This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 318458.

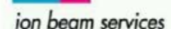
ICT Project no. 318458

SUPERTHEME

Circuit Stability Under Process Variability and Electro-Thermal-Mechanical Coupling

D6.5: Final report on More Moore simulation and support provided to the ITRS

	Name	**Organisation**	**Date**
Edited	J. Lorenz	Fraunhofer IISB	January 15, 2016
Reviewed	P. Ellinghaus, M. Nedjalkov	TU Wien	January 20, 2015
Final approval	J.M. Park	Ams AG	January 25, 2015



Contents

Abstract 3

1 Introduction 3

2 Selected Results Obtained 4

2.1 Outline of Software Development Results 4

Integrated process and equipment simulation 5

Extension of statistical device simulator GARAND 6

Process-aware compact models 7

2.2 Simulation Support Provided to the ITRS 8

2.3 More Moore Application Examples 9

Combined simulation of systematic and stochastic variations 9

Impact of correlation of variations for LELE Double Patterning 10

Impact of correlation of variations for FinFET based SRAM 13

Impact of self-heating on statistical variability 15

3 Inventory of Published Results 17

3.1 More Moore Publications in 2013 17

3.2 More Moore Publications in 2014 17

3.3 More Moore Publications in 2015 18

3.4 More Moore Publications in 2016 18

4 Conclusions 18

References 18

Abstract

This document gives an overview of results from the SUPERTHEME project obtained in the *More Moore* area. The initial objective was to focus this part of the project activities on support to the International Technology Roadmap for Semiconductor ITRS. Due to later changes in the scope of the ITRS the capabilities of the SUPERTHEME software were then demonstrated on state-of-the-art planar and FinFET architectures, simulating the impact of selected systematic and stochastic process variations on transistors and SRAM circuits. Besides the development, enhancement and integration of simulation modules a key result has been the development and demonstration of process and correlation aware compact models. This document includes four simulation examples which demonstrate some of the new capabilities implemented within SUPERTHEME: The simultaneous simulation of the impact of systematic and stochastic process variations; the impact of correlations resulting from Litho-Etch-Litho-Etch double patterning on an SRAM circuit; the impact of correlations of variations for FinFET based SRAM; and the impact of self-heating on statistical variability.

1 Introduction

Among the physical limitations which challenge progress in nanoelectronics for aggressively scaled *More Moore*, *Beyond CMOS* and advanced *More-than-Moore* applications, process variability and the interactions between and with electrical, thermal and mechanical effects have got more and more critical. Effects from various sources of process variations, both systematic and stochastic, influence each other and lead to variations of the electrical, thermal and mechanical behavior of devices, interconnects and circuits. Correlations are of key importance because they drastically affect the percentage of products which meet the specifications. Whereas the comprehensive experimental investigation of these effects is largely impossible, modeling and simulation (TCAD) offers the unique possibility to predefine process variations and trace their effects on subsequent process steps and on devices and circuits fabricated, just by changing the corresponding input data. This important requirement for and capability of simulation is among others highlighted in the International Technology Roadmap for Semiconductors ITRS.

Within the SUPERTHEME project, the most important weaknesses which limit the use of current TCAD software to study the influence of both systematic and stochastic process variability and its interaction with electro-thermal-mechanical effects have been removed, and the study of correlations has been enabled. The project has efficiently combined the use of commercially available software and leading-edge background results of the consortium with the implementation of the key missing elements and links. It has bridged the critical gap between variability simulation on process and device/interconnect level, and has included the treatment of correlations. The capabilities of the software system have been demonstrated both on advanced analog circuits and on aggressively scaled transistors.

Whereas due to the Call for Proposals addressed by SUPERTHEME benchmarks conducted in the project focused on advanced *More-than-Moore* applications, the software development and integration results obtained are both important for advanced analog applications, where especially matching plays a key role, and for aggressively scaled digital applications. Concerning these latter *More Moore* applications, within SUPERTHEME some simulation studies were planned in support of the International Technology Roadmap for Semiconductors ITRS. Due to changes in the focus of the ITRS which occurred during the SUPERTHEME project period the *More Moore* simulation studies were then shifted to address the variability of state-of-the-art planar and FinFET transistors, especially focusing on variation-aware compact models.

In the following, some exemplary results in the *More Moore* area are sketched. Finally, an orientation on publications from the project is given.

2 Selected Results Obtained

2.1 Outline of Software Development Results

In the following, a small set of selected examples for software development results required especially for *More Moore* applications is shown, whereas a full description of the advancements achieved is neither possible in a public document nor within the space available. These software results largely enabled the application examples summarized below in section 2.3.

The impact of stochastic variations such as Random Dopant Fluctuations (RDF), Line Edge Roughness (LER) or Metal Gate Granularity (MGG) on the properties of devices and circuits was broadly discussed in many publications from various groups including especially the SUPERTHEME partners GSS and GU. Here, the source of variations is largely the granularity of matter, including the problem that for very small devices only a few ions are implanted into the channel regions. In order to efficiently simulate the impact of these variations on device properties the GSS atomistic device simulator GARAND [1] uses tensor-based meshes. In addition, due to shrinking feature sizes and ever shallower dopant profiles device geometries and dopant distributions have also been increasingly affected by systematic variations of patterning or doping processes, because these have in several respects approached physical limits: E.g. the smallest wavelength used in optical lithography is 193 nm. For dry lithography with optimum (very large) lenses this would allow the printing of dense lines down to about 100 nm pitch or about 50 nm feature size. Current aggressively scaled transistors with gate lengths of 30 nm and below are only possible with technological tricks like immersion lithography and/or double patterning, or computational lithography. With the nominal processes having approached physical limits, variations of the processes have become more severe, and critically affect key device parameters such as transistor gate length. In turn, process variations must be traced from their source at equipment level through device geometry and dopant distributions up to their effects on transistor and circuit performance. Figure 1 shows the various levels of variations addressed within SUPERTHEME. In addition to software developed by SUPERTHEME partners - especially for topography and device simulation as well as for compact model extraction - SUPERTHEME employs external tools for equipment simulation and parts of Sentaurus Process for process simulation.

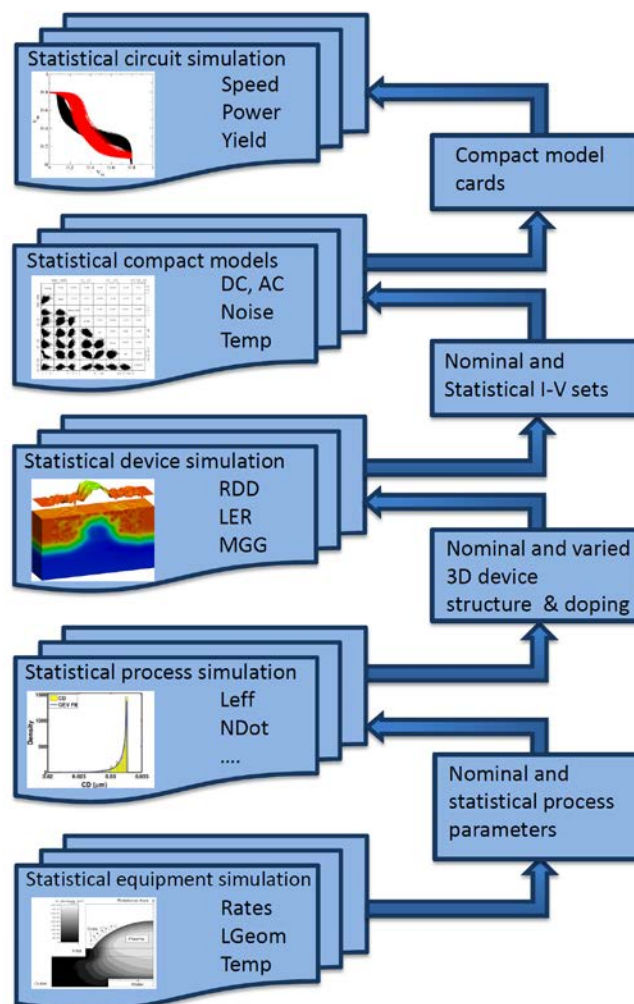


Fig.1: Impacts of variations to be considered at various levels of simulation.

Integrated process and equipment simulation

In order to be able to trace the impact of systematic variations caused by the process equipment, a close coupling between equipment and process simulation is important. This is state-of-the-art in current advanced lithography simulators such as Dr.LiTHO [2] from the SUPERTHEME partner Fraunhofer IISB: Here, equipment properties such as details of the illumination source, the mask and the imaging system are considered in the simulations of the distortions of patterns transferred from the mask into the photoresist. In turn, also the change of resist patterns due to inevitable variations e.g. of the distance between the imaging system and the resist (the so-called focus) can be simulated. Figure 2 shows the standard so-called Bossung plot where the target feature size is printed depending on focus and illumination dose threshold: Whereas the central dashed line represents the combinations of focus and dose threshold which result in printing the nominal feature sizes (“critical dimensions” CD), both outer dashed lines represent combinations of focus and dose threshold which lead to a 10% higher CD.

