed logic. The output-high is approximately at $V_{DD} - V_T$ and the output-low is non-zero. As predicted, the high-to-low transition delay is much faster than the low-to-high transition delay. This means that it is a valid, although crude approximation to assume that the propagation delay is half that of the low-to high transition. This principle is applied in the simple analytical model.

It is also confirmed that the low-to-high transition is approximately linear up to the half transition, verifying that the analytical approximation of a linear transition is valid. Additionally some over- and undershoots in the transients

are observed. These are due to the input capacitance and are lost when using the Miller-effect simplification in the analytical solution.

9.1.2 Self-driven

However, the use of an ideal input signal is far from the case in a typical circuit. This is due to the output from one gate, being the input to the next. The conventional method to illustrate this fact is the use of a ring oscillator. As this is very difficult to realize with analytical expressions, the ring oscillator is approximated as a linear chain of inverters. In order to study the self-driven characteristic, the input signal is taken as the output of the 19th stage in the chain and the output of the 20th stage is used as the self-driven output characteristic. These input and output signals are shown in Figure 39.

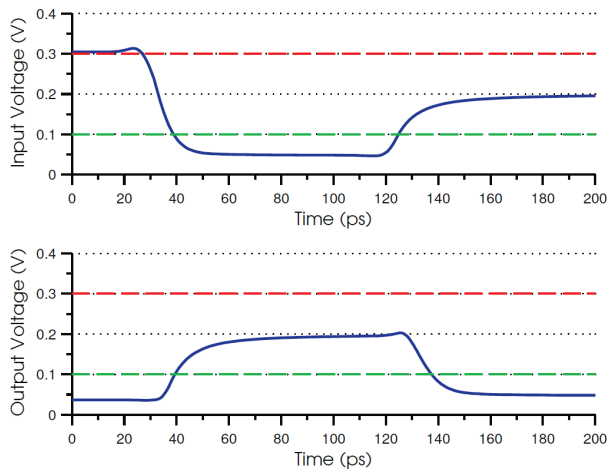


Figure 39 – Self-driven output characteristic, simulated as the transition between the 19th and 20th stage in a linear chain of inverters.

9.1.3 Propagation delay

In order to determine the propagation delay in the numerical solution, the linear inverter chain is used once again. This time the output of the stage in the middle of the chain is compared to the output at the end of the chain. By comparing the displacement of the input pulse, the propagation delay is calculated as the delay divided by the number of inverters stages between the studied stages.

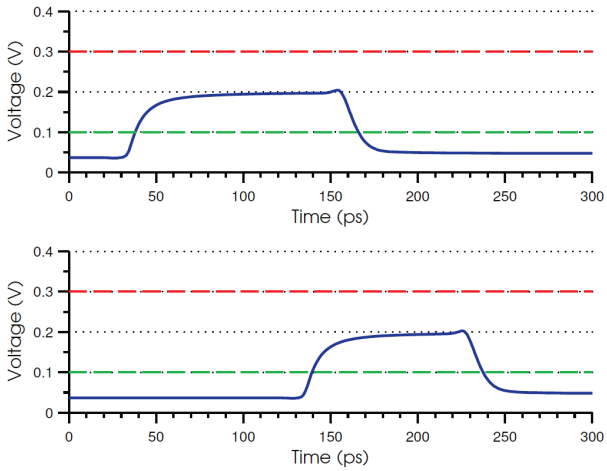


Figure 40 – Propagation delay, illustrated as the difference of characteristics at the (a) 10th and (b) 20th stages of the linear chain of inverters

9.2 Currents

To determine the power dissipation of the inverter, the currents through the individual transistors need to be considered. The currents through the load and PDN transistors respectively, are presented in Figure 41(a). It is clearly observed that the current is close to zero when the input is low and a large static current is flowing through the inverter when the input is high.

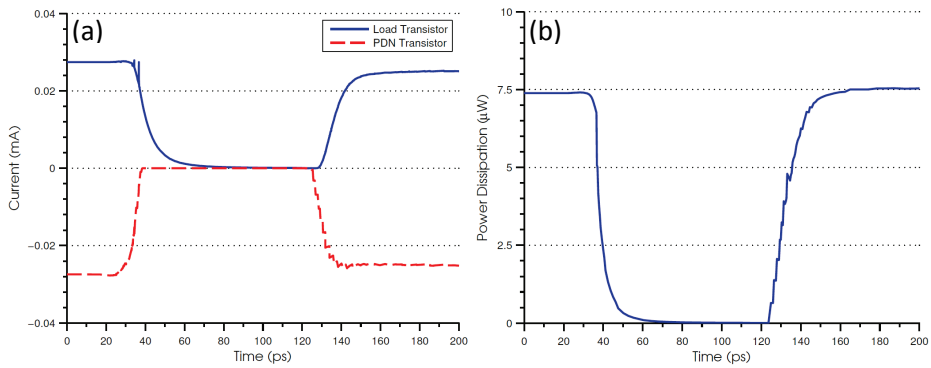


Figure 41 – (a) Currents flowing through the load and PDN transistors respectively. (b) Power dissipation of the inverter.

9.3 Power dissipation

Now that both the current and voltage is known, it is possible to derive the power dissipation of the inverter, as shown in Figure 41(b). As predicted by

the current levels, the only power dissipation that need to be considered is caused by the short-circuit when both transistors are on, i.e. for a high input level. There is some power dissipation at the transitions, but it is negligible in comparison to the static power consumption, with reservation for very fast switching speeds.

9.4 Voltage levels

By constructing 3D-plots, it is possible to investigate the influence of supply voltage and threshold voltage on the inverter behaviour at the same time. In the following plots, blue corresponds to a generally favourable value, while red denotes the worst possible combination. In the plots, there can be seen a displaced borderline between the PDN linear and saturation regions. This is due to the small approximation done in the expression for the linear region. However, this small error should not significantly degrade the results.

The light green region is covering the non-regenerative values due to output-low, while the light blue region covers the values that are not regenerative, due to the output-high restriction. The light red area is the recommended noise margin and values inside this area may not be safe to use without error correction methods.

9.4.1 Propagation delay

In Figure 42, the propagation delay dependence on the voltage levels is shown. It can be seen that the propagation delay is short along the left boundary of the triangle formation, with the fastest speed at the upper left corner. However, the value at the bottom corner is not much worse and will be shown to have much lower power dissipation due to the lower voltage levels.

9.4.2 Average power dissipation

The power dissipation has a somewhat inverted form in comparison to the propagation delay, as shown in Figure 43. However, they share a common acceptable point, which is at the power dissipation minima at the bottom corner. This will be clearly reflected in the PDP and EDP plots.

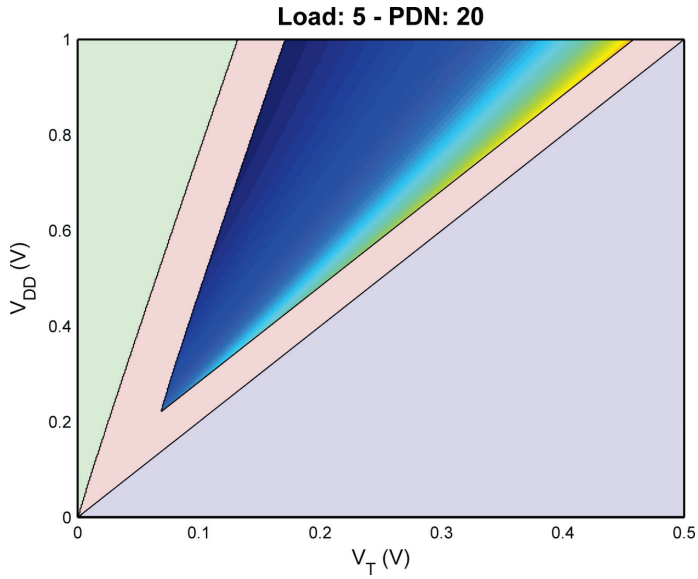


Figure 42 – The propagation delay for an inverter with 5 nanowires in the load transistor and 20 in the PDN transistor. The colour scale is logarithmic from blue to red corresponding to short respectively long propagation delays.

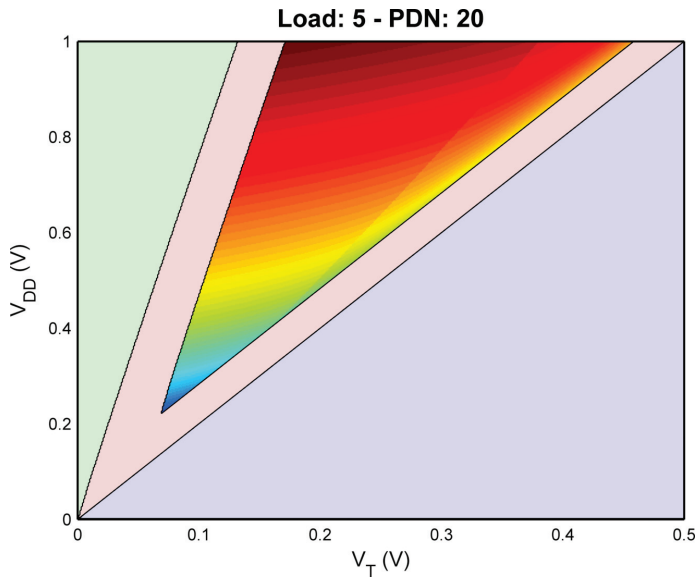


Figure 43 – The average power dissipation for an inverter with 5 nanowires in the load transistor and 20 in the PDN transistor. The colour scale is logarithmic from blue to red corresponding to low respectively high power dissipation.

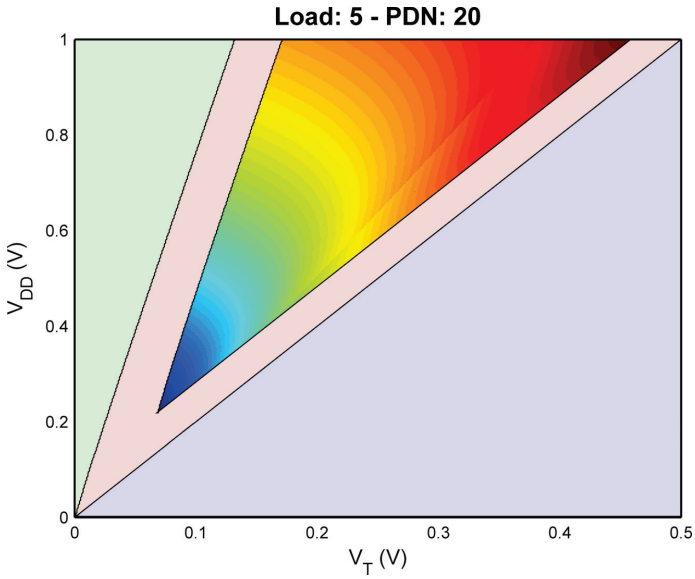


Figure 44 – Power-delay product for an inverter with 5 nanowires in the load transistor and 20 in the PDN transistor. The colour scale is logarithmic from blue to red corresponding to low respectively high power-delay product.

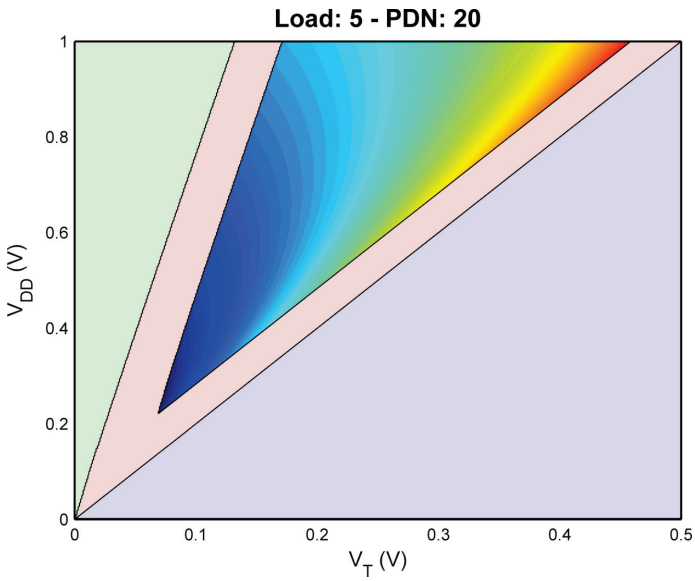


Figure 45 – Energy-delay product for an inverter with 5 nanowires in the load transistor and 20 in the PDN transistor. The colour scale is logarithmic from blue to red corresponding to low respectively high energy-delay product.

9.4.3 PDP and EDP

In Figure 44, it is seen that the PDP minimum is located at the bottom corner of the triangle formation. The same case is true for EDP, as shown in Figure 45. It is apparent that the optimization of the NWFET inverter design is simply to use the lowest possible voltage levels. However, this point is dependent on the regeneracy limitations, as well as noise. In the optimization section, it is shown that by varying e.g. the number of nanowires in the respective transistors, the location of this point may be adjusted.

9.5 Wire dependence

By automated calculation of device metrics in the software, it is possible to quantify the performance dependence on the number of nanowires in the respective transistors. In the following figures, values are omitted due to non-regeneracy.

9.5.1 Delay times

The delay times are shown in Figure 46. Again, it is confirmed that the high-to-low transition is much faster than the low-to-high transition and that this relation is maintained for different inverter compositions. Both of the delays increase with the number of nanowire in the PDN. However, the high-to-low transition delay is expected to decrease with the number of PDN nanowires, because of the lower resistance in the PDN. The reason for the observed behaviour is that the parasitic capacitance is increasing with the number of nanowires in the transistor, and thereby counteracting the positive effect of more nanowires.