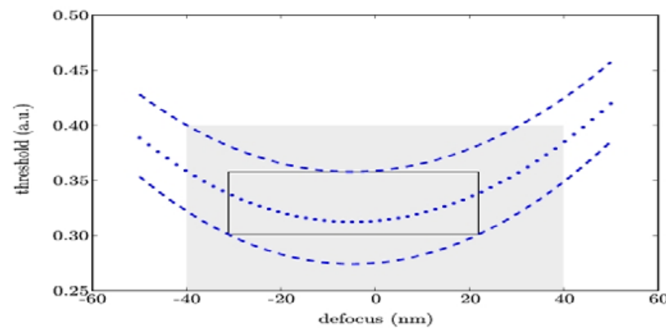


Fig. 2: Typical process window in optical lithography.

In order to treat process variations in deposition and etching, a close link between reactor-scale equipment and feature-scale process simulation must be implemented. Due to the large variety of process equipment used in semiconductor industry a generic integration is not possible, but the link can only be demonstrated for exemplary cases. Within SUPERTHEME the computational fluid dynamics program ESI CFD ACE [3] and the plasma simulation program Q-VT [4] were linked with the feature-scale deposition and etching simulators DEP3D and ANETCH of Fraunhofer IISB. These combinations were then used to simulate and study the impact of various etching, deposition and plasma oxidation equipment on the structures generated. As an example Figure 3 shows the setup for equipment simulation of plasma-enhanced chemical vapor deposition (PECVD) of SiO_2 using CFD ACE, whereas Figure 4 shows exemplarily the flux consumed at the surface and the SiO_2 deposition rate at all surface positions.

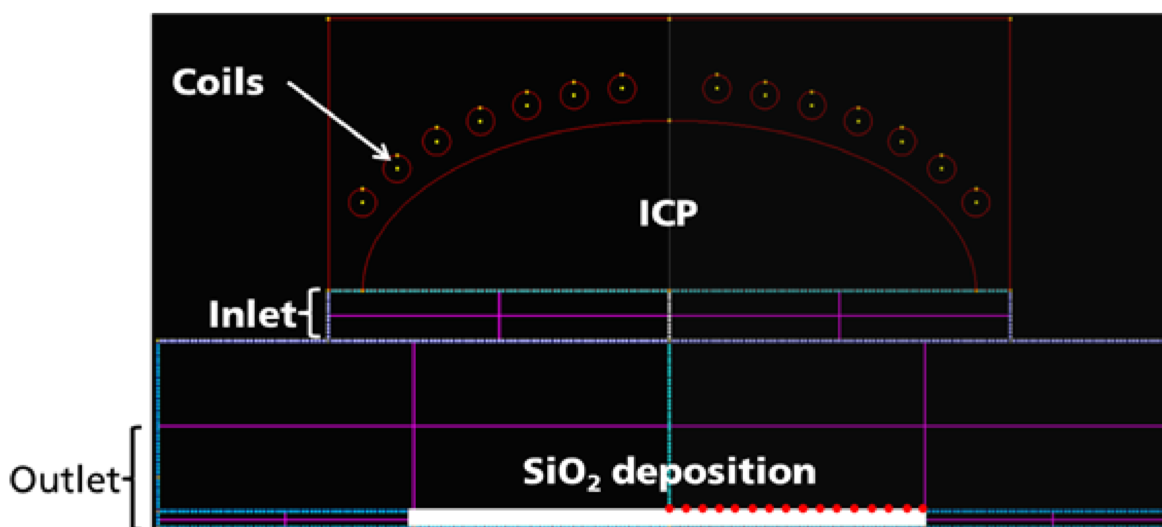


Fig. 3: Reactor geometry used for the simulation of SiO_2 deposition on equipment level.

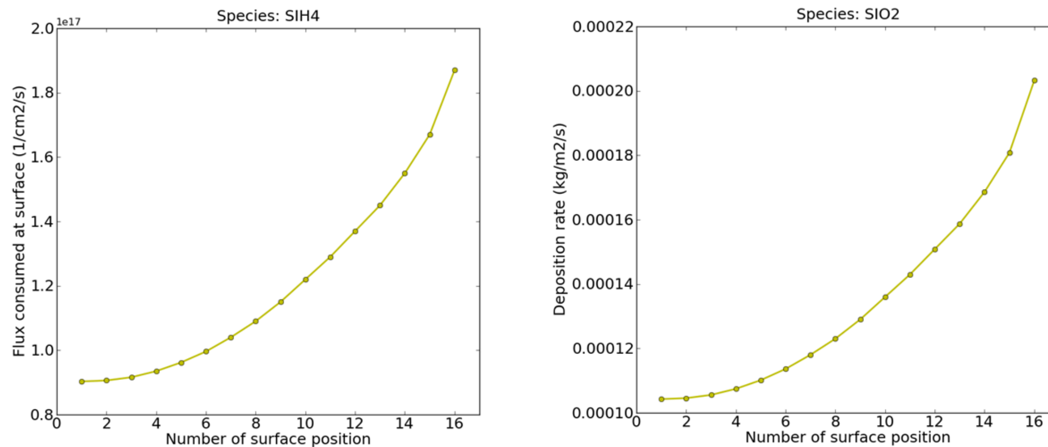


Fig.4: Simulation of SiO_2 deposition for the reactor described in Figure 2. SiH_4 flux consumed at surface (left), SiO_2 deposition rate (right) which originates from SiH_4 and SiH_2 contributions. The surface positions 1 to 16 correspond to the red circles in Figure 2.

In another SUPERTHEME activity the feature-scale topography simulation tools from Fraunhofer IISB and TU Wien were linked with the industrial standard process simulator Sentaurus Process [5] via writing the simulated geometries in DF-ISE file format [6].

Extension of statistical device simulator GARAND

An important software result obtained within SUPERTHEME was to enable GARAND to read and simulate devices with non-rectangular shapes resulting from real patterning steps.

In the version of the GSS atomistic device simulation tool, GARAND, available at the beginning of the SUPERTHEME project 3D statistical variability simulations were carried out for 'template' transistors with fixed, typically idealised rectangular 3D geometry and independently of the systematic variability simulations. Real patterning steps, however, lead to non-rectangular shapes, e.g. with rounded corners and sloped sidewalls, where the distance between edges and/or the shape may vary due to process variations. In order to enable GARAND to model non-ideal 3D geometries coming from process simulation enhanced capabilities were implemented for handling arbitrary 3D device geometries arising from complex patterning steps. This was done via transferring structures and other data described in the DF-ISE format, which is a commonly used format for the output of process simulation results. The structures defined in DF-ISE format are imported in GARAND from an intermediary mesh created by the structure translator MONOLITH [7]. Extensive code development in GARAND has been carried out, which allows GARAND to import an arbitrary device structure defined initially in a DF-ISE file and translated by MONOLITH in an intermediate file format. Various further improvements were made e.g. to enable the superposition of sources of statistical and systematic variability and to assure that the results obtained are independent of the final mesh used, provided of course that the mesh is fine enough. Furthermore, an optimised import process has been worked out for the DF-ISE structures produced by Sentaurus and the Fraunhofer tools. Figure 5 shows the basic flow of this transfer of arbitrary device geometries and dopant distributions into GARAND. This extension of GARAND has especially been described in a paper published at ULIS 2014 [8].

Important enhancements to GARAND were not only done with regard to algorithm and interfaces, but also concerning the implementation of physical models. This includes especially the implementation of a Wigner approach for the simulation of charge trapping in gate dielec-

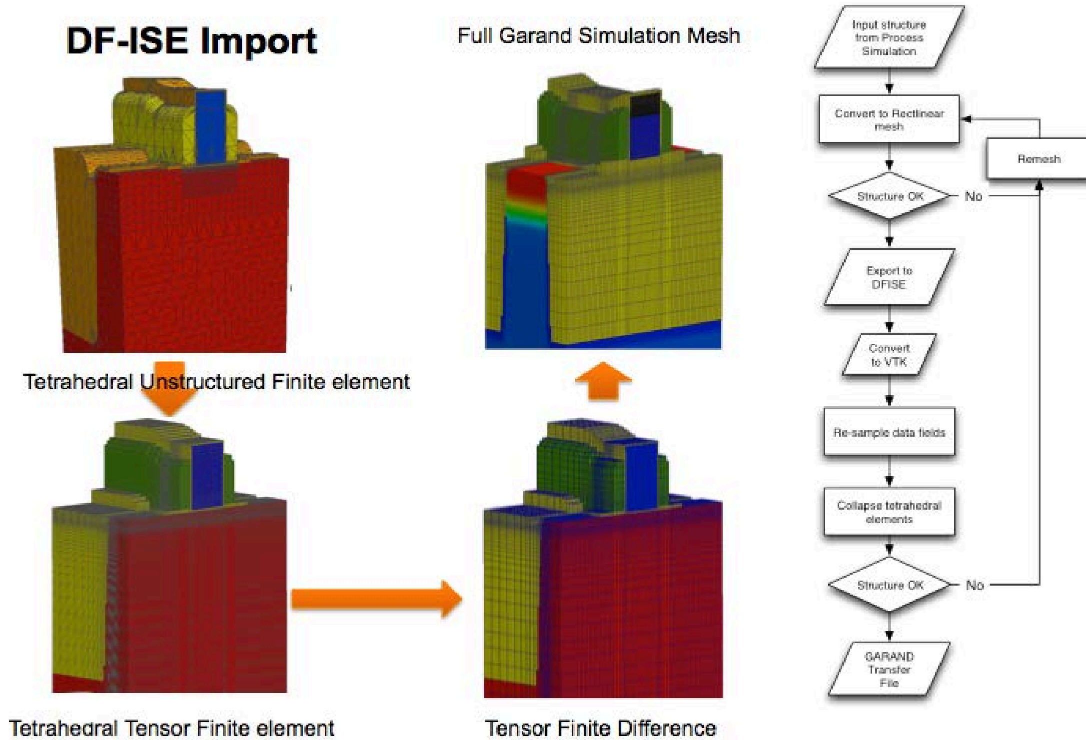


Fig. 5: Schematic description of the process of mesh conversion from Sentaurus process to GARAND.

trics [9] and the solution of the coupled Heat Flow, Poisson, and Current Continuity Equations for coupled electro-thermal simulations [10].

Process-aware compact models

The simulation sequence outlined above allows for the simulation of the impact of variations from equipment up the device level. In order to utilize the results of variability on circuit level, process-aware compact models are required.

In turn, within SUPERTHEME a hierarchical compact model extraction strategy was developed. It is based on BSIM4 compact models [11] and employs the GSS compact model extractor Mystic [12]. In short this approach can be summarized as follows:

- Nominal compact models are extracted for different geometrical device parameters, e.g. gate length L and width W of a planar transistor.

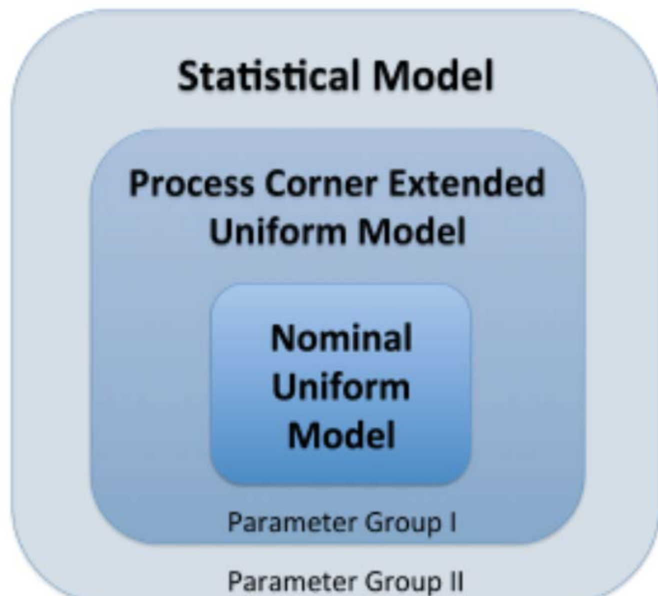


Fig. 6: Schematic view of the unified compact modelling strategy. It features nominal uniform model, and two groups of compact model parameters, and enhanced statistical models.

- A process corner extended compact model is extracted which additionally describes the dependence of electrical device parameters on the geometrical parameters considered. This results in so-called “group 1” compact model parameters.
- Finally, so-called “group 2” parameters are extracted to capture the impact of the statistical variations considered (RDF, LER, MGG) on the device characteristics.

This approach is illustrated in Figure 6 and is described more in detail in one of the papers which resulted from SUPERTHEME [13].

2.2 Simulation Support Provided to the ITRS

The simulation support provided by SUPERTHEME to the ITRS was described in the specific public deliverable D6.2. In the following, the main results are briefly summarized, whereas more information is given in the preceding deliverable.

Fraunhofer IISB calibrated Sentaurus Device [14] from Synopsys using experimental data for 20 nm CMOS transistors published by Samsung [15]. The following approach was then used for the assessment of the scaling of nominal bulk, single-gate (SG) fully-depleted (FD) SOI, and multigate (MG) FD NMOS FETs:

- Geometrical sizes and channel doping of the transistors were used as specified in the 2011 issue of the ITRS [16].
- For simplicity, a uniform doping of $3 \cdot 10^{20}$ cm^{-3} was used for the source and drain regions, which were aligned to the edges of the gate with zero overlap.
- Appropriate physical models were selected, as detailed elsewhere [17].

Different scaling scenarios were considered and simulated for all three device architectures in question and compared with the 2011 ITRS specifications for the device performance, namely:

- For bulk NMOS: a) Standard ITRS scaling; b) channel doping fixed to $5 \cdot 10^{18}$ cm^{-3} instead of increasing from $4.5 \cdot 10^{18}$ cm^{-3} in 2011 to $9 \cdot 10^{18}$ cm^{-3} in 2017; c) in addition to the fixed channel doping a scaling of *EOT* more aggressive than foreseen in the ITRS was assumed. This latter scaling plan led to a good agreement with the ITRS specifications for the drain current. Furthermore, with this scaling plan especially for smaller channel lengths significantly lower DIBL and inverse subthreshold swing were predicted.
- For SG FD SOI NMOS: a) Standard ITRS scaling; b) channel doping fixed to $1 \cdot 10^{17}$ cm^{-3} instead of increasing from $6 \cdot 10^{18}$ cm^{-3} in 2013 to $9 \cdot 10^{18}$ cm^{-3} in 2017 and then decreasing to $1 \cdot 10^{17}$ cm^{-3} in 2019; c) channel doping fixed to $1 \cdot 10^{15}$ cm^{-3}; d) in addition to a fixed channel doping of $2.5 \cdot 10^{18}$ cm^{-3} a scaling of *EOT* more aggressive than foreseen in the ITRS was assumed. This latter scaling plan led to a good agreement with the ITRS specifications for the drain current. Furthermore, with this scaling plan DIBL stayed below the tolerable value of 100 mV/V, and the inverse subthreshold swing for smaller channel lengths gets somewhat lower that specified by the ITRS.
- For DG FD SOI NMOS: a) Standard ITRS scaling; b) silicon body thickness adjusted to 40% of the channel length, instead of using more than twice the value for the single-date fully depleted SOI transistor; c) additionally, *EOT* was adapted (higher than ITRS specifications at large channel lengths, lower at small channel lengths), see Figure 7. This latter scaling plan led to a good agreement with the ITRS specifications for the drain current, see Figure 8. Furthermore, with this scaling plan DIBL was strongly reduced and stayed significantly below the tolerable value of 100 mV/V. Also the inverse subthreshold swing was strongly reduced.
- In the 2011 ITRS bulk NMOS was specified for channel lengths down to 14 nm, and SG FD SOI NMOS down to 11.7. For these values the drain current of both architectures was simulated to be rather similar to that of the DG FD SOI NMOS.

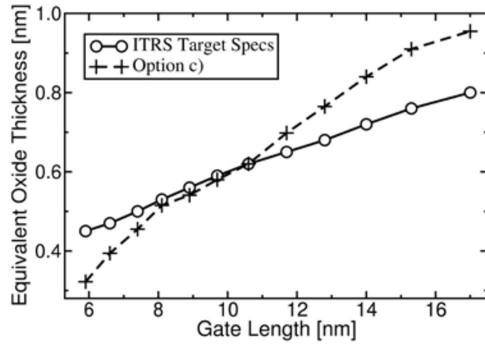


Fig. 7: Equivalent oxide thickness *EOT* as used in the scaling option c) for the DG FD SOI NMOS transistor, compared with original specifications for *EOT* from the 2011 ITRS [17].