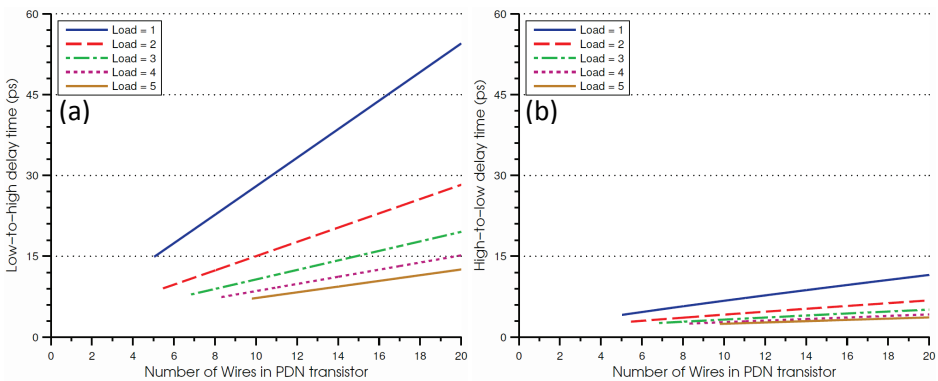


Figure 46 – (a) Low-to high delay time and (b) High-to-low delay time.

The combined propagation delay is shown in Figure 47(a). It is clearly observed that the propagation delay increases linearly with an increased number of wires in the PDN transistor. With increasing number of nanowires in the load transistor, the propagation delay instead decreases.

9.5.2 Average power dissipation

The dependence of the average power dissipation on the number of nanowires is shown in Figure 47(b). As is the case for the delay times, the power dissipation is at its minimum when the number of nanowires in the PDN transistor is as low as possible. However, for the number of nanowires in the load transistor, the behaviour is reversed. In this case, a low number of nanowires are desired in the load transistor as well.

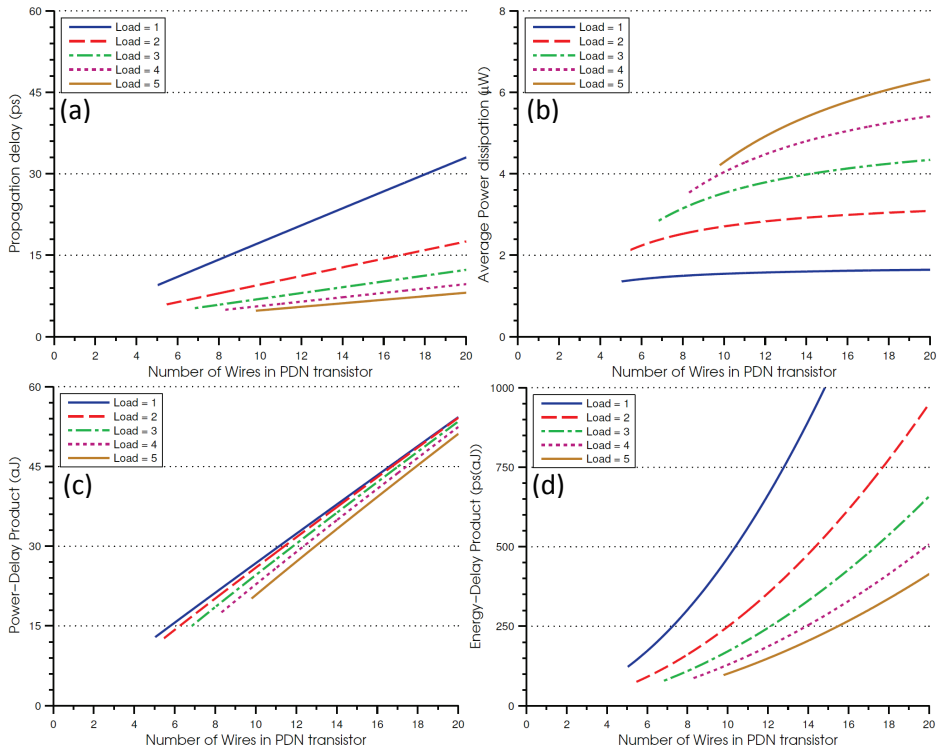


Figure 47 – (a) Propagation delay, (b) Average power dissipation, (c) Power-delay product and (d) Energy-delay product.

9.5.3 PDP and EDP

By using the derived behaviour of the propagation delay and power dissipation, it is possible to plot the power-delay and energy-delay products, as shown in Figure 47. In this figure, the restriction of the regeneracy is clearly illustrated. If regeneracy is not considered, the curves would extend further and the minima change. It is also seen that the minima are not the same for the PDP and EDP curves and that different inverter compositions may be used to optimize for speed and low power consumption respectively.

Part X OPTIMIZATION METHOD

10.1 Method

The optimization of the inverter design is done primarily based on the energy-delay product, but also power-delay product as well as pure propagation delay and power dissipation optimization is done. The PDP design might be most suitable for low-power applications, while the EDP optimized design will be more suited for high-performance applications.

In order to automate the optimization procedure, the optimization module of the software suite is used. A screenshot from this program is seen in Figure 48. When the optimization is completed, a report of the results may be exported. This report is designed to be used as a recipe in the manufacture of a new, optimized circuit. Additionally, the report contains a summary of the modelling and circuit simulation, as well as all calculated properties and figures of merit. An example of such a report is included in the appendix.

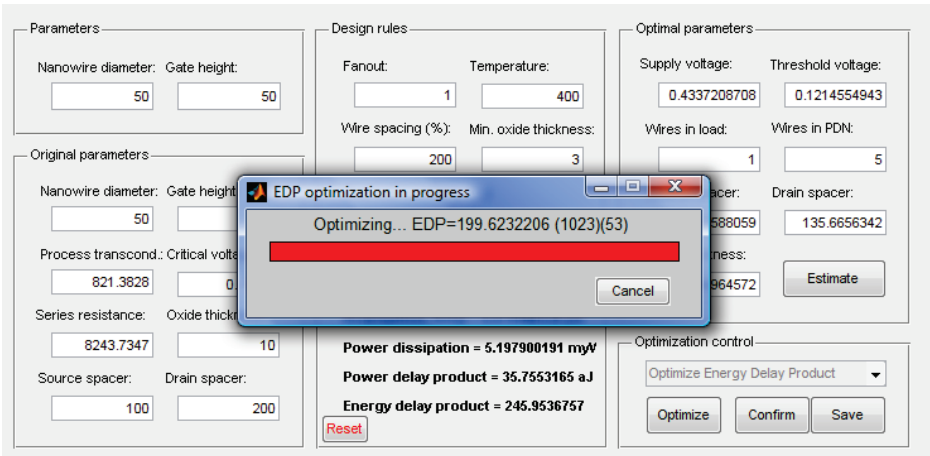


Figure 48 – A screenshot of the optimization module of the software suite.

By using an intelligent algorithm, similar to the one used for the modelling part, the parameters of the advanced analytical model are optimized. These values are then fed to the numerical solver for accuracy refinement. An absolute quality evaluation is problematic, as the optimum value is not known. Instead, a relative quality evaluation is used to allow for intelligent step size adjustments. The step sizes in this case are set through a complex mixture of derived hardcoded limits as well as closely monitoring the

incremental success percentage. Additionally, some variable sets are grouped together and optimized alternately.

As the complexity and risk for finding local minima, instead of the global minima, is rising exponentially with additional number of variables, primarily seven parameters are chosen for the optimization. The others are estimated by using minimum dimensions defined by design rules.

10.2 Optimization variables

The criteria for a variable to be included in the optimization, is that it is of significant interest and the optimum value is not easily derived. In this case, seven different variables are chosen.

10.2.1 Voltage levels

Of immense importance for both propagation delay and power dissipation is the supply voltage, which supports the obvious inclusion of this variable into the optimization variables. It is also important to find the optimum threshold voltage to complement the supply voltage.

10.2.2 Nanowire quantity

As it has been shown in earlier sections, the nanowire quantities in the load and PDN transistors seriously affects the circuit performance and power consumption. It is therefore a simple choice to include them among the optimization variables.

10.2.3 Spacer layers

When the device is scaled, it is important that the spacer layers and gate oxide thickness scale as well. However, this scaling relation is not linear, even as a crude approximation. This motivates the added complexity of including these variables in the optimization process. It will later be shown that the thickness of the spacer layers greatly influences the balance between the series resistance and the parasitic capacitances.

10.3 Design rules

The boundary conditions of the optimizer algorithm are set by a defined set of design rules. These rules are somewhat arbitrary, as the processing limitations are not known, but they should at least provide a plausible starting point. The

design rules are easy to change in the optimization module of the accompanying software.

10.3.1 Fan-out

Fan-out is set to one in order to get appropriate benchmarking data, but should be set higher, to be able to design arbitrary logic circuits. However, this can be easily compensated for, as the behaviour of an increasing fan-out is highly predictable. Fan-out should not be significantly reflected in the optimization minima and only affect the size of the figures of merits.

10.3.2 Nanowire spacing

The minimum nanowire spacing, i.e. the distance between the outer limits of two neighbouring nanowires, is assumed to be four times the nanowire diameter.

10.3.3 Electrode configuration

In the optimum configuration, the electrodes are assumed to be in a perpendicularly networked structure. The widths of the electrodes are set to double the wire diameter, including dielectric, while the electrode height is set to the gate height.

10.3.4 Minimum gate dielectric thickness

In the optimization, it is assumed that the leakage current through the gate oxide can be neglected. This assumption is not necessarily true, as the devices are scaled towards very small dimensions. Preliminary work at the department has shown that the magnitude of the leakage current starts to significantly degrading at an oxide thickness of three nanometres. This motivates the decision to set a restriction of the gate oxide thickness to be a minimum of 3 nm.

10.4 Parameter scaling

As the process scales it is important to make sure that the model parameters used, are valid in the new region. In order to compensate for this effect, the following methods are used to scale the different model parameters.

10.4.1 Series resistance

The resistivity of the nanowire is assumed to be evenly distributed in the nanowire, although this may not be the case if different levels of doping are present. As the original topographical parameters of the device are known, it is easy to recalculate the extracted series resistance to a resistivity. The series resistance shows a linear dependence on the spacer layer thicknesses, which means that it can be recalculated for the new dimensions. As the diameter of the nanowire scales down, so does the area of the series resistance, which means that it needs to be compensated accordingly. In this simple resistance model, there is no compensation for eventual changes in the contact resistances.

10.4.2 Process transconductance parameter

The process transconductance is divided by the initial gate capacitance, in order to get the mobility. In this optimization, it is assumed that the mobility remains constant with decreased transistor dimensions. As the design is scaled down, the gate capacitance is recalculated and multiplied with the mobility in order to get the scaled process transconductance parameter.

10.4.3 Critical voltage

As predicted by (2), the critical voltage needs to be recalculated by scaling with the new channel length. It is here assumed that the mobility, as well as the saturation velocity, remains constant with the downscaling.

10.5 Noise

As the component is assumed to work in large-scale implementations at room temperature (or slightly above), there are some noise restrictions. The restriction of the noise is determined by the voltage levels of the inverter and is formulated as:

$$\begin{cases} V_T > V_{OL} + V_N \\ V_T < V_{OH} - V_N \end{cases} \tag{91}$$

The noise, V_N, is varying with temperature, here set to 400 K, and should be lower than specified in the equation below:

$$V_N = \beta \sqrt{\frac{kT}{C_N}} \approx 0.1 \text{ V} \tag{92}$$

Here C_N is the effective transistor capacitance, combined from both intrinsic and parasitic elements:

$$C_N = NC_G + G_{GS} + C_{SD} + \frac{C_{SD}\,(NC_G + C_{GS})}{C_{GD}} \tag{93}$$

The factor β is slightly dependant on the implementation complexity, as well as the switching speeds [**22**]. The value of β is influencing the false switching frequency, v, in the following manner:

$$v(\beta) = \frac{N_T}{\sqrt{3}\tau_p} \mathrm{e}^{-\frac{\beta^2}{2}} \tag{94}$$

Here, N_T is the number of transistors in the desired application. Figure 49 shows the false switches as a function of β.

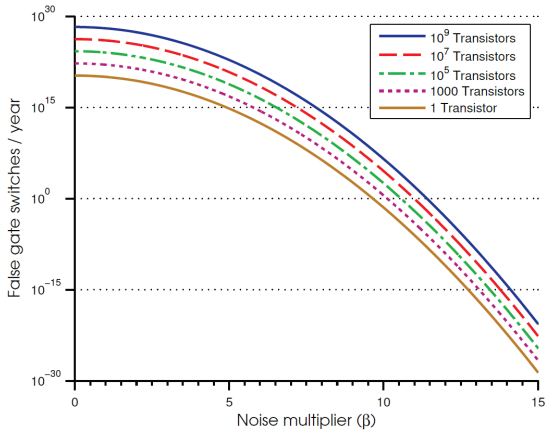


Figure 49 – False switching frequency for different transistor magnitudes.

It is set as a requirement that less than one false switch may occur per year, which corresponds to a value for β of about 12. This should provide a sufficiently safe noise margin even for operation without error correction. It can be seen that this value is largely independent on the implementation complexity and switching frequency.

10.5.1 Regeneracy limitations

Through the optimization simulations, it becomes apparent that it is desirable to use a supply voltage as low as possible in combination with an appropriate threshold voltage. However, a logic gate should be regenerative to be usable in arbitrary circuits.

The limitations imposed by the output-high and output-low voltage levels have been investigated in earlier chapters. Added to this is the noise influence, which further raises the allowed voltage levels. These limitations are illustrated in Figure 50 as different coloured regions, where the optimum voltage level is marked by a circle.