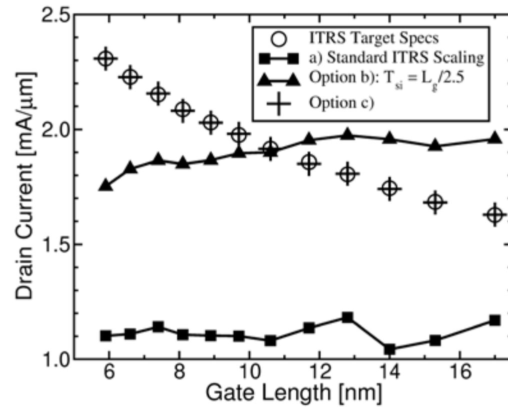


Fig. 8: Drain current specifications for the DG FD SOI NMOS transistor compared with simulation results using a) original ITRS scaling specifications; b) silicon body thickness adjusted to 40% of L_g; c) additionally, *EOT* adjusted as shown in Fig. 7 [17].

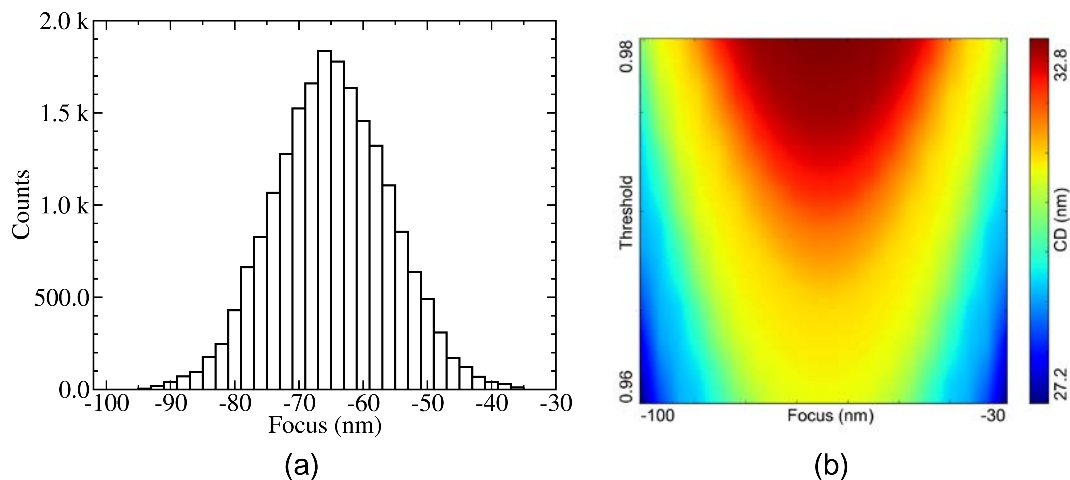
The simulation study performed in this work has demonstrated that all three MOSFET architectures suggested by the ITRS have the potential to achieve the electrical performance specified by the ITRS for the channel lengths in question, however with significantly modified scaling of *EOT*, channel doping, and/or silicon body thickness. More details of the work are given in a paper [17].

2.3 More Moore Application Examples

In the following, some key *More Moore* application results from the SUPERTHEME project are summarized, with more details being given in the SUPERTHEME publications referenced.

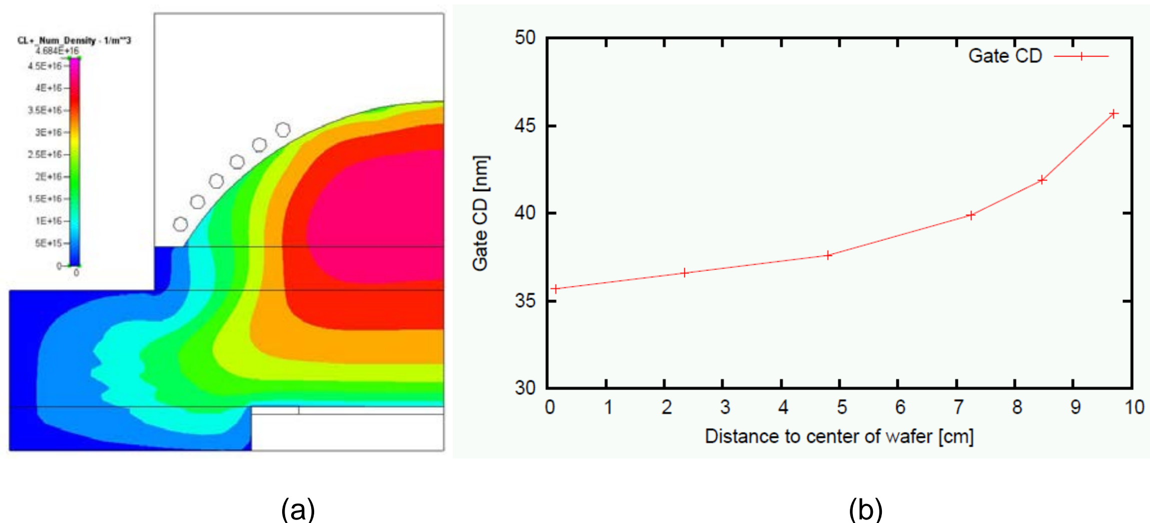
Combined simulation of systematic and stochastic variations

As outlined above in section 2.1, a key innovation within SUPERTHEME has been the rigorous simulation of the combined impact of systematic and stochastic variations. A first example for this was presented at SISPAD 2014, see reference [S6] in section 3 below. Here, the combined impacts of Gaussian variations of focus and dose threshold in optical lithography (see Figure 9), position on the wafer during polysilicon etching in an inductively coupled plasma (ICP) reactor (see Figure 10), and stochastic variations (RDF) are simulated. Figure 11 shows the combined impact of these variations on the threshold voltage of bulk transistors with a nominal gate length of 23.5 nm and a nominal gate width of 33 nm. Both the variations of the gate length caused by focus and dose threshold variations in lithography and the Random Dopant Fluctuations (RDF) lead to relatively broad distributions of the threshold voltage (and of course also other device properties). The dependence of the etch bias on the position on the wafer modifies and shifts this distribution for different positions on the wafer. In turn, the overall distribution of the threshold voltage for the whole wafer is broader than the distributions for a fixed etch bias. The overall distribution is most strongly affected by the distribution for the transistors near the edge of the wafer, because the number of transistors with a fixed distance from the wafer center scales with the square of this distance, and the etch bias depends on that distance.



(a) (b)

Fig. 9: (a) Assumed statistical distribution of defocus (similar for threshold); (b) channel length L(P) (colors) from process simulation, depending on the parameters P defocus and intensity threshold [S6].



(a) (b)

Fig. 10: (a) Simulated density of the Cl^{+} ions in an ICP reactor for a rotationally symmetric geometry. The rotation axis is on the right boundary; (b) simulated dependence of the gate CD value on the position on the wafer [S6].

Impact of correlation of variations for LELE Double Patterning

A first example for the impact of correlations was presented at SISPAD 2013 [S2]. Here, an SRAM circuit consisting of bulk transistors with nominal gate lengths and widths of 20 nm for the inner flip-flop and 25 nm for the access transistors was considered. Due to limitations of optical lithography such small features must be fabricated with Double Patterning. For the example studied we assumed the Litho-Etch-Litho-Etch (LELE) variant of Double Patterning. Unlike usual lithography, which uses a single mask, Double Patterning splits a critical mask level into two incremental masks, which are exposed separately in two incremental lithography steps. Figure 12 shows the SRAM cell after simulation of double patterning for fixed values of dose and focus. The black features are the active silicon areas, the dark gray features are the polysilicon areas generated with the first lithography step of the Double Patterning

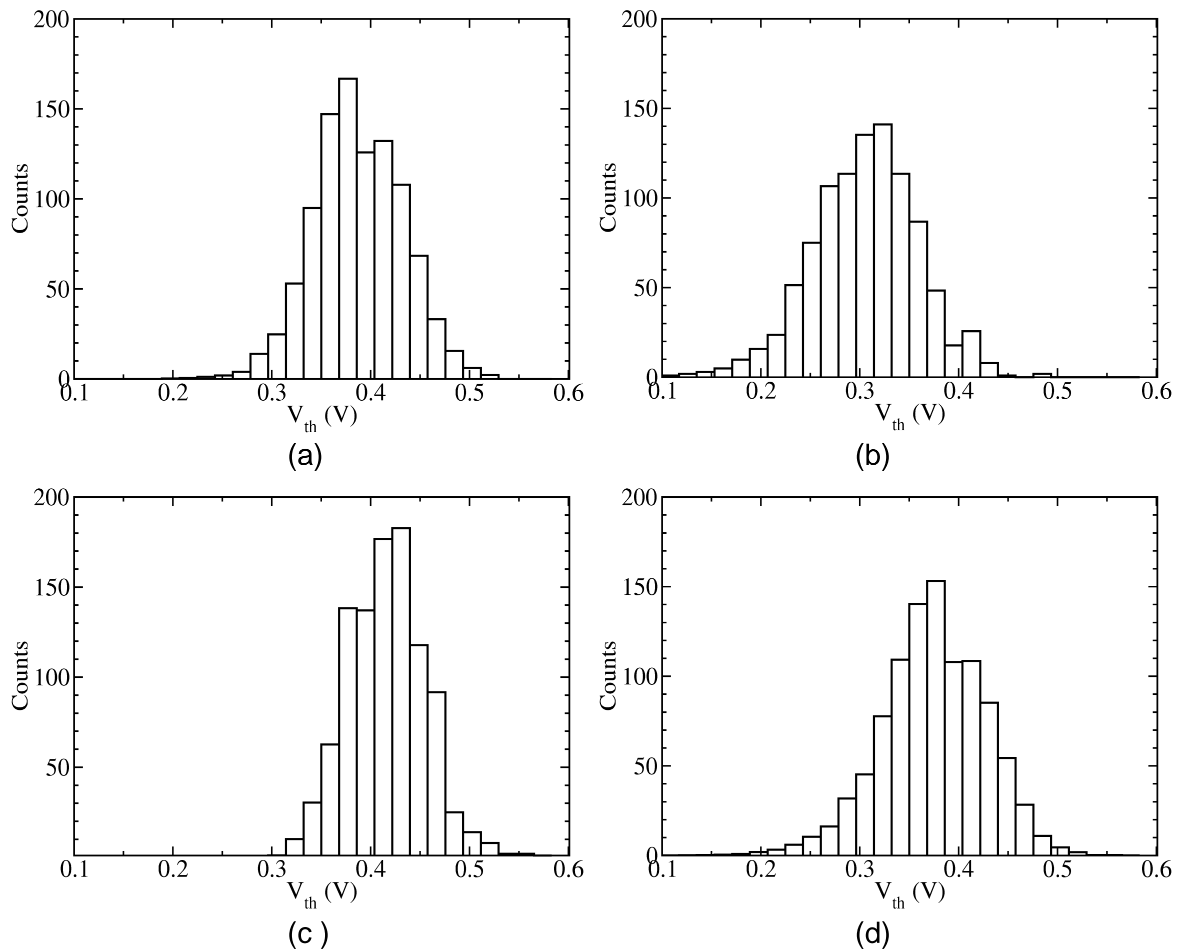


Fig. 11: Overall distribution of threshold voltage at a drain voltage of 0.05 V caused by process and stochastic variations. In all cases, variation of lithography focus and dose is assumed as shown in Fig. 9, and RDF (a): Transistors with nominal etch bias of 5 nm; (b) transistors at center of wafer (etch bias 10 nm); (c) transistors near edge of wafer (etch bias 0 nm); (d) overall distribution of all transistors on the wafer taking variation of etch bias into account [S6].

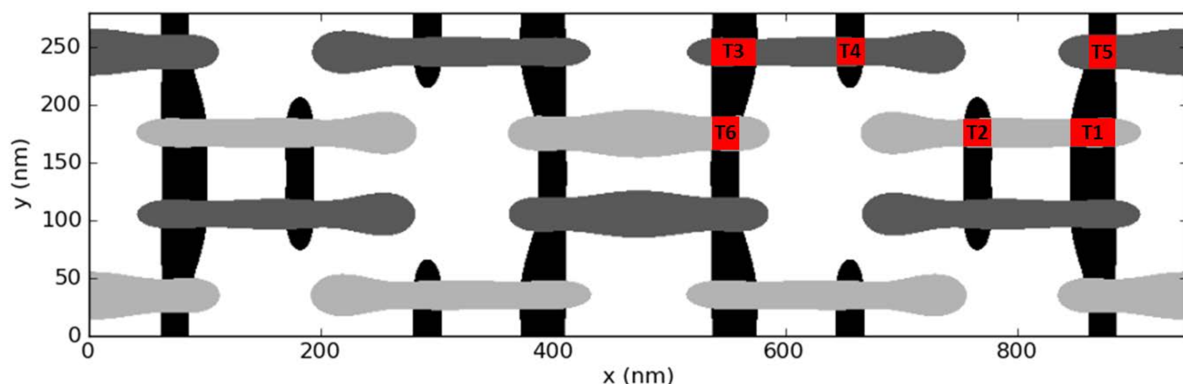


Fig. 12: Final result of the lithography simulation of the 6T SRAM cell exemplarily for one focus / threshold position. The black features are the active silicon areas, the dark gray features are the polysilicon areas generated with the first lithography step of the double patterning process and the light gray features are the polysilicon areas generated with the second lithography step of the double patterning process [S2].

process and the light gray features are the polysilicon areas generated with the second lithography step of the double patterning process. Due to periodic boundary conditions four 6T SRAM cells have to be simulated. T1 / T2 / T3 / T4 are the flip-flop transistors, T5 / T6 are the access transistors. Due to the two lithography steps of the double patterning process the focus / threshold variations inside the two groups T3 / T4 / T5 and T6 / T2 / T1 are the same but between the two groups the variations are independent from each other.

Assuming Gaussian distributions for the variations of focus and dose threshold in both incremental lithography steps the gate length distributions of the six transistors have two distinct properties: First, they are different from each other, both due to the two incremental lithography steps used and due to their different positions in the cell layout which due to the proximity effect influence the feature sizes printed in the lithography process. Second and especially important, due to the division of the cell layout into two incremental steps for the patterning of the gate structures, within each of the two transistor triples (T1, T2, T6) and (T3, T4, T5) the variations of the gate lengths correlate, but gate length do not correlate between the two triples. Besides this, the layout used also leads to different channel width variations for the six transistors, although the channel width is defined by a single exposure lithography step.

In order study the effect of these correlations on the SRAM cell we first set the defocus and the intensity threshold to their minimum and maximum values and observed the resulting change of the SRAM performance parameters [S2]. Figure 13 shows the results of such sensitivity analysis for two variants of correlated variations. In the first variant shown in Figure 13 (left), defocus and intensity threshold were set simultaneously in all lithography levels to their minimum and then to their maximum values. In Figure 13 (right), defocus and intensity threshold were set in anti-phase in the two lithography levels for polysilicon structuring, i. e. in Var F1 defocus was at its minimum in the first illumination but at its maximum in the second illumination; in Var F2 defocus was at its maximum in the first illumination but at its minimum in the second illumination; in Var D1 the intensity threshold was at its minimum in the first illumination but at its maximum in the second illumination; and in Var D2 the intensity threshold was at its maximum in the first illumination but at its minimum in the second illumination. The results of such sensitivity analysis are shown in Fig. 13. If defocus is set synchronously to its minimum a lower Read static noise margin (SNM) than in nominal setting is obtained. On the other hand, if defocus in Var F1 was at its minimum in the first illumination but at its maximum in the second illumination, we have very little difference in Read SNM in comparison to the optimum setting of defocus and intensity threshold in the middle of the process window (Nominal). In contrast, in Var F2, when defocus was at its maximum in the first illumination but at its minimum in the second illumination the value of the Read SNM has a value much lower than nominal SNM or SNM in Var F1. A synchronous variation of the

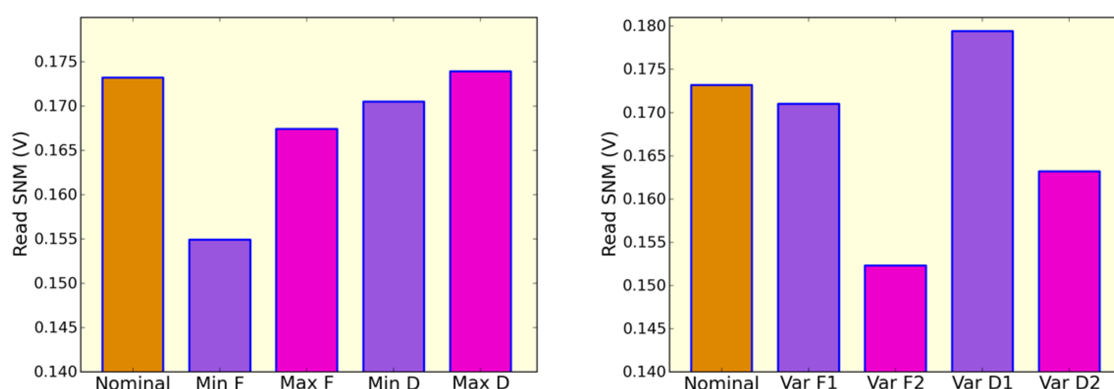


Fig. 13: Read Static Noise Margin variations of the SRAM circuit from minimum/maximum sensitivity analysis [S2].

intensity threshold leads to moderate change of Read SNM, while Var D1 variation results in an SNM significantly higher than the nominal and Var D2 leads to SNM significantly lower than the nominal. These correlations are further discussed in the literature [S2].