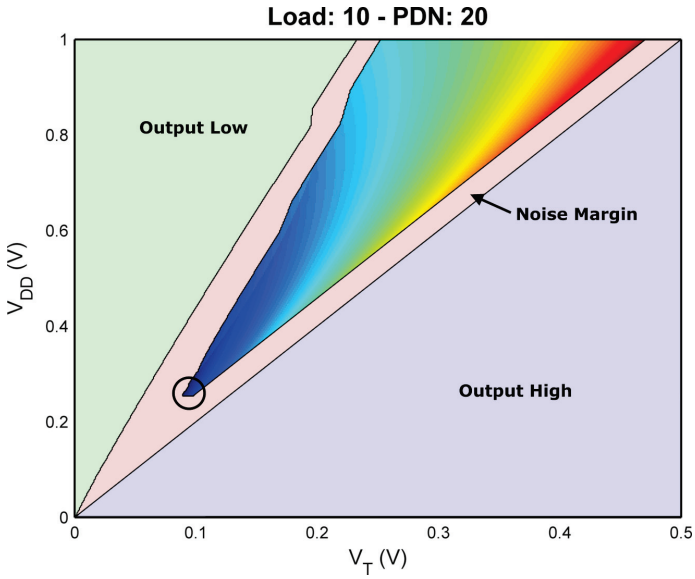


Figure 50 – The regenerative triangle showing regeneracy limitations due to output-low and output-high, as well as the recommended noise margin.

Note that although the noise region seems just like a small border region, it imposes a significant limitation on the voltage levels. The triangle formation shown in the figure is further on referred to as the regenerative triangle.

10.5.2 Noise reduction

In order to minimize the allowed voltage levels, it is apparent that the circuit noise must be reduced. The primary tool to achieve this is by increasing the number of wires in the transistors, thereby increasing the total gate capacitance. This increased capacitance increases the noise resistance of the circuit at the cost of reduced performance, as described by (92). As there generally are fewer wires in the load transistor, it is more affected by noise than the PDN transistor, which in turn limiting the noise resistance. One of the difficulties in the optimization is finding an acceptable balance between high noise resistance and a high circuit performance.

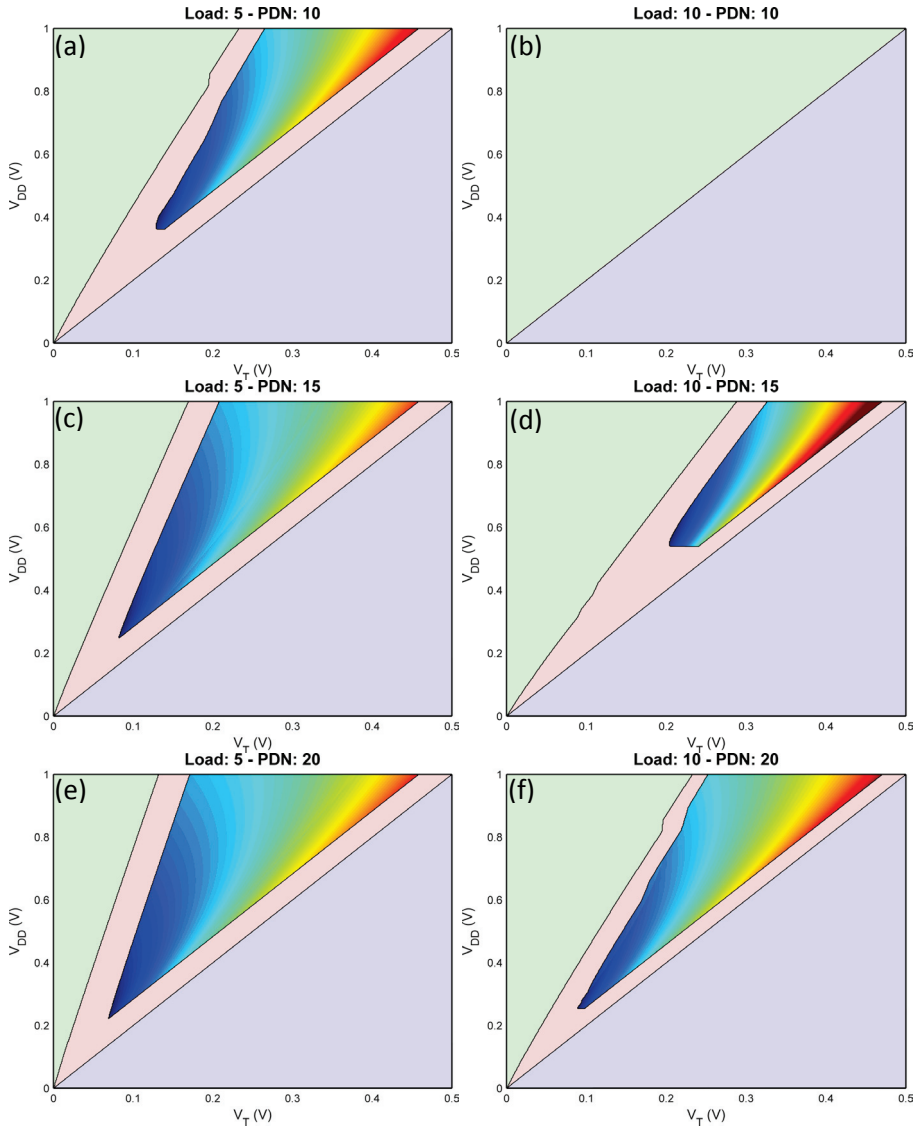


Figure 51 – Regeneracy triangle size with changing number of wires in the load and PDN transistors.

10.5.3 Wire ratios

By adjusting the ratio between the number of wires in the load and PDN transistor, the size of the regenerative triangle can be changed, as seen in Figure 51. Increasing the number of wires in the PDN transistor allows the triangle to expand further into the output-low territory, while an increase in

the number of wires in the load transistor counteracts these changes. It is also observed that increasing the nanowire quantity from 1:2 to 2:4, lowers the noise floor, as predicted in the previous section.

Part XI OPTIMIZATION RESULTS

By using the results from the optimization module, the behaviour of a number of interesting parameters, with increased downscaling, is derived. It should be noted that the construction of the following plots is a very computer intense task and in order to achieve the necessary power, a cluster of over 10 PCs were used over the duration of about four work days. It is estimated that around 300 million parameter sets has been tested, resulting in about 3000 different local minima.

The scaling is performed as a function of the nanowire diameter, with a step size of one nanometre, between 10 nm to 100 nm. As the genetic algorithm used during the analysis, is prone to get stuck at local optimization minima, care has to be taken to ensure that the global minimum is found. In order to distinguish between local and global minima in the optimization plots, the global minima are highlighted in respect to the other sets.

By some preliminary testing, it becomes apparent that the only meaningful optimization of the design is to optimize the energy-delay product. The obvious minima for power dissipation are when the series resistance is infinite and thus the static power consumption becomes zero. This means that the optimization predicts infinitely thick spacer layers and this instability is inherited in the optimization of the power-delay product. In addition, the propagation delay optimization is rather uninteresting, as it finds minima with diminishing returns of speed, at an unreasonable cost of power dissipation.

11.1 Topography

To achieve optimal performance, it is important to know the optimum topographical parameters of the inverter design. In the following plots, the optimum of both spacer layer thicknesses and the number of nanowires in the design is predicted.

11.1.1 Spacer layers

As the wire dimensions are scaled down, it is suspected that the spacer layers should be downscaled as well. The results are illustrated in Figure 52 and it is observed that the scaling is performed in essentially two regimes: one with very thick spacer layers and one with thicknesses around that of the nanowire diameter.

Thick spacer layers are observed for nanowires with diameter above 67 nm , while the thinner spacer layers are reserved for the nanowires with a small diameter. For diameters between 67 and 56 nm, a smaller, intermediate scaling regime is observed. The oxide thickness is generally as small as possible, but with some small fluctuations.

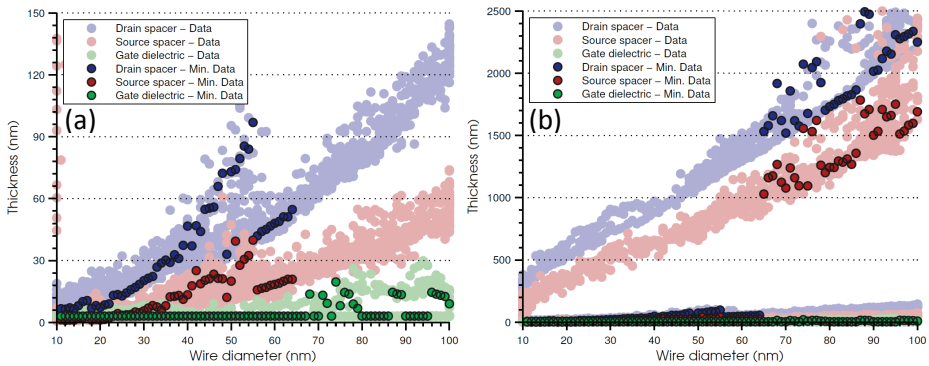


Figure 52 – Effect of scaling down the nanowire diameter on the spacer layers. Notice the different thickness scales in (a) and (b).

11.1.2 Array sizes

As well as adjusting the thickness of the spacer layers, the optimal ratio of nanowires in the load and PDN transistor should change with the scaling. The scaling dependence on the number of nanowires is shown in Figure 53. It is noticed that the optimum wire quantities are kept low for most of the scaling. However, at diameters around 70 nm and below 25 nm, it is observed that the wire numbers increases drastically. The reason for the kink at 70 nm is found in the nanowire ratio figure.

For most of the scaling, the ratio is held at either three to one or two to one. The strange behaviour for diameters around 70 nm is because the optimum wire ratio is somewhere between these two ratios. However, since only integer number of wires is allowed, the number of wires has to increase to accommodate the ratio. As the wire ratio changes abruptly in this region, this discontinuity is reflected, also in the scaling of the other parameters. However, the sharp increase in both the number of wires and the wire ratio for low diameters cannot be explained by the same theory.

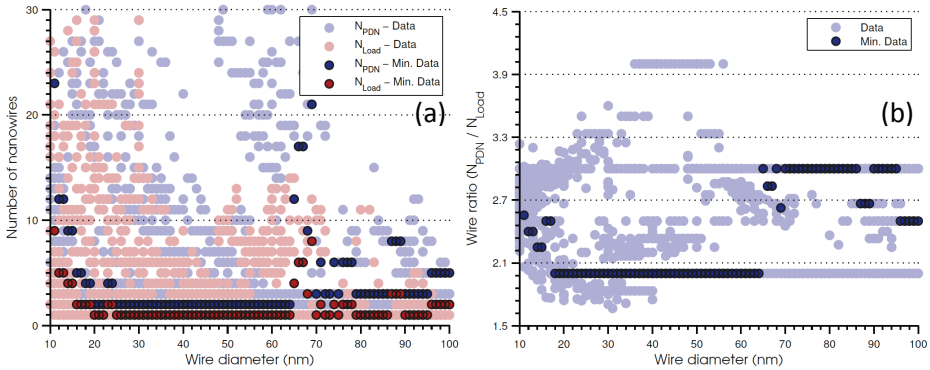


Figure 53 – Effect of scaling down the nanowire diameter on the number of nanowires (a) and optimum nanowire ratio in (b).

11.2 Series resistance

When the number of nanowires and spacer layer thicknesses are changed by the scaling, the capacitances and series resistance change as well. The series resistance of each nanowire is directly proportional to the thickness of the spacer layers. It is therefore not surprising to have both a region of low resistance and one of high resistance, as shown in Figure 54(a). However, the resistance has a stronger dependence on the cross-sectional area of the nanowire. This explains the increase in resistance for smaller diameters.

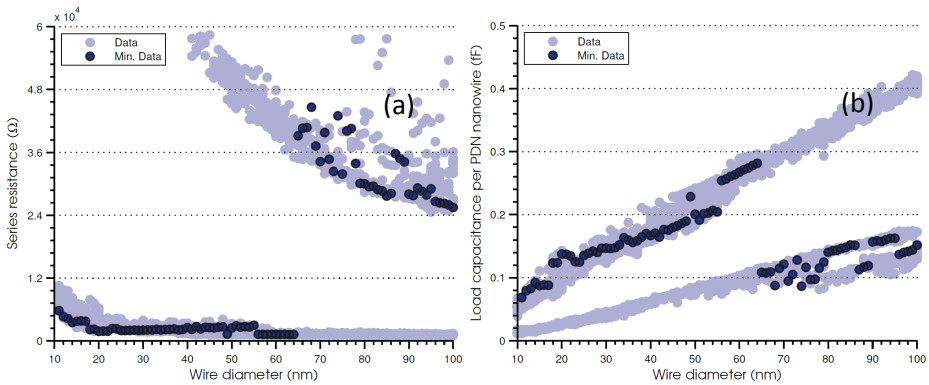


Figure 54 – Effect of scaling down the nanowire diameter on the series resistance in (a) and on the load capacitance, scaled with the number of nanowires in PDN, in (b).

11.3 Load capacitance

In Figure 54(b), it is observed that the load capacitance decreases approximately linearly, in two different segments corresponding to different

optimization domains. Keeping the capacitance change constant probably is an optimum compromise between noise resistance and device performance.

The reason for the capacitance decrease is partly caused by a smaller nanowire area, effectively reducing the wire capacitances, and partly by the smaller electrodes, due the downscaling. As the nanowire quantity takes discreet and not continuous steps, discontinuities between the minima are caused. The sharp discontinuity is probably due to thinner spacer layers, thereby increasing the overlap capacitance.

11.4 Voltage levels

As an optimal topographical design now has been proposed, the next step is to study the optimal voltage levels to drive the inverter. The interesting voltage levels are depicted in Figure 55. As in the previous plots, discontinuities are seen at 70 nm and below 25 nm.

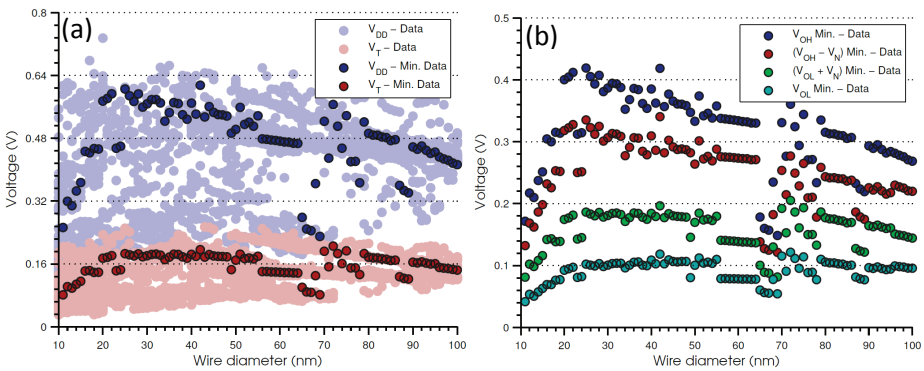


Figure 55 – (a) Effect of scaling down the nanowire diameter on the voltage levels. (b) The static voltage levels of the inverter, the threshold voltage is constant at the output-low with noise margin level.

In Figure 55(a), the proposed threshold voltages of the transistors and the corresponding optimal supply voltages are shown. It is observed that the voltages are approximately constant, with a small decreasing trend, for the entire interval, with the exclusion for very small diameters. The supply voltage is confirmed to be low and does not go above 0.55 V. This strongly indicates that low supply voltages are a requirement for optimum performance of this inverter design. The threshold voltage is also approximately constant over a large interval followed by a decrease for the thin nanowires. The level is held

below 0.2 V, which is slightly higher than the threshold voltage of the measured transistors.

In Figure 55(b), the static voltage levels are plotted. The main thing that can be said about this data is that the output-low level with added noise is at the same level as the threshold voltage. This confirms that the optimal voltage levels can be determined, simply by studying the intersection of the output-high and output-low voltages with added noise margin. This principle is demonstrated by the circle in the regeneracy triangle in Figure 50.

11.4.1 Supply voltage determination

As it is difficult to establish a general trend in supply voltage, it is weighted with the total gate width in the PDN transistor, as shown in Figure 56(a). It is clearly observed that the supply voltage decreases with an increasing number on nanowires or nanowire diameter. As the local minima points indicate, the relationship seems to be exponential in nature.

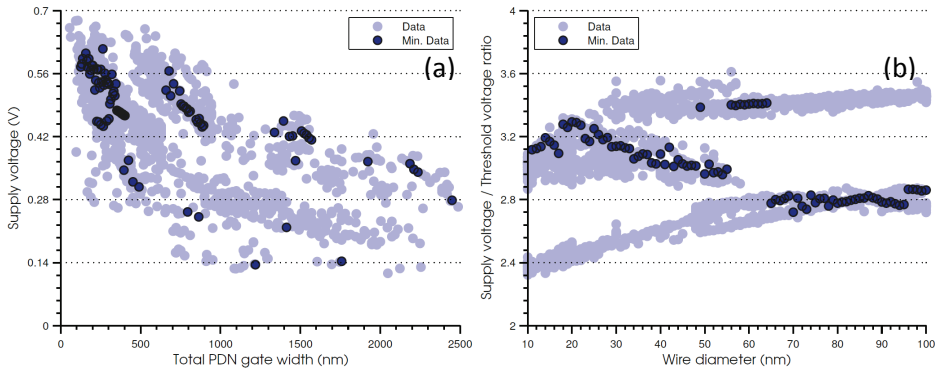


Figure 56 – (a) Determining the optimum supply voltage by taking the total PDN gate width. (b) The optimum ratio between supply voltage and threshold voltage.

11.4.2 Ratio

In Figure 56(b), the ratios between the supply voltages and the threshold voltages are presented. It is noted that the ratio between the two voltages are essentially constant around 3, but the local minima points indicate a more complicated relationship.

11.5 Performance

As the optimum design and voltage levels and their downscaling behaviour are now concluded, the performance metrics of the device are studied.

11.5.1 Propagation delay and power dissipation

In Figure 57, it is clearly observed that scaling down the wire diameter decreases the propagation delay at the cost of increased power dissipation. For both figures of merit, essentially two different scaling behaviours are seen and, again, the global scaling behaviour is changed at a diameter of 60-70 nm.

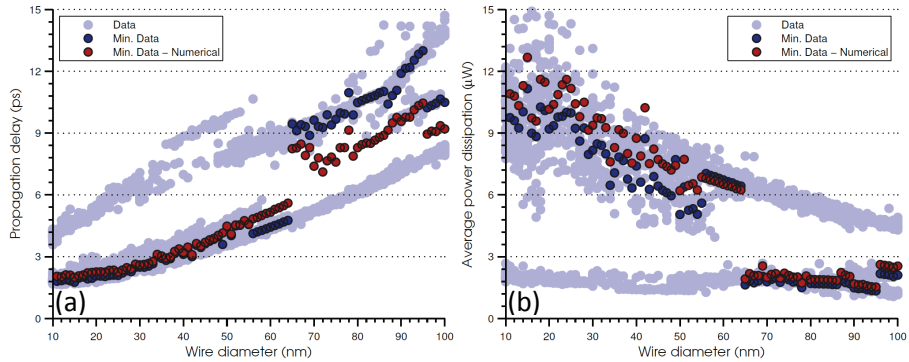


Figure 57 – Effect of scaling down the nanowire diameter on the propagation delay in (a) and power dissipation in (b).

The analytical modelling is compared to the corresponding numerical values and it is concluded that the two models correspond very well. This indicates that the assumptions and simplifications done in the analytical solution are correct and properly applied.

11.5.2 PDP and EDP

In Figure 58, the power-delay and energy-delay products are shown, again compared to numerical data.

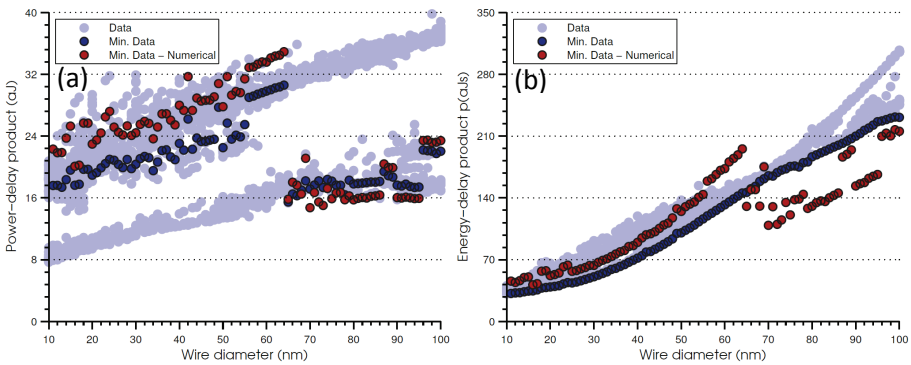


Figure 58 – Effect of scaling down the nanowire diameter on the power-delay product in (a) and energy-delay product in (b).

It is clearly observed that the trend for the energy-delay product, as well as the power-delay product, gets progressively smaller with decreasing nanowire dimensions. It is thus concluded that further downscaling is advisable. It is important to take into account that the optimization was EDP-based and therefore the PDP behaviour is not optimally derived.

The comparison between the numerical and analytical PDP values shows that they correspond reasonably well. However, as can be seen in the EDP plot, the analytical expressions may result in either an underestimation or an overestimation of the metric, depending on the optimization behaviour.

Gate length compensation

In order to simplify benchmarking against competing technologies, the EDP and PDP is divided by the gate length. The result is illustrated in Figure 59.

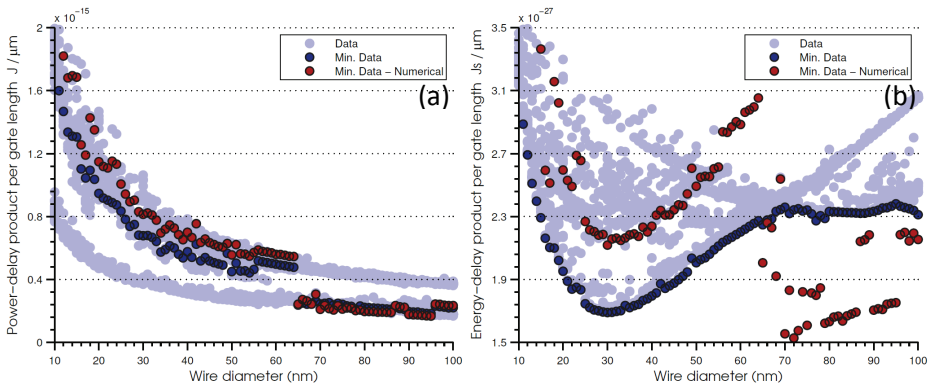


Figure 59 – Effect of scaling down the nanowire diameter on the scaled power-delay product in (a) and scaled energy-delay product in (b).

Another EDP behaviour is now becoming apparent, with diminishing returns for downscaling, below about 30 nm. From the local minima data points, two different parabolas are seen, which represents two different scaling regimes.

As for the power-delay product, the case is less clear, as this analysis predict that larger diameter wires may be better suited for low-power applications. From this point of view, the gains from downscaling diminish as the nanowires are scaled down.

The difference between the numerical and analytical models is low for the gate length scaled PDP. However, the contrast is larger between the models

for the scaled EDP. Additionally, it is seen that another global minimum is present in the analytical model. This is due to the propagation delay being underestimated by the analytical model for high nanowire diameters.

It is probable that a numerical optimization would behave similarly to the analytical optimization, although this may not be determined without a massive amount of computing power. However, the models correspond well for the low nanowire diameters and it is probable that the position of the global minimum is correctly predicted.

Optimization	NW Diameter	V_{DD} (V)	V_T (V)	N_{load}	N_{PDN}	l (nm)
EDP	10 nm	0.136	0.058	137	374	64
	30 nm	0.568	0.182	1	2	144
	50 nm	0.502	0.170	1	2	224
	70 nm	0.523	0.192	1	3	304
	100 nm	0.413	0.144	2	5	455
Optimization	**NW Diameter**	t_{gs} (nm)	t_{gd} (nm)	t_{ox} (nm)	W_E (nm)	H_E (nm)
EDP	10 nm	2054	333	3	32	10
	30 nm	5.6	20	3	72	30
	50 nm	20	73	3	112	50
	70 nm	1077	1518	3	152	70
	100 nm	1691	2250	9	236	100
Optimization	**NW Diameter**	τ_p (ps)	P_A (µW)	PDP (aJ)	EDP p(aJs)	
EDP	10 nm	3.97	1.99	7.89	31.4	
	30 nm	2.49	8.17	20.3	50.7	
	50 nm	4.46	5.05	22.5	100	
	70 nm	9.63	1.78	17.2	165	
	100 nm	10.5	2.10	22.0	231	

TABLE 6 – OPTIMUM PARAMETERS FOR DIFFERENT TRANSISTOR DESIGNS.

11.6 Scaling summary

As extraction of parameter from plots may be a tedious task, the results are summarized for a few key nanowire diameters in Table 6.

11.6.1 Model comparison

In the following table, the optimization results are compared using the different solution methods, derived from the Jansson-Berg model, as described in section 8.9.

NW Diameter	Model	τ_p (ps)	P_A (μW)	PDP (aJ)	EDP p(aJs)
10 nm	Numerical	3.33	2.47	8.23	27.4
	Advanced analytical	3.97	1.99	7.89	31.3
	Simple analytical	9.11	1.75	15.9	145
	Simple analytical without R_s	4.73	3.37	15.9	75.4
50 nm	Numerical	4.48	6.20	27.8	125
	Advanced analytical	4.46	5.05	22.5	100
	Simple analytical	11.0	1.35	14.8	163
	Simple analytical without R_s	10.7	1.39	14.8	158
100 nm	Numerical	9.20	2.54	23.4	215
	Advanced analytical	10.5	2.09	22.0	231
	Simple analytical	33.8	0.83	28.2	952
	Simple analytical without R_s	29.6	0.95	28.2	834

TABLE 7 – MODEL COMPARISON FOR OPTIMIZED PARAMETERS.

It is concluded that the advanced analytical model has an exceptional correspondence to the numerical model. The simple analytical model is a lot worse, although it is surprisingly accurate, given that the model is based on very rough assumptions. It is interesting to note that the simple analytical model corresponds better to the numerical data when it is not incorporating a series resistance. This may be due to the capacitance simplifications counteracting the lack of a dependence on the series resistance.

11.7 Optimal transistor characteristics

As the optimum transistor design, based on EDP optimization, is now concluded, it is possible to predict the transistor characteristics of this downscaled design.

11.7.1 Scaled EDP optimum design

The optimal design predicted by the gate-length scaled energy-delay product has a nanowire diameter of 30 nm. In order to get an impression of how this transistor design would behave, the transistor characteristics are simulated. The simulated characteristics are presented in Figure 60, using the parameter set found in Table 8.