Impact of correlation of variations for FinFET based SRAM

Further shrinking of transistors has led to the introduction of transistor architectures which employ buried oxide layers for the prevention of parasitic current and improvement of switching behavior. These so-called "Silicon-on-Insulator" (SOI) technologies have two major variants: First, the extension of planar device architectures with SOI substrates and potentially a second (buried) gate; second and most promising, the replacement of bulk silicon or of a thin silicon layer in case of SOI technology by just a narrow bar of semiconducting material, with the gate electrode more or less wrapped around. This leads to a three-dimensional architecture and includes several variants of so-called FinFETs or nanowires.

FinFETs are typically generated by etching narrow bars from thin SOI layers, and placing the gate on both sides and optionally in addition also on top of the fins. Optimum switching behavior is achieved by using undoped channels and especially by reducing one lateral dimension of the fin (typically the width) to much less (about one third) of the channel length.

In turn, the most critical feature size for a FinFET transistor is the fin width: E.g. for a 14 nm node FinFET transistor the gate length is about 20 nm and the fin width about 10 nm [18]. The standard technology to pattern these FinFET transistors employs Self-Aligned Double Patterning (SADP) for the fin, and then LELE for gate formation. In the SADP step, first a core fin is structured by immersion optical lithography. Then spacers are formed by deposition on these hard masks and subsequent back etching.

For the example studied within SUPERTHEME, the fin layout shown in Figure 14 has been addressed. SADP based on 193 nm optical immersion (water) lithography has been employed, as illustrated in Figure 15. Patterning of the carbon fin core by lithography has been simulated with Dr.LiTHO [2]. Subsequent chemical vapor deposition (CVD) of an oxide layer has been simulated with DEP3D [19], followed by anisotropic back etching using a dry etching process, which was emulated as a perfectly anisotropic one using Sentaurus Process [5]. This finally results in spacers which exhibit a doubled pattern density compared with the initial hard mask. Using further etching steps, the pattern defined by the spacers is then transferred to the underlying layers and eventually to the silicon leading to the fin structures. This means that the bottom widths of the spacers generated by the SADP technique define the dimensions of the mask for creating the fin structures by dry etching of the silicon material. The result is shown in Figure 15 (right). A cross section of this structure is shown in Figure 16 for a CD of 35 nm of the initial carbon lines generated by lithography, which is the nominal value. It can be seen that the width of the inner spacers is smaller than the width of the outer spacers which is a result of the physics of the oxide deposition process: Due to shadowing effects in CVD the thickness of the layer between the lines is smaller than the thickness in the outer regions of the carbon lines. Because the outer lines are not affected by shadowing from the fin cores, the width of the outer lines is constant, independent of the variations of the pitch of the core fins. However, variations of the core fin pitch change the shadowing and therefore lead

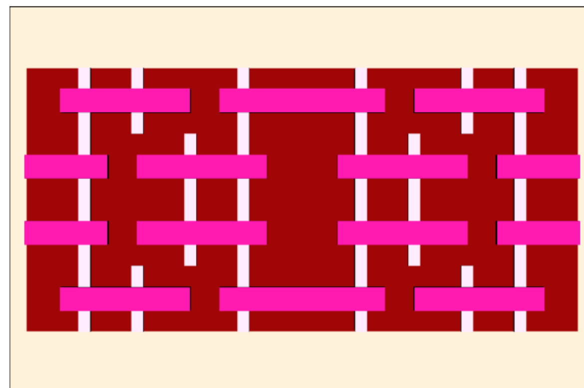


Fig. 14: Layout of silicon fins (vertical lines) and gate lines (horizontal lines) in a 4-cell fragment of an SRAM cell field

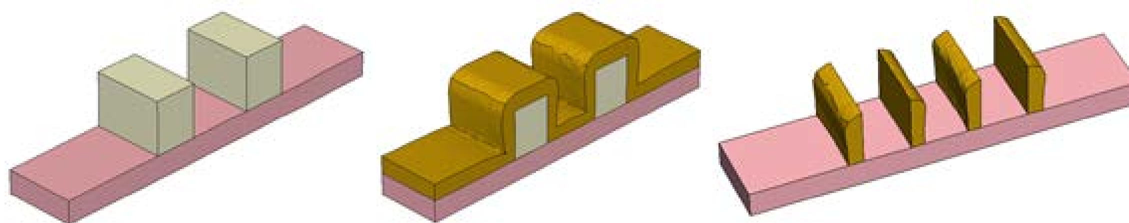


Fig. 15: Sequence for creating the pattern structure for fin structuring: Carbon lines patterned by the lithography step (left), deposition of an oxide layer (middle), back etching to create the spacers (right).

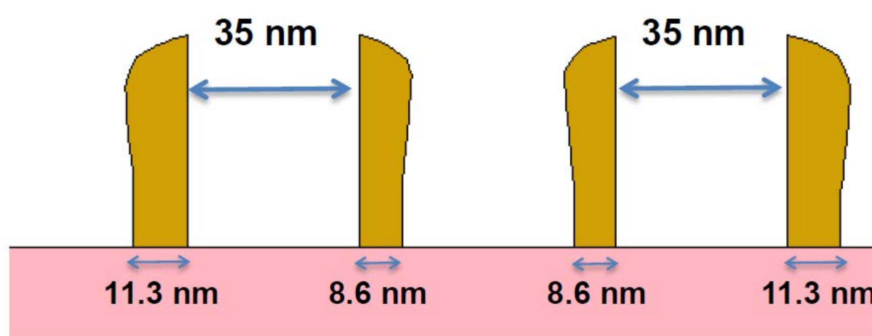


Fig. 16: Analysis of the bottom width of the spacers for carbon lines with the nominal CD of 35 nm.

to variations of the width of the inner lines, although these are smaller: In the example studied a variation of the carbon lines by 3.5 nm is reduced to about 0.3 nm for the inner lines.

Process compact models developed within SUPERTHEME for the fin and gate structuring were used for the final computation of the statistical fin widths and gate lengths distributions of the investigated 6T SRAM cell. For this study the variation parameters focus position and intensity threshold, serving as input for the compact models, were assumed to have a Gaussian distribution around the best focus and best intensity threshold, respectively, of the corresponding process windows. As an example Figure 17 shows the width distributions for the fin cores and for the inner fins.

Following the simulation of device patterning described above, statistical device simulation of a double-gate SOI FinFET has been performed employing the GARAND [1] tool from GSS. Because this study has focused on the interactions between systematic variations introduced by the structuring processes applied and statistical variations which result from the granularity of matter, the default model from GARAND was employed for the doping profiles. The fin height/width are 25nm and 10nm, respectively, and the physical gate length is 20 nm. The low channel doping is 1×10^{15} cm^{-3} and the maximum of the conformal source/drain doping is assumed to be 3×10^{20} cm^{-3}. The (110) sidewall orientation, beneficial for enhancing the hole mobility, has been adopted. Stress engineering has also been considered, and its impact is captured in a strain dependent mobility model. The drift diffusion device simulations are calibrated in using a comprehensive 3D ensemble Monte-Carlo simulation performance study. Density-gradient quantum corrections are included in the drift-diffusion simulations and are mandatory for such small-scale thin-body transistors.

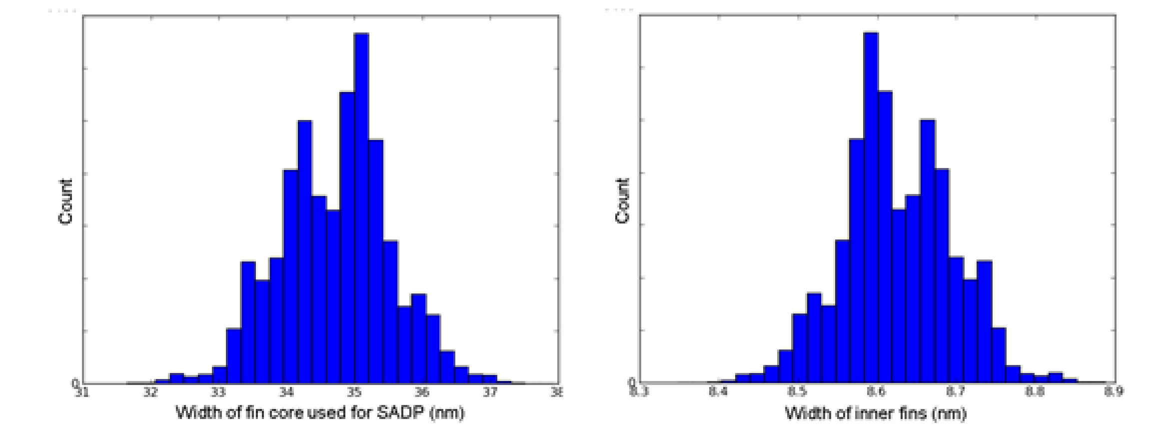


Fig. 17: Width distribution of the fin core used for SADP (left) and distribution of the resulting inner fin widths (right), the outer fins have a constant width of 11.3 nm