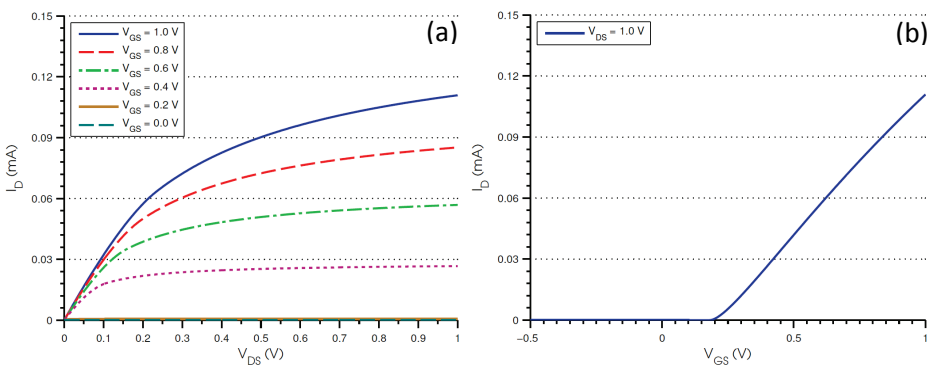


Figure 60 – Transistor characteristics for the gate-length scaled EDP-optimized design with 30 nm nanowire diameter.

	V_T (V)	V_C (V)	k' (μA/V²)	R_s (kΩ)	N
Value	0.1808	0.048	1637	2.065	1

TABLE 8 – TRANSISTOR PARAMETERS FOR THE EDP-OPTIMIZED DESIGN WITH 30 NM NANOWIRE DIAMETER.

Reduction of series resistance

One of the greatest limitations of this transistor design is the large series resistance. It would be interesting to see how the performance would improve if the resistance is reduced, e.g. by doping of the nanowire. In Figure 61, the transistor characteristics is shown, when the resistivity is reduced by a factor of ten. It is seen that the characteristic is greatly improved.

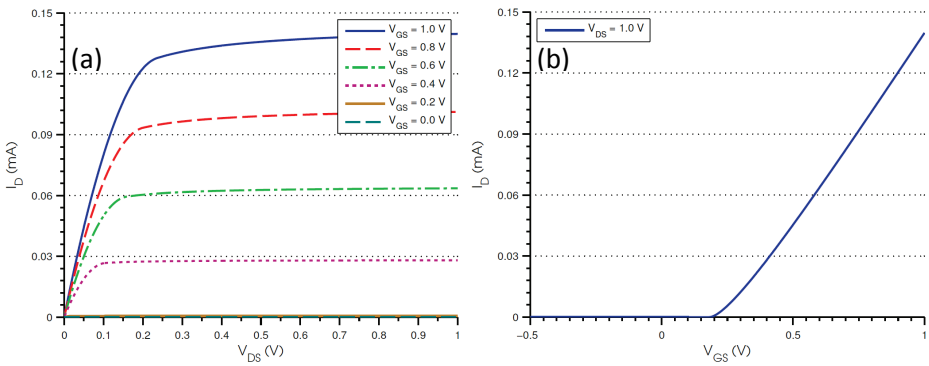


Figure 61 – Transistor characteristics for the gate-length scaled EDP-optimized design with 30 nm nanowire diameter, with a reduction of the series resistance by a factor of ten.

11.7.2 Minimum feature size transistor

As scaling towards nanowire diameters of 10 nm is on the roadmap, it would be interesting to investigate the performance of these transistors further. By using the parameters in Table 9, the transistor characteristics in Figure 62 are constructed.

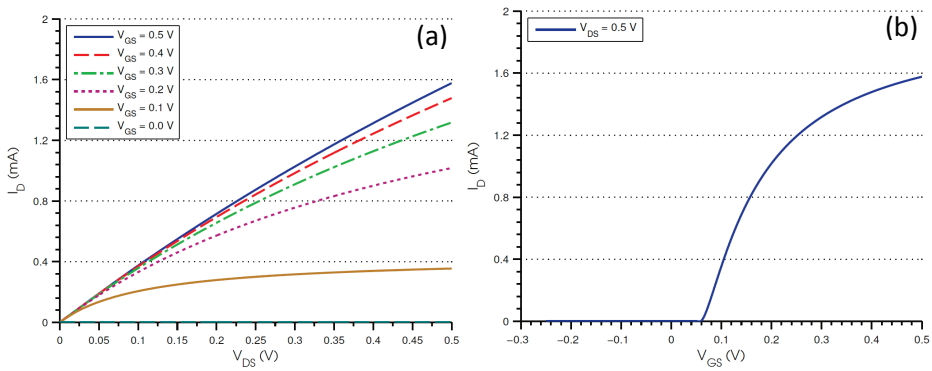


Figure 62 – Transistor characteristics for the 10 nm nanowire diameter design.

	V_T (V)	V_C (V)	k' (μA/V^2)	R_s (kΩ)	N
Value	0.0576	0.016	3447	34.92	137

TABLE 9 – TRANSISTOR PARAMETERS FOR THE 10 NM NANOWIRE DIAMETER DESIGN.

It is clearly observed that the characteristics are highly resistive and this probably explains why the design for optimum performance per gate length is for nanowires with a diameter of 30 nm. As the resistance of a cylinder increases with a square behaviour on the radius, but only decreases linearly with the height, the resistance is a large problem for small designs.

Reduction of series resistance

As the resistance is reduced a factor of ten, the characteristics of the transistor is greatly improved, as can be seen in Figure 63. It is thus concluded, that in order to scale the transistor design to very low dimensions, reduction of the series resistance is essential.

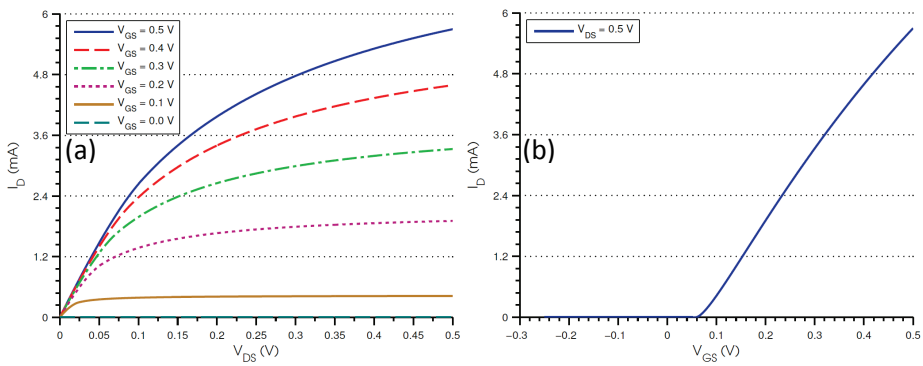


Figure 63 – Transistor characteristics for the 10 nm nanowire diameter design, with a reduction of the series resistance by a factor ten.

Part XII CONCLUSIONS

12.1 Modelling

In this report, a modified MOSFET model is introduced, based on the proper deep-submicron MOSFET model. By adding a series resistance and introducing a dependence on the nanowire quantity, in addition to neglecting the channel length modulation, the Jansson-Berg model is defined. This model corresponds very well to measurement data.

In addition, a specialised nanowire model is investigated and modified for maximum flexibility and simplicity. It is also reformulated to take the corresponding form of conventional MOSFET models, as well as introducing series resistance and a dependence on the number of nanowires in the array. This model, and similar models, may be more suitable for future use, as more of the physical parameters may be pre-calculated. Nevertheless, a good fitting with a physically plausible parameter set, is shown to fit well to measurement data.

To simulate the parasitic capacitances in the transistor design, a model consisting of 21 separate capacitances is derived and the relative contributions of these are quantified.

12.2 Circuit design

The Jansson-Berg model is analysed in the application of simulating a digital inverter. As no p-type transistors are available, it is decided to use a ratioed logic design. As an exact analytical solution of the resulting differential equation is near impossible, it is concluded that series solutions are necessary. Although simplifications are made, the analytical solutions achieve a good correspondence to a purely numerical analysis.

To provide a complete workflow from measurement, through transistor modelling and circuit simulation to an optimized inverter design, an in-house computer application has been developed. By implementing intelligent algorithms and extraction methods through regression analysis, the workflow is highly automated. The software suite reduces the time, needed to do a complete analysis cycle, from weeks, using manual analysis, to a matter of minutes.

It is confirmed that ratioed logic has large drawbacks in comparison with conventional CMOS implementations. The most significant is the large static power consumption, resulting from a direct current path from supply to ground, when the input signal is high. Additionally, the existence of static output-low and output-high voltages, that are far from ideal, severely restricts the operation range of the devices, as it results in non-regenerative gates.

12.3 Optimization

As the optimum voltage levels in the device are very low, the influence of thermal noise is a very significant and limiting factor. The thermal noise decreases with an increased gate capacitance and the main method to combat the noise is to increase the number of nanowires in the transistor array. However, an increase in the number of nanowires in the array increases the parasitic capacitances, which in turn reduces the device performance. One of the difficulties in the optimization is finding the optimum balance between noise resistance and performance.

As the dimensions of the inverter are downscaled, the propagation delay is decreased, at the cost of increased power dissipation. These performance metrics are combined to form the energy-delay product and it is observed that it decreases with shrinking dimensions. It is thus concluded that continued downscaling, of the nanowire transistor design, is desired.

Part XIII APPENDIX

13.1 Simple MOSFET sub-threshold modelling

In the simple MOSFET models, instead of setting the current in the sub-threshold region to zero, the sub-threshold current can be modelled in the following way [**9**].

$$I_D = k(\varsigma - 1)\varphi_{th}^2 e^{\frac{V_{GS}-V_T}{\varsigma\varphi_{th}}}\left(1 - e^{-\frac{V_{DS}}{\varphi_{th}}}\right) \tag{95}$$

$$\left[\varsigma = \varsigma_0 + \frac{C_d}{C_{ox}}\right] \tag{96}$$

Here, ς_0 is an empirical parameter. This model has been validated to model the sub-threshold current adequately, but the results are omitted in this report.

13.2 Sub-threshold NWFET model

The sub-threshold current of the NWFET model is presented in (97) with its corresponding band structure parameter, α, defined in (98).

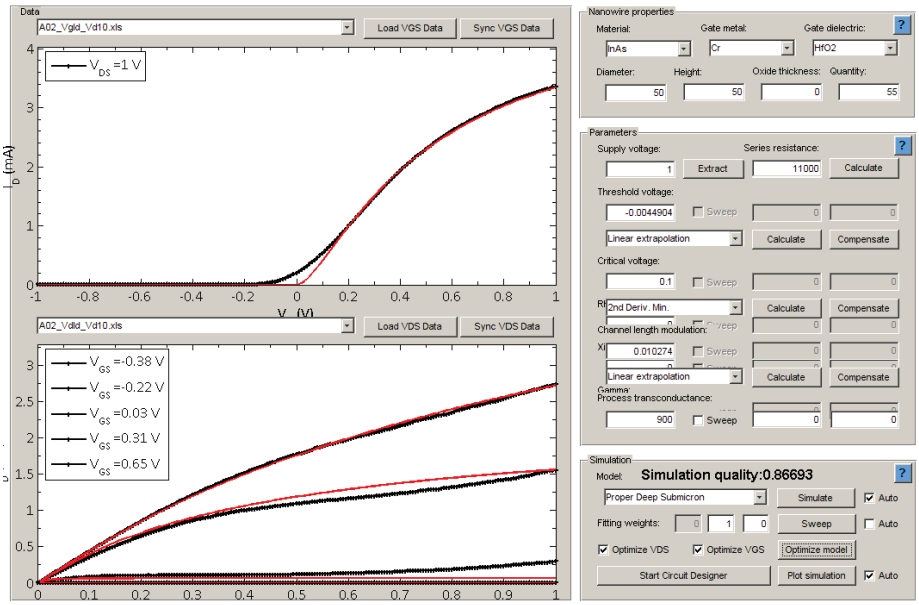
$$I_{DS} = \frac{\phi_{th}\mu C_{ox}}{L_{eff}}\,\text{lambertw}\left(\frac{\alpha}{\phi_{th}C_{ox}}\,\text{e}^{\frac{V_{GS}-\phi_{ms}}{\phi_{th}}}\right) \cdot \left(V_{DS} + \phi_{th}\left(1 - \text{e}^{-\frac{V_{DS}}{\phi_{th}}}\right)\right) \tag{97}$$

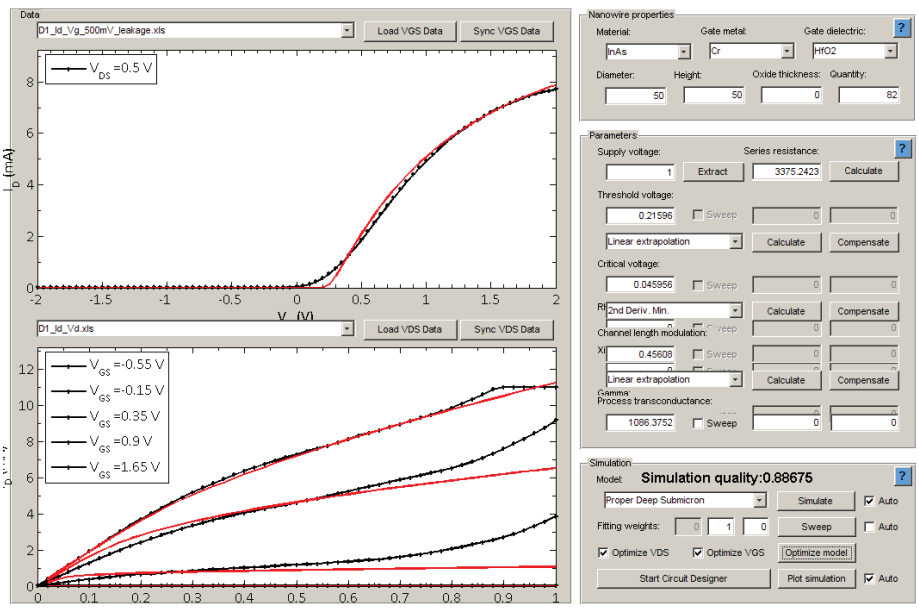
$$\left[\alpha = N_0 \sum_n \sum_v g_v \sqrt{m_{d,v}} \cdot \sqrt{\pi kT} \cdot \text{e}^{\frac{-E_{v,n}}{kT}}\right] \tag{98}$$

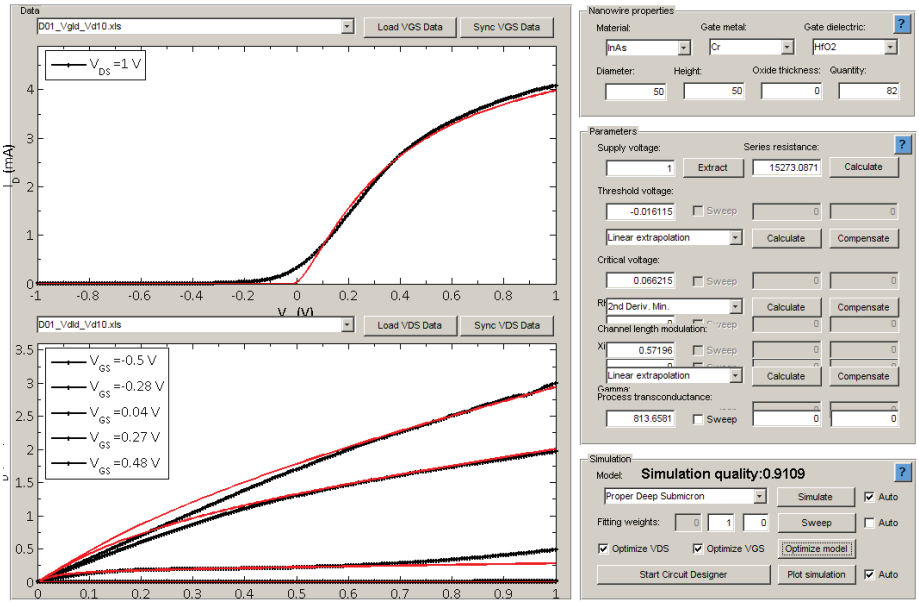
The modelling of the sub-threshold current has been shown to follow measurement data closely, but confirmation of this is out of scope of this project.

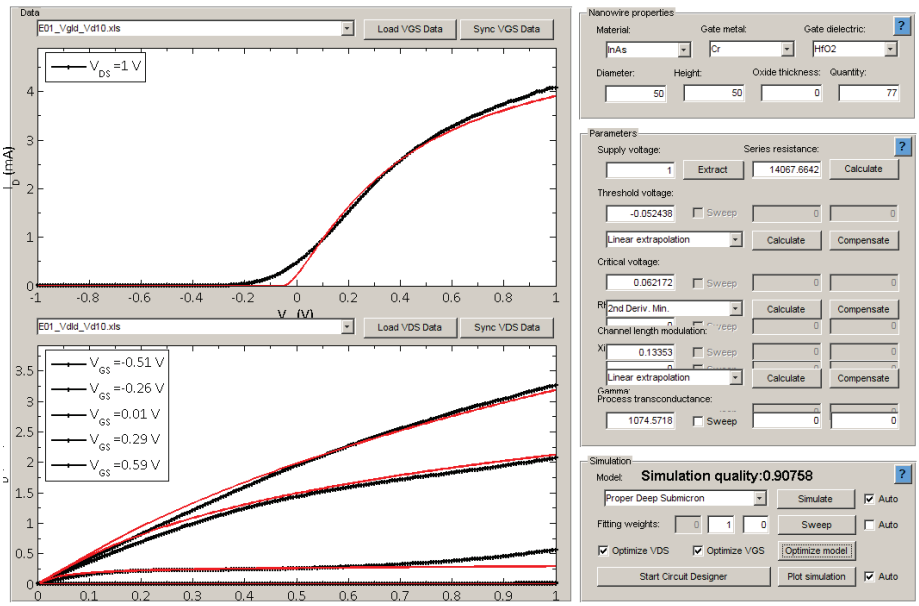
13.3 Additional model fitting data

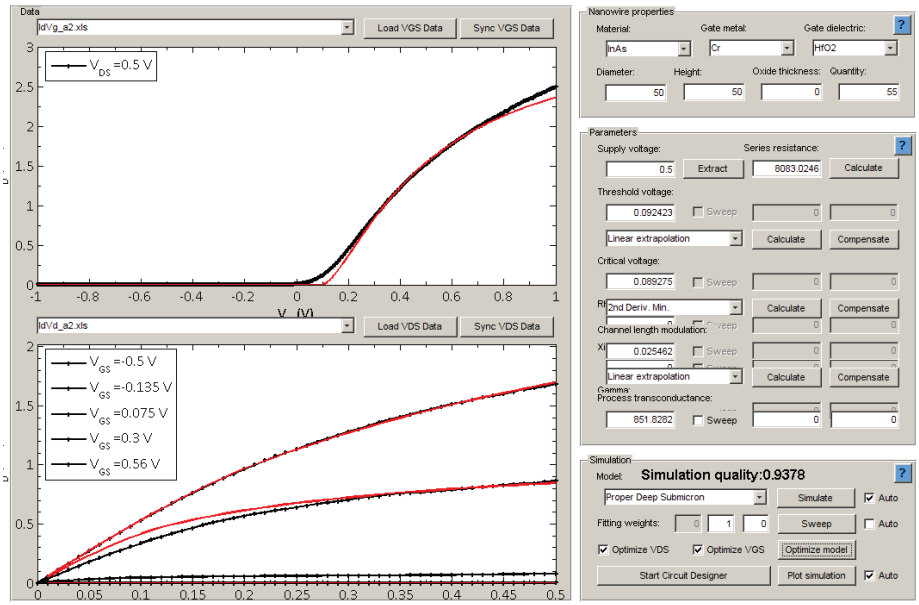
Some examples of the fitting results obtained by the modelling software are presented below. The figures presents both the $I_D - V_{GS}$ and $I_D - V_{DS}$ characteristics, as well as their corresponding parameter sets. The simulation quality values are also shown, as defined in section 5.1.2.

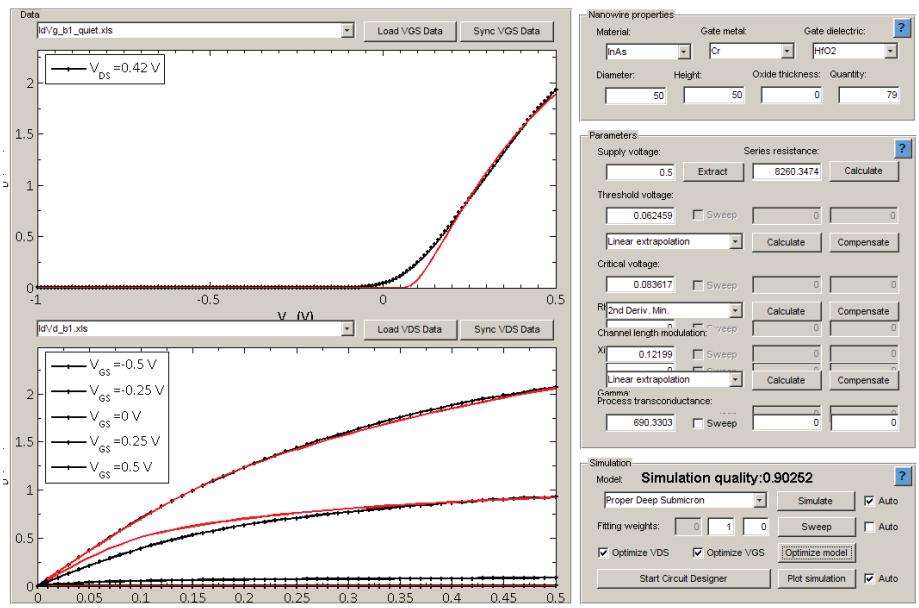

Data
A02_Vgld_Vdt0.xls
Load VGS Data
Sync VGS Data
V_DS=1 V
I_D (mA)
V (V)
A02_Vdld_Vdt0.xls
Load VDS Data
Sync VDS Data
V_GS=-0.38 V
V_GS=-0.22 V
V_GS=0.03 V
V_GS=0.31 V
V_GS=0.65 V
Nanowire properties
Material:
Gate metal:
Gate dielectric:
InAs
Cr
HfO2
Diameter:
Height:
Oxide thickness:
Quantity:
50
50
0
55
Parameters
Supply voltage:
Series resistance:
1
Extract
11000
Calculate
Threshold voltage:
-0.0044904
Sweep
Linear extrapolation
Calculate
Compensate
Critical voltage:
0.1
2nd Deriv. Min.
Channel length modulation:
0.010274
Process transconductance:
900
Simulation
Model:
Simulation quality:0.86693
Proper Deep Submicron
Simulate
Auto
Fitting weights:
Sweep
Optimize VDS
Optimize VGS
Optimize model
Start Circuit Designer
Plot simulation


Data
D1_Id_Vg_500mV_leakage.xls
Load VGS Data
Sync VGS Data
V_DS=0.5 V
I_D (mA)
V (V)
D1_Id_Vd.xls
Load VDS Data
Sync VDS Data
V_GS=-0.55 V
V_GS=-0.15 V
V_GS=0.35 V
V_GS=0.9 V
V_GS=1.65 V
Nanowire properties
Material:
Gate metal:
Gate dielectric:
InAs
Cr
HfO2
Diameter:
Height:
Oxide thickness:
Quantity:
50
50
0
82
Parameters
Supply voltage:
Series resistance:
1
Extract
3375.2423
Calculate
Threshold voltage:
0.21596
Sweep
Linear extrapolation
Calculate
Compensate
Critical voltage:
0.045956
2nd Deriv. Min.
Channel length modulation:
0.45608
Process transconductance:
1086.3752
Simulation
Model:
Simulation quality:0.88675
Proper Deep Submicron
Simulate
Auto
Fitting weights:
Sweep
Optimize VDS
Optimize VGS
Optimize model
Start Circuit Designer
Plot simulation


Data
D01_Vgld_Vd10.xls
Load VGS Data
Sync VGS Data
V_{DS} =1 V
I_D (mA)
V (V)
D01_Vdld_Vd10.xls
Load VDS Data
Sync VDS Data
V_{GS} =-0.5 V
V_{GS} =-0.28 V
V_{GS} =0.04 V
V_{GS} =0.27 V
V_{GS} =0.48 V
Nanowire properties
Material:
Gate metal:
Gate dielectric:
InAs
Cr
HfO2
Diameter:
Height:
Oxide thickness:
Quantity:
50
50
0
82
Parameters
Supply voltage:
Series resistance:
1
Extract
15273.0871
Calculate
Threshold voltage:
-0.016115
Sweep
Linear extrapolation
Calculate
Compensate
Critical voltage:
0.066215
2nd Deriv. Min.
Channel length modulation:
0.57196
Process transconductance:
813.6581
Simulation
Model:
Simulation quality:0.9109
Proper Deep Submicron
Simulate
Auto
Fitting weights:
Sweep
Optimize VDS
Optimize VGS
Optimize model
Start Circuit Designer
Plot simulation


Data
E01_Vgld_Vd10.xls
Load VGS Data
Sync VGS Data
V_{DS} =1 V
I_D (mA)
V (V)
E01_Vdld_Vd10.xls
Load VDS Data
Sync VDS Data
V_{GS} =-0.51 V
V_{GS} =-0.26 V
V_{GS} =0.01 V
V_{GS} =0.29 V
V_{GS} =0.59 V
Nanowire properties
Material:
Gate metal:
Gate dielectric:
InAs
Cr
HfO2
Diameter:
Height:
Oxide thickness:
Quantity:
50
50
0
77
Parameters
Supply voltage:
Series resistance:
1
Extract
14067.6642
Calculate
Threshold voltage:
-0.052438
Sweep
Linear extrapolation
Calculate
Compensate
Critical voltage:
0.062172
2nd Deriv. Min.
Channel length modulation:
0.13353
Process transconductance:
1074.5718
Simulation
Model:
Simulation quality:0.90758
Proper Deep Submicron
Simulate
Auto
Fitting weights:
Sweep
Optimize VDS
Optimize VGS
Optimize model
Start Circuit Designer
Plot simulation


Data
IdVg_a2.xls
Load VGS Data
Sync VGS Data
V_DS =0.5 V
V (V)
IdVd_a2.xls
Load VDS Data
Sync VDS Data
V_GS =-0.5 V
V_GS =-0.135 V
V_GS =0.075 V
V_GS =0.3 V
V_GS =0.56 V
Nanowire properties
Material:
Gate metal:
Gate dielectric:
InAs
Cr
HfO2
Diameter:
Height:
Oxide thickness:
Quantity:
50
50
0
55
Parameters
Supply voltage:
Series resistance:
0.5
Extract
8083.0246
Calculate
Threshold voltage:
0.092423
Sweep
Linear extrapolation
Calculate
Compensate
Critical voltage:
0.089275
2nd Deriv. Min.
Channel length modulation:
0.025462
Process transconductance:
851.8282
Simulation
Model:
Simulation quality:0.9378
Proper Deep Submicron
Simulate
Auto
Fitting weights:
Sweep
Optimize VDS
Optimize VGS
Optimize model
Start Circuit Designer
Plot simulation