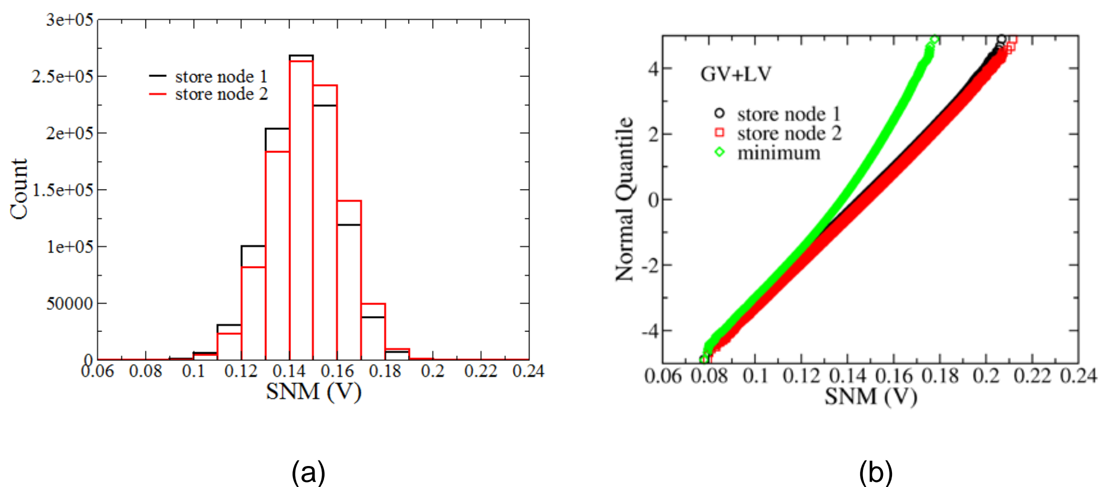


Fig. 18: Histogram plot (a) and Quantile vs. SNM plot (b) of SNM distributions subject to global CD variation and statistical local variability.

Subsequently, the statistical circuit simulation engine RandomSpice [20] was used to carry out the statistical compact model generation using GSS ModelGen technology, which handles non-Gaussian statistical parameter distributions with complex correlations. The hierarchical compact modeling approach outlined above in section 2.1 was used to simulate the combined effect of the patterning variations discussed above and the relevant statistical variations Random Dopant Fluctuations (RDF), Gate Edge Roughness (GER) and Fin Edge Roughness (FER) on an SRAM circuit. Here, due to the split of the gate patterning into two incremental lithography steps outlined above, again two triples of store transistors have different mean values and distributions of channel length. In turn, two distributions are obtained for the static noise margin, labelled in the following as “store node 1” and “store node 2”. Figure 18 shows the variation of the static noise margin of the SRAM cell considered.

Impact of self-heating on statistical variability

Advanced devices which employ SOI substrates are even more influenced by thermal effects, due to the smaller thermal conductivity of the buried oxide and the reduction of thermal

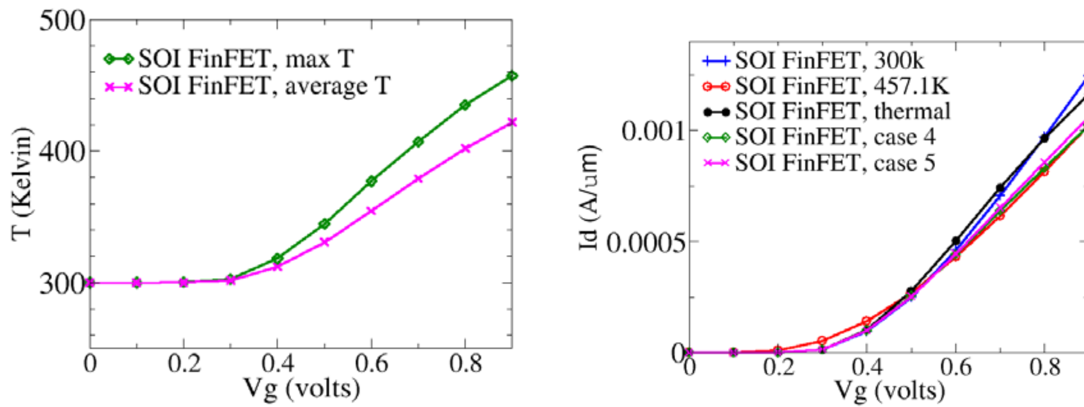


Fig. 19: Simulation results for the SOI FinFET: (a) maximum temperature and average temperature in the fin at high drain voltage; (b) Simulated Id-Vg characteristics at high drain voltage obtained from the electro-thermal simulations and with different assumptions for uniform temperature simulation results without self-heating.

conductivity in thin semiconducting films due to phonon ballistic transport. Therefore, also the self-heating of devices and its interactions with process variations must be considered.

New models and methods for the calculation of thermal conductivity in FinFETs have been implemented into GARAND [S8] and have been used for the simulation of bulk and SOI FIn-FETs with gate lengths of 25 nm. Both bulk and SOI FinFETs are significantly affected by self-heating. As an example Figure 19 shows the temperature increase during self-heating of an SOI FinFET and its impact on transistor performance in comparison with devices at room temperature and at 457.1K, which is the maximum temperature observed in the thermal simulations. Furthermore, “case4” and “case 5” identify simulations at a constant temperature defined by the maximum and average temperature, respectively, at each individual V_d and V_g

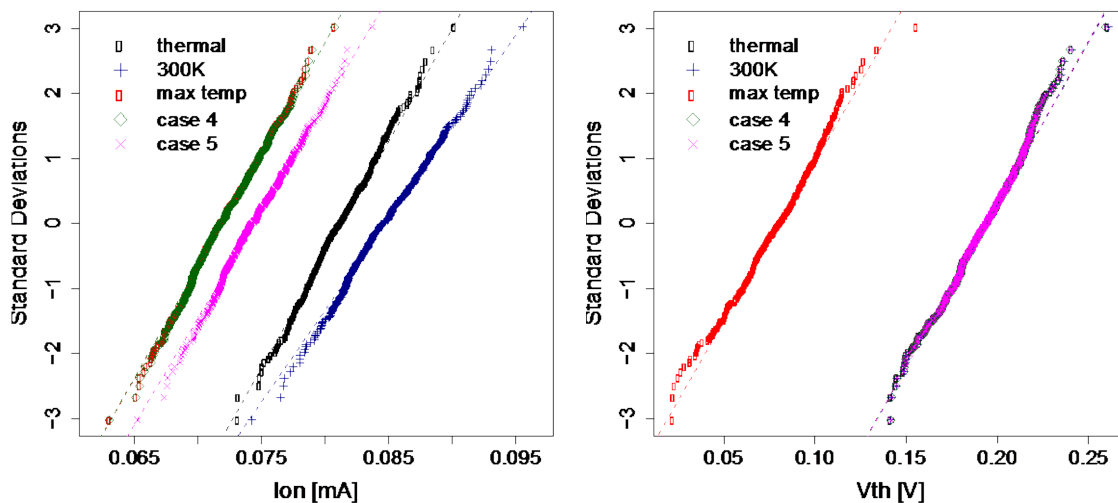


Fig. 20: Comparison of the statistical distribution of ON-current and threshold voltage for the SOI FinFET obtained from the electro-thermal simulations and with different assumptions for uniform temperature simulation results without self-heating.

bias condition obtained from the electro-thermal simulations. Figure. 20 shows the combined effect of self-heating and statistical variations (FER, GER and MGG) on the SOI FinFET. Further results have been published in various papers [S13-S15, S17, S18].

3 Inventory of Published Results

Many journal and conference publications were made based on SUPERTHEME results especially in the area of *More Moore* which has since long been the focal area for work on modeling and simulation at the academic SUPERTHEME partners Fraunhofer, University of Glasgow and TU Wien, but also at the software house Gold Standard Simulations GSS.