Data
IdVg_b1_quiet.xls
Load VGS Data
Sync VGS Data
V_DS =0.42 V
V (V)
IdVd_b1.xls
Load VDS Data
Sync VDS Data
V_GS =-0.5 V
V_GS =-0.25 V
V_GS =0 V
V_GS =0.25 V
V_GS =0.5 V
Nanowire properties
Material:
Gate metal:
Gate dielectric:
InAs
Cr
HfO2
Diameter:
Height:
Oxide thickness:
Quantity:
50
50
0
79
Parameters
Supply voltage:
Series resistance:
0.5
Extract
8260.3474
Calculate
Threshold voltage:
0.062459
Sweep
Linear extrapolation
Calculate
Compensate
Critical voltage:
0.083617
2nd Deriv. Min.
Channel length modulation:
0.12199
Process transconductance:
690.3303
Simulation
Model:
Simulation quality:0.90252
Proper Deep Submicron
Simulate
Auto
Fitting weights:
Sweep
Optimize VDS
Optimize VGS
Optimize model
Start Circuit Designer
Plot simulation

13.4 Example of a generated report

WIGFET Modelling and Simulation Report

(14-Sep-2009 10:17:39)

Dataset VGS: IdVg_a2.xls

Dataset VDS: IdVd_a2.xls

Transistor Parameters

Nanowire material: InAs

Nanowire diameter: 50 nm

Nanowire gate height: 50 nm

Number of nanowires: 55

Gate metal: Cr

Gate dielectric: HfO2

Gate dielectric thickness: 10 nm

Transistor Topography

Nanowire spacing: 200 nm

Electrode widths: 100 nm

Electrode heights: 100 nm

Source spacer material: SiO_2

Source spacing: 100 nm

Drain spacer material: S1805

Drain spacing: 200 nm

Transistor Modelling

Model used: Proper Deep Submicron

Modelling quality: 93.6326%

Modelling quality VG: 95.8632%

Modelling quality VDS: 91.4019%

Process transconductance: 775.6085 $\mu A/V^2$

Threshold voltage: 0.092423 V

Critical voltage: 0.1035 V

Channel length modulation: 0

Series resistance: 8111.0302 Ω

Capacitance Calculation

Oxide capacitance: 40.5241 aF

Overlap capacitances:

Gate-Source: 2.3796 aF

Gate-Drain: 0.68084 aF

Source-Drain: 1.3899 aF

Wire capacitances:

Wire (Gate-Source): 62.1028 aF

Wire (Gate-Drain): 39.1048 aF

Wire-X (Gate-Source): 1.6629 aF

Wire-X (Gate-Drain): 0.64315 aF

Wire-Y (Gate-Source): 7.4642 aF

Wire-Y (Gate-Drain): 5.7415 aF

Fringe capacitances:

Fringe (Gate-Source): 5.5092 aF

Fringe (Gate-Drain): 1.7717 aF

Fringe (Source-Drain): 1.7542 aF

Fringe-X (Gate-Source): 2.2608 aF
Fringe-X (Gate-Drain): 0.79868 aF
Fringe-Y (Gate-Source): 2.4245 aF
Fringe-Y (Gate-Drain): 0.836 aF
Edge capacitances:
Edge-X (Gate-Source): 0.92256 aF
Edge-X (Gate-Drain): 0.35245 aF
Edge-Y (Gate-Source): 0.92256 aF
Edge-Y (Gate-Drain): 0.35245 aF
Edge-Y (Source-Drain): 0.43854 aF

Inverter capacitances

Input Capacitance: 0.089806 fF
Load Capacitance: 0.40306 fF
Load Transistor Capacitances:
Intrinsic capacitance: 0.040524 fF
Gate-Source Parasitic: 0.073682 fF
Gate-Drain Parasitic: 0.042967 fF
Source-Drain Parasitic: 0.0040212 fF
PDN Transistor Capacitances:
Intrinsic capacitance: 0.081048 fF
Gate-Source Parasitic: 0.15329 fF
Gate-Drain Parasitic: 0.089806 fF
Source-Drain Parasitic: 0.0075286 fF
Total Intrinsic Capacitance: 0.4611 fF
Total Parasitic Capacitance: 0.12157 fF

Circuit Simulation Parameters

Simulation Time: 500 ps

Number of Iterations: 25

Manual Capacitance Input: Off

Miller Effect Simplification: Off

Supply Voltage: 0.3 V

Fan-out: 1

Number of Nanowires in Load Transistor: 1

Number of Nanowires in PDN Transistor: 2

Figures of Merit

Nanowire Mobility: 1503.2088 cm^2/Vs

Propagation Delay: 4.8566 ps

Average Power Dissipation: 3.5868 µW

Power Delay Product: 17.4194 aJ

Energy Delay Product: 84.599 p(aJs)

Output-high: 0.30634 V

Output-low: 0.022702 V

Noise Margin: 0.069721 V

Regenerative: True

Report Generated by:

TDS – NWFET Modelling Suite 2009

Developed at Lund Institute of Technology

13.5 Full capacitance model

The complete expressions for all the 21 different parasitic capacitances are presented below. No simplifications have been made on the geometrics of the layout, e.g. the electrode widths W_D and W_S, are not held equal. For a general description of the parasitic capacitances in the quadratic array, see Figure 10.

13.5.1 Electrode overlap Capacitances

$$C_{GD,o} = \frac{\varepsilon_0 \varepsilon_{GD} \left(W_G W_D - \pi (R + t_{ox})^2\right)}{d_{GD}}$$

$$C_{GS,o} = \frac{\varepsilon_0 \varepsilon_{GS} \left(W_G W_S - \pi (R + t_{ox})^2\right)}{d_{GS}}$$

$$C_{SD,o} = \frac{\varepsilon_0 \varepsilon_{GS} \left(W_S l_x - \pi (R + t_{ox})^2\right)}{d_{DS}}$$

13.5.2 Nanowire-coupled transistors

$$C_{GD,w} = 4R\varepsilon_0 \varepsilon_{GD} \left(\ln\left(\frac{a_1}{\frac{\varepsilon_{GD}}{\varepsilon_{ox}} t_{ox}} \right) + \frac{1}{\pi} \ln\left(\frac{a_2}{\frac{\varepsilon_{GD}}{\varepsilon_{ox}} t_{ox} + \frac{W_G}{2} - R - t_{ox}} \right) \right)$$

$$\begin{bmatrix} a_1 = T_{GD,w,1} + \sqrt{\left(\frac{\varepsilon_{GD}}{\varepsilon_{ox}} t_{ox}\right)^2 + T_{GD,w,1}^2} \\ a_2 = T_{GD,w,2} + \sqrt{\left(\frac{\varepsilon_{GD}}{\varepsilon_{ox}} t_{ox} + \frac{W_G}{2} - R - t_{ox}\right)^2 + T_{GD,w,2}^2} \\ T_{GD,w,1} = d_{GD} \mathrm{e}^{\frac{\frac{\varepsilon_{GD}}{\varepsilon_{ox}} t_{ox} + \frac{W_G}{2} - R - t_{ox} - \sqrt{\left(\frac{\varepsilon_{GD}}{\varepsilon_{ox}} t_{ox}\right)^2 + d_{GD}^2}}{\left(\frac{W_G}{2} - R - t_{ox}\right)\tau}} \\ T_{GD,w,2} = d_{GD} \mathrm{e}^{\frac{\frac{\varepsilon_{GD}}{\varepsilon_{ox}} t_{ox} + \frac{l_y}{2} - R - t_{ox} - \sqrt{\left(\frac{\varepsilon_{GD}}{\varepsilon_{ox}} t_{ox} + \frac{W_G}{2} - R - t_{ox}\right)^2 + d_{GD}^2}}{\left(\frac{l_y - \mathrm{W_G}}{2}\right)\tau}} \end{bmatrix}$$

$$C_{GS,w} = 4R\varepsilon_0\varepsilon_{GS}\left(\ln\left(\frac{b_1}{\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}}\right)+\frac{1}{\pi}\ln\left(\frac{b_2}{\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{W_G}{2}-R-t_{ox}}\right)\right)$$

$$\left[\begin{array}{l} b_1 = T_{GS,w,1}+\sqrt{\left(\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}\right)^2+T_{GS,w,1}^2} \\ b_2 = T_{GS,w,2}+\sqrt{\left(\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{W_G}{2}-R-t_{ox}\right)^2+T_{GS,w,2}^2} \\ T_{GS,w,1} = d_{GS}\mathrm{e}^{\frac{\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{W_G}{2}-R-t_{ox}-\sqrt{\left(\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}\right)^2+d_{GS}^2}}{\left(\frac{W_G}{2}-R-t_{ox}\right)\tau}} \\ T_{GS,w,2} = d_{GS}\mathrm{e}^{\frac{\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{l_y}{2}-R-t_{ox}-\sqrt{\left(\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{W_G}{2}-R-t_{ox}\right)^2+d_{GS}^2}}{\left(\frac{l_y-W_G}{2}\right)\tau}} \end{array}\right]$$

$$C_{GD,w,y} = \frac{4R}{\pi}\varepsilon_0\varepsilon_{GD}\ln\left(\frac{T_{GD,w,y}+\sqrt{\left(\frac{\varepsilon_{GD}}{\varepsilon_{ox}}t_{ox}+\frac{l_y-W_G}{2}\right)^2+T_{GD,w,y}^2}}{\frac{\varepsilon_{GD}}{\varepsilon_{ox}}t_{ox}+\frac{l_y-W_G}{2}}\right)$$

$$\left[T_{GD,w,y} = d_{GD}\mathrm{e}^{\frac{\frac{\varepsilon_{GD}}{\varepsilon_{ox}}t_{ox}+l_y-W_G-\sqrt{\left(\frac{\varepsilon_{GD}}{\varepsilon_{ox}}t_{ox}+\frac{l_y-W_G}{2}\right)^2+d_{GD}^2}}{\left(\frac{l_y-W_G}{2}\right)\tau}}\right]$$

$$C_{GD,w,x} = \frac{4R}{\pi}\varepsilon_0\varepsilon_{GD}\ln\left(\frac{T_{GD,w,x}+\sqrt{\left(\frac{\varepsilon_{GD}}{\varepsilon_{ox}}t_{ox}+\frac{l_x-W_G}{2}\right)^2+T_{GD,w,x}^2}}{\frac{\varepsilon_{GD}}{\varepsilon_{ox}}t_{ox}+\frac{l_x-W_G}{2}}\right)$$

$$\left[T_{GD,w,x} = d_{GD}\mathrm{e}^{\frac{\frac{\varepsilon_{GD}}{\varepsilon_{ox}}t_{ox}+l_x-R-t_{ox}-\sqrt{\left(\frac{\varepsilon_{GD}}{\varepsilon_{ox}}t_{ox}+\frac{l_x-W_G}{2}\right)^2+d_{GD}^2}}{\left(\frac{W_G}{2}-R-t_{ox}\right)\tau}}\right]$$

$$C_{GS,w,y} = \frac{4R}{\pi}\varepsilon_0\varepsilon_{GS}\ln\left(\frac{T_{GS,w,y}+\sqrt{\left(\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{l_y-W_G}{2}\right)^2+T_{GS,w,y}^2}}{\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{l_y-W_G}{2}}\right)$$

$$\left[T_{GS,w,y} = d_{GS}\mathrm{e}^{\frac{\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+l_y-W_G-\sqrt{\left(\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{l_y-W_G}{2}l_y\right)^2+d_{GS}^2}}{\left(\frac{l_y-W_G}{2}\right)\tau}}\right]$$

$$C_{GS,w,x} = \frac{4R}{\pi}\varepsilon_0\varepsilon_{GS}\ln\left(\frac{T_{GS,w,x}+\sqrt{\left(\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{l_x-W_G}{2}\right)^2+T_{GS,w,x}^2}}{\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{l_x-W_G}{2}}\right)$$

$$\left[T_{GS,w,x} = d_{GS}\mathrm{e}^{\frac{\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+l_x-R-t_{ox}-\sqrt{\left(\frac{\varepsilon_{GS}}{\varepsilon_{ox}}t_{ox}+\frac{l_x-W_G}{2}\right)^2+d_{GS}^2}}{\left(\frac{W_G}{2}-R-t_{ox}\right)\tau}}\right]$$

13.5.3 Electrode fringe capacitances

$$C_{DS,f} = \frac{2(l_x-W_G)}{\pi}\varepsilon_0\varepsilon_{GS}\ln\left(1+\frac{H_D+H_S}{d_{DS}}\right).$$

$$C_{GD,f} = \frac{4}{\pi}\varepsilon_0\varepsilon_{GD}\left(W_D\ln\left(\frac{c_1}{d_{GD}}\right)+W_G\ln\left(\frac{c_2}{d_{GD}}\right)\right)$$

$$\left[c_z = T_{GD,f,z}+\sqrt{d_{GD}^2+T_{GD,f,z}^2}\right.$$

$$T_{GD,f,1} = \frac{l_x-W_G}{2}\mathrm{e}^{\frac{L_G+d_{GD}-\sqrt{d_{GD}^2+\left(\frac{l_x-W_G}{2}\right)^2}}{L_G\tau}}$$

$$\left.T_{GD,f,2} = \frac{l_y-W_D}{2}\mathrm{e}^{\frac{H_D+d_{GD}-\sqrt{d_{GD}^2+\left(\frac{l_y-W_D}{2}\right)^2}}{H_D\tau}}\right]$$

$$C_{GS,f} = \frac{4}{\pi}\varepsilon_0\varepsilon_{GS}\left(W_S \ln\left(\frac{d_1}{d_{GS}}\right) + W_G \ln\left(\frac{d_2}{d_{GS}}\right)\right)$$

$$\begin{bmatrix} d_z = T_{GS,f,z} + \sqrt{d_{GS}^2 + T_{GS,f,z}^2} \\ T_{GS,f,1} = \frac{l_x - W_G}{2}\,\mathrm{e}^{\frac{L_G + d_{GS} - \sqrt{d_{GS}^2 + \left(\frac{l_x - W_G}{2}\right)^2}}{L_G \tau}} \\ T_{GS,f,2} = \frac{l_y - W_S}{2}\,\mathrm{e}^{\frac{H_S + d_{GS} - \sqrt{d_{GS}^2 + \left(\frac{l_y - W_S}{2}\right)^2}}{H_S \tau}} \end{bmatrix}$$

$$C_{GD,f,x} = \frac{4W_D}{\pi}\varepsilon_0\varepsilon_{GD} \ln\left(\frac{g}{d_{GD} + \frac{l_x - W_G}{2}}\right) \quad \text{if } d_{GD} > \frac{l_x - W_G}{2}$$

$$\begin{bmatrix} g = \frac{l_x - W_G}{2} + T_{GD,f,x} + \sqrt{d_{GD}^2 + T_{GD,f,x}^2 + (l_x - W_G)T_{GD,f,x}} \\ T_{GD,f,x} = \left(\frac{l_x}{2} - R - t_{ox}\right)\mathrm{e}^{\frac{L_G + d_{GD} - \sqrt{d_{GD}^2 + \left(\frac{l_x}{2} - R - t_{ox}\right)^2 + \left(\frac{l_x}{2} - R - t_{ox}\right)(l_x - W_G)}}{L_G \tau}} \end{bmatrix}$$

$$C_{GD,f,y} = \frac{4W_G}{\pi}\varepsilon_0\varepsilon_{GD} \ln\left(\frac{h}{d_{GD} + \frac{l_y - W_D}{2}}\right) \quad \text{if } d_{GD} > \frac{l_y - W_D}{2}$$

$$\begin{bmatrix} h = \frac{l_y - W_D}{2} + T_{GD,f,y} + \sqrt{d_{GD}^2 + T_{GD,f,y}^2 + (l_y - W_D)T_{GD,f,y}} \\ T_{GD,f,y} = \left(\frac{l_y}{2} - R - t_{ox}\right)\mathrm{e}^{\frac{H_D + d_{GD} - \sqrt{d_{GD}^2 + \left(\frac{l_y}{2} - R - t_{ox}\right)^2 + \left(\frac{l_y}{2} - R - t_{ox}\right)(l_y - W_D)}}{H_D \tau}} \end{bmatrix}$$

$$C_{GS,f,x} = \frac{4W_S}{\pi}\varepsilon_0\varepsilon_{GS} \ln\left(\frac{i}{d_{GS} + \frac{l_x - W_G}{2}}\right) \quad \text{if } d_{GS} > \frac{l_x - W_G}{2}$$

$$\begin{bmatrix} i = \frac{l_x - W_G}{2} + T_{GS,f,x} + \sqrt{d_{GS}^2 + T_{GS,f,x}^2 + (l_x - W_G)T_{GS,f,x}} \\ T_{GS,fri,x} = \left(\frac{l_x}{2} - R - t_{ox}\right)\mathrm{e}^{\frac{L_G + d_{GS} - \sqrt{d_{GS}^2 + \left(\frac{l_x}{2} - R - t_{ox}\right)^2 + \left(\frac{l_x}{2} - R - t_{ox}\right)(l_x - W_G)}}{L_G \tau}} \end{bmatrix}$$

$$C_{GS,f,y} = \frac{4W_G}{\pi}\varepsilon_0\varepsilon_{GS}\ln\left(\frac{j}{d_{GS}+\frac{l_y-W_S}{2}}\right) \quad \text{if } d_{GS} > \frac{l_y-W_S}{2}$$

$$\left[j = \frac{l_y-W_S}{2} + T_{GS,f,y} + \sqrt{d_{GS}^2 + T_{GS,f,y}^2 + (l_y - W_S)T_{GS,f,y}}\right]$$

$$\left[T_{GS,fri,y} = \left(\frac{l_y}{2} - R - t_{ox}\right)\mathrm{e}^{\frac{H_S+d_{GS}-\sqrt{d_{GS}^2+\left(\frac{l_y}{2}-R-t_{ox}\right)^2+\left(\frac{l_y}{2}-R-t_{ox}\right)(l_y-W_S)}}{H_S\tau}}\right]$$

13.5.4 Array edge capacitances

$$C_{DS,e,y} = \frac{W_D+W_S}{2\pi}\varepsilon_{GS}\ln\left(1+\frac{H_D+H_S}{d_{DS}}\right)$$

$$C_{GD,e,x} = \frac{W_D\varepsilon_0\varepsilon_{GD}}{\pi}\frac{L_G+H_D}{\frac{l_x-W_G}{\alpha_{GD,e,x}}+(L_G+H_D)\left(\ln\left(1+\frac{L_G+H_D}{d_{GD}}\right)+e^{-\frac{l_x-W_G}{6d_{GD}}-\frac{1}{3}}\right)^{-1}}$$

$$\left[\alpha_{GD,e,x} = e^{-\frac{l_x-W_G}{2d_{GD}+L_G+H_D}}\right]$$

$$C_{GD,e,y} = \frac{W_D\varepsilon_0\varepsilon_{GD}}{\pi}\frac{L_G+H_D}{\frac{l_y-W_D}{\alpha_{GD,e,y}}+(L_G+H_D)\left(\ln\left(1+\frac{L_G+H_D}{d_{GD}}\right)+e^{-\frac{l_y-W_D}{6d_{GD}}-\frac{1}{3}}\right)^{-1}}$$

$$\left[\alpha_{GD,e,y} = e^{-\frac{l_y-W_D}{2d_{GD}+L_G+H_D}}\right]$$

$$C_{GS,e,x} = \frac{W_S\varepsilon_0\varepsilon_{GS}}{\pi}\frac{L_G+H_S}{\frac{l_x-W_G}{\alpha_{GS,e,x}}+(L_G+H_S)\left(\ln\left(1+\frac{L_G+H_S}{d_{GS}}\right)+e^{-\frac{l_x-W_G}{6d_{GS}}-\frac{1}{3}}\right)^{-1}}$$

$$\left[\alpha_{GS,e,x} = e^{-\frac{l_x-W_G}{2d_{GS}+L_G+H_S}}\right]$$

$$C_{GS,e,y} = \frac{W_S\varepsilon_0\varepsilon_{GS}}{\pi}\frac{L_G+H_S}{\frac{l_y-W_S}{\alpha_{GS,e,y}}+(L_G+H_S)\left(\ln\left(1+\frac{L_G+H_S}{d_{GS}}\right)+e^{-\frac{l_y-W_S}{6d_{GS}}-\frac{1}{3}}\right)^{-1}}$$

$$\left[\alpha_{GS,e,y} = e^{-\frac{l_y-W_S}{2d_{GS}+L_G+H_S}}\right]$$

13.6 Output-low solution

The two possible solutions of the output-low voltage for the inverter are presented below. Firstly, the pull-down transistor is in the linear region and secondly, the pull-down transistor is in saturation region for a saturated load transistor. The solution is written on the general form, for a solution to a second grade equation.

$$V_{OL} = -\frac{b}{2} - \sqrt{\frac{b^2}{4} - c}$$

$$\begin{bmatrix} b_{LS} = 2V_{DSAT}(V_{DD} - V_T) - N_{load}\, V_{DD}\, \sigma_S(V_{DD} - V_{OL}, V_{DD}) \cdot \frac{R_s}{N_{PDN}} \\ c_{LS} = N_{load}\, V_{DD}\, \sigma_S(V_{DD} - V_{OL}, V_{DD}) \cdot \frac{2(1+R_s k V_{DSAT}\,(V_{DD} - V_T))}{N_{PDN}\, k} \end{bmatrix}$$

$$\begin{bmatrix} b_{SS} = \frac{(N_{PDN} - N_{load})R_s k V_{DSAT}^2\,(V_{DD} - V_T) V_{DSAT}^2\,(V_{DD} - V_{OL})}{2N_{PDN}\, V_{DSAT}^2\,(V_{DD} - V_T) - 2N_{load}\, V_{DSAT}^2\,(V_{DD} - V_{OL})} - V_{DD} \\ c_{SS} = \frac{N_{load}\, V_{DD}\, R_s k V_{DSAT}^2\,(V_{DD} - V_T) V_{DSAT}^2\,(V_{DD} - V_{OL})}{2N_{PDN}\, V_{DSAT}^2\,(V_{DD} - V_T) - 2N_{load}\, V_{DSAT}^2\,(V_{DD} - V_{OL})} \end{bmatrix}$$

Part XIV REFERENCES

[1] Claes Thelander et al., "Development of a Vertical Wrap-Gated InAs FET," *IEEE Transactions on electron devices*, vol. 55, no. 11, pp. 3030-3036, November 2008.

[2] Luryi Choi et al., "Extracting Mobility Degradation and Total Series Resistance of Cylindrical Gate-All-Around Silicon Nanowire Field-Effect Transistors," *IEEE Electron Device Letters*, vol. 30, no. 6, pp. 665-667, June 2009.

[3] Benjamin Iñíguez et al., "Explicit Continuous Model for Long-Channel Undoped Surrounding Gate MOSFETs," *IEEE Transactions on Electronic Devices*, vol. 52, no. 8, pp. 1868-1873, August 2005.

[4] E. Gnani, S. Reggiani, S. Rudan, and G. Baccarani, "A New Approach to the Self-consistent Solution of the Schrödinger-Poisson Equations in Nanowire MOSFETs," in *Solid-State Device Research conference*, Bologna, Italy, September 2004, pp. 177-180.

[5] Bipul C. Paul et al., "An Analytical Compact Circuit Model for Nanowire FET," *IEEE Transactions on Electron Devices*, vol. 54, no. 7, pp. 1637-1644, July 2007.

[6] Simon Min Sze, *Semiconductor Devices: Physics and Technology*, 2nd ed. New York, New York, United States of America: John Wiley & Sons, inc., 1985.

[7] Betty Lise Anderson and Richard L. Anderson, *Fundamentals of Semiconductor Devices*. New York, New York, United States of America: McGraw-Hill Professional, 2005.

[8] Aditya Bansal, Bipul C. Paul, and Kaushik Roy, "An analytical fringe capacitance model for interconnects using conformal mapping," *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, pp. 2765-2774, April 2006.

[9] Narain Arora, *MOSFET modeling for VLSI simulation: Theory and practice*,

Chih-Tang Sah, Ed. Danvers, Massachusetts, United States of America: World Scientific Publishing Co. Pte. Ltd., 2007.

[10] Jan M. Rabaey, Anantha Chandrakasan, and Borivoje Nikolić, *Digital Integrated Circuits: A Design Perspective*, 2nd ed. New Jersey, New Jersey, Unites States of America: Prentice Hall, Pearson Education Inc., 2003.

[11] William B. Shockley, "A Unipolar "Field-Effect" Transistor," *Proceedings of the IRE*, vol. 40, no. 11, pp. 1365-1376, 1952.

[12] C. T. Sah, "Characteristics of the metal-oxide-semiconductor transistors," *IEEE Transactions on Electronic Devices*, vol. 11, no. 7, pp. 324-345, 1964.

[13] X. Jiang, A. P. Jayasumana, and S. Chiao, "A proper deep submicron MOSFET model (PDSMM) and its applications for delay modeling of CMOS inverters," in *6th International Conference on Solid-State and Integrated-Circuit Technology*, vol. 2, 2001, pp. 875-878.

[14] Harold Shichman and David A. Hodges, "Modeling and simulation of insulated gate field-effect transistor switching circuits," *IEEE Journal of Solid-State Circuits*, vol. 3, no. 3, pp. 285-289, 1968.

[15] Carlos Galup-Montoro and Márcio Cherem Schneider, *MOSFET modeling for circuit analysis and design*. Singapore, Republic of Singapore: World Scientific Publishing Co. Pte. Ltd., 2007.

[16] Narain Arora, *MOSFET modeling for VLSI simulation: Theory and practice*, Chih-Tang Sah, Ed. Danvers, Massachusetts, United States of America: World Scientific Publishing Co. Pte. Ltd., 2007.

[17] Erik Lind, Martin P Persson, Yann-Michel Niquet, and Lars-Erik Wernersson, "Band Structure Effects on the Scaling Properties of [111] InAs Nanowire MOSFETs," *IEEE Transactions on Electron Devices*, vol. 56, pp. 201-205, February 2009.

[18] Adelmo Ortiz-Conde, Francisco J. Garzía Sánchez, and Juin J. Liou, "On the Extraction of Threshold Voltage, Effective Channel Length and Series

Resistance of MOSFETs (Invited)," in *Diagnostics and Yield Symposium*, vol. 39, Warsaw, 2000, pp. 1-17.

[19] Jen Shuang Wong, Jian Guo Ma, Kiat Seng Yeo, and Manh Anh Do, "A New Approach for the Extraction of Threshold Voltage for MOSFET's," in *Technical Proceedings of the 2001 International Conference on Modeling and Simulation of Microsystems*, vol. 1, Hilton Head Island, South Carolina, 2001, pp. 534-537.

[20] Lidia Dobrescu, M. Petrov, D. Dobrescu, and C. Ravariu, "Threshold voltage extraction methods for MOS transistors," in *International Semiconductor Conference*, vol. 1 & 2, 2000, pp. 371-374.

[21] Sung-Mo Kang and Yusuf Leblebici, *CMOS Digital Integrated Circuits: Analysis and Design*, 3rd ed. New York, New York, United States of America: McGraw-Hill Professional, 2003.

[22] Laszlo B. Kish, "End of Moore's law: thermal (noise) death of integration in micro and nano electronics," *Physics Letters A*, vol. 305, no. 3-4, pp. 144-149, December 2002.

[23] Robert Chau et al., "Benchmarking nanotechnology for high-performance and low-power logic transistor applications," *IEEE Transactions on Nanotechnology*, vol. 4, pp. 153-158 , March 2005.

Part XV INDEX

15.1 Figure index

15.2 Table index