In the following the SUPERTHEME papers related to *More Moore* applications, published or at least accepted until the end of the project (end of 2015) are listed. Because SUPERTHEME started in October 2012 the first results were published in 2013. The full list of SUPERTHEME publications is also given on the SUPERTHEME WWW page at http://www.supertheme.eu/en/project/publications.html

3.1 More Moore Publications in 2013

[S1] S.M. Amoroso, L. Gerrer, J.M. Sellier, I. Dimov, M. Nedjalkov, S. Selberherr, A. Asenov, Quantum Insights in Gate Oxide Charge-Trapping Dynamics in Nanoscale MOSFETs, in: Simulation of Semiconductor Processes and Devices (SISPAD) 2013, IEEE, p. 25

[S2] P. Evanschitzky, A. Burenkov, J. Lorenz, Double Patterning: Simulating a Variability Challenge for Advanced Transistors, in: Simulation of Semiconductor Processes and Devices (SISPAD) 2013, IEEE, p. 105

[S3] J.M. Sellier, M. Nedjalkov, I. Dimov, and S. Selberherr, Decoherence and Time Reversibility: The Role of Randomness at Interfaces, J. Appl. Phys. 114 (2013) 174902

3.2 More Moore Publications in 2014

[S4] S.M. Amoroso, L. Gerrer, M. Nedjalkov, R. Hussin, C. Alexander, A. Asenov, Modeling Carrier Mobility in Nano-MOSFETs in the Presence of Discrete Trapped Charges: Accuracy and Issues, IEEE Trans. Electr. Dev. 61 (2014),p. 1292

[S5] A. Burenkov, J. Lorenz, Y. Spiegel, F. Torregrosa, Simulation of AsH3 Plasma Immersion Ion Implantation into Silicon, in: Proceedings International Conference on Ion Implantation Technology (IIT) 2014

[S6] J. Lorenz, E. Bär, A. Burenkov, P. Evanschitzky, A. Asenov, L. Wang, X. Wang, A.R. Brown, C. Millar, D. Reid, Simultaneous Simulation of Systematic and Stochastic Process Variations, in: Proceedings of Conference on Simulation of Semiconductor Processes and Devices 2014 (SISPAD 2014), p. 289

[S7] L. Wang, A. R. Brown, M. Nedjalkov, C. Alexander, B. Cheng, C. Millar, A. Asenov, 3D Coupled Electro-Thermal Simulations for SOI FinFET with Statistical Variations Including the Fin Shape Dependence of the Thermal Conductivity, 2014 IEEE 12th International Conference on Solid-State and Integrated Circuit Technology (ICSICT 2014), Guilin, China, Oct. 2014

[S8] L. Wang, A. R. Brown, M. Nedjalkov, C. Alexander, B. Cheng, C. Millar, A. Asenov, 3D Coupled Electro-Thermal FinFET Simulations Including the Fin Shape Dependence of the Thermal Conductivity, in: Proceedings of Conference on Simulation of Semiconductor Processes and Devices 2014 (SISPAD 2014), p. 269

[S9] L. Wang, A.R. Brown, C. Millar, A. Burenkov, X. Wang, A. Asenov, J. Lorenz, Simulation for Statistical Variability in Realistic 20 nm MOSFET, in: Proceedings of the 15th International Conference on Ultimate Integration on Silicon (ULIS), 2014, p. 5

[S10] X. Wang, D. Reid, L. Wang, A. Burenkov, C. Millar, B. Cheng, A. Lange, J. Lorenz, E. Baer, A. Asenov, Variability-Aware Compact Model Strategy for 20-nm Bulk

MOSFETs, in: Proceedings of Conference on Simulation of Semiconductor Processes and Devices 2014 (SISPAD 2014), p. 293

3.3 More Moore Publications in 2015

[S11] A. Burenkov, J. Lorenz, Y. Spiegel, F. Torregrosa, Simulation of Plasma Immersion Ion Implantation into Silicon, in: Proceedings of Conference on Simulation of Semiconductor Processes and Devices 2015 (SISPAD 2015), p. 218

[S12] R. Nagy, A. Burenkov, J. Lorenz, Numerical Evaluation of the ITRS Transistor Scaling, J. Comput. Electron. 14 (2015), p. 192

[S13] L. Wang, A. R. Brown, M. Nedjalkov, C. Alexander, B. Cheng, C. Millar, A. Asenov, Impact of Self-Heating on the Statistical Variability in Bulk and SOI FinFETs, IEEE Trans. Electr. Dev. 62 (2015), p. 2106

[S14] L. Wang, T. Sadi, M. Nedjalkov, A. R. Brown, C. Alexander, B. Cheng, C. Millar, A. Asenov. An Advanced Electro-Thermal Simulation Methodology For Nanoscale Device, in: Proceedings of IEEE 2015 International Workshop on Computational Electronics (IWCE 2015), p. 1

[S15] L. Wang, A. R. Brown, M. Nedjalkov, C. Alexander, B. Cheng, C. Millar, A. Asenov, 3D Electro-Thermal Simulations of Bulk FinFETs with Statistical Variations, in: Proceedings of Conference on Simulation of Semiconductor Processes and Devices 2015 (SISPAD 2015), p. 112

[S16] X. Wang, D. Reid, L. Wang, A. Burenkov, C. Millar, J. Lorenz, A. Asenov, Hierarchical Variability-Aware Compact Models of 20nm Bulk CMOS, in: Proceedings of Conference on Simulation of Semiconductor Processes and Devices 2015 (SISPAD 2015), p. 325

3.4 More Moore Publications in 2016

[S17] A. Burenkov, J. Lorenz, Simulation of thermo-mechanical effect in bulk-silicon FinFETs, Materials Science in Semiconductor Processing 42 (2016) 242-246

[S18] L. Wang, T. Sadi, M. Nedjalkov, A. R. Brown, C. Alexander, B. Cheng, C. Millar, A. Asenov, Simulation Analysis of the Electro-thermal Performance of SOI FinFETs, in: Proc. of Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS) 2016

4 Conclusions

Due to changes which occurred in the scope and actions of the International Technology Roadmap for Semiconductors the activities on *More Moore* applications within the SUPERTHEME project focused on the simulation of variability for bulk and FinFET transistors. Integration of existing, enhanced and simulation tools newly developed within SUPERTHEME have enabled these applications. The simulation of the combined effects of systematic and stochastic variations from equipment up to circuit level has been enabled. Especially, also the treatment of correlations has been included which requires to track these throughout each step of the simulation sequence. In order to enable both circuit simulation and the upward transfer of variability information into design variability- and correlation-aware compact models have been extracted.

References

[1] GARAND, http://www.goldstandardsimulations.com/
[2] Dr.LiTHO, http://www.drlitho.com
[3] CFD ACE, https://www.esi-group.com/

[4] Q-VT Plasma Processing Simulator, Quantemol Ltd., London, 2013
[5] Sentaurus Process, http://www.synopsys.com
[6] DF-ISE, http://www.synopsys.com
[7] Monolith, www.goldstandardsimulations.com
[8] L. Wang, A.R. Brown, C. Millar, A. Burenkov, X. Wang, A. Asenov, J. Lorenz, Simulation for Statistical Variability in Realistic 20 nm MOSFET, in: Proceedings of the 15th International Conference on Ultimate Integration on Silicon (ULIS), 2014, p. 5
[9] S.M. Amoroso, L. Gerrer, J.M. Sellier, I. Dimov, M. Nedjalkov, S. Selberherr, A. Asenov, Quantum Insights in Gate Oxide Charge-Trapping Dynamics in Nanoscale MOSFETs, in: Simulation of Semiconductor Processes and Devices (SISPAD) 2013, IEEE, p. 25
[10] L. Wang, A. R. Brown, M. Nedjalkov, C. Alexander, B. Cheng, C. Millar, A. Asenov, 3D Coupled Electro-Thermal FinFET Simulations Including the Fin Shape Dependence of the Thermal Conductivity, in: Proceedings of Conference on Simulation of Semiconductor Processes and Devices 2014 (SISPAD 2014), p. 269
[11] B. Cheng, et al., "Statistical Variability Compact Modeling Strategies for BSIM4 and PSP," IEEE Design and Test of Computers, Vol. 27, No. 2, pp. 26–35, Mar./Apr. 2010.
[12] Mystic, www.goldstandardsimulations.com
[13] X. Wang, D. Reid, L. Wang, A. Burenkov, C. Millar, B. Cheng, A. Lange, J. Lorenz, E. Baer, A. Asenov, Variability-Aware Compact Model Strategy for 20-nm Bulk MOSFETs, in: Proceedings of Conference on Simulation of Semiconductor Processes and Devices 2014 (SISPAD 2014), p. 293
[14] Sentaurus Device, http://www.synopsys.com
[15] H.-J. Cho, K.-I. Seo, W.C. Jeong et al., Bulk planar 20nm high-k/metal gate CMOS technology platform for low power and high performance applications,in: IEDM 2011, pp. 15.1.1.-15.1.4, IEEE, Piscataway and NJ, 2011
[16] International Technology Roadmap for Semiconductors, see www.itrs.net
[17] R. Nagy, A. Burenkov, J. Lorenz, Numerical Evaluation of the ITRS Transistor Scaling, J. Comput. Electron. 14 (2015), p. 192
[18] K. Schuegraf, M.C. Abraham, A. Brand, M. Naik, R. Thakur, *Semiconductor logic technology innovation to achieve sub-10 nm manufacturing*, IEEE J. Electron Devices Soc. 1 (2013) 66–75.
[19] DEP3D, physical deposition simulator, release 0.6.0, Fraunhofer IISB, Erlangen, Germany, 2008
[20] RandomSpice, www.goldstandardsimulations.com