Received 6 January 2020; revised 25 February 2020 and 25 March 2020; accepted 18 April 2020. Date of publication 28 April 2020; date of current version 14 July 2020. The review of this article was arranged by Editor A. Khakifirooz.

Digital Object Identifier 10.1109/JEDS.2020.2990718

A Benchmark Study of Complementary-Field Effect Transistor (CFET) Process Integration Options Done by Virtual Fabrication

B. VINCENT [1], J. BOEMMELS [2] (Member, IEEE), J. RYCKAERT [2], AND J. ERVIN [3]

1 Coventor SARL, Lam Research Company, 91140 Villebon-Sur-Yvette, France
2 Department of Logic, IMEC, 3001 Leuven, Belgium
3 Coventor inc., Lam Research Company, Fremont, CA 94538 USA

CORRESPONDING AUTHOR: B. VINCENT (e-mail: bvincent@coventor.com)

This work was supported by the Electronic Component Systems for European Leadership Joint Undertaking PIn3S Project.

ABSTRACT Four process flow options for Complementary-Field Effect Transistors (C-FET), using different designs and starting substrates (Si bulk, Silicon-On-Insulator, or Double-SOI), were compared to assess the probability of process variation failures. The study was performed using virtual fabrication techniques without requiring fabrication of any actual test wafers. In the study, Nanosheet-on-Nanosheet stacked channels provided superior process integration robustness compared to Nanowire-On-Fin stacked channels. For the Nanowire-On-Fin option, using an SOI substrate as the starting material (compared to Si bulk or DSOI) also strongly reduced process variation failure rates.

INDEX TERMS CMOS, CFET, DSOI, nanosheet, SOI, sensitivity analysis, virtual fabrication, 3 nm node.

I. INTRODUCTION

Sub-5 nm logic nodes will require an extremely high level of innovation to overcome the inherent real estate limitations at this increased device density. One approach to increasing device density is to utilize the vertical device dimension (z-direction), and stack devices on top of each other instead of conventionally side-by-side. Reference [1] the fabrication of Complementary-Field Effect Transistor (CFET) technology recently described in [2] suggests that it is possible to directly fabricate n-MOS transistors on top of p-MOS transistors or vice-versa. This architecture will require new metal wiring designs and Buried Power Rails (BPR) [3] within the substrate, as depicted in Figure 1. The design will be disruptive and will require the development of specific new processing steps such as dielectric selective deposition on metal.

In this study, we have used Coventor's SEMulator3D platform [4] with four different CFET reference designs/process flows to understand the design-process interaction risks of such technology development. This work was done in advance of any first mask tape-out and/or first wafer start. Certain process assumptions (such as etch selectivity, recess control or metal filling) were then considered prior to having silicon data. The type and level of assumptions were the same for the four designs/process flows considered during this study, to provide a fair comparison between each of the four alternatives. Each process flow studied had a specific channel architecture (Fin, nanowire or nanosheet) and a specific process flow with different Si starting substrates: bulk, Silicon-On-Insulator (SOI) [5], or Double Silicon-On-Insulator (DSOI) [6]. In this paper, we use the terms "Fin" to denote a vertical sheet of Silicon in contact with the Si substrate, "Nanowire" a vertical sheet of Silicon isolated from the Si substrate, and "Nanosheet" a horizontal sheet of Silicon isolated from the Si substrate. We compare the four process flows in terms of their robustness to process variation and identify the one with the lowest likelihood of specific processing failures. The targeted node dimensions were the same for all 4 or the flows: Fin or Nanowire width (or Nanosheet thickness) of 5nm with a 24 nn pitch, Fin height (or Nanosheet width) of 30 nm, and Gate CD of 14 nm with a 42 nm pitch.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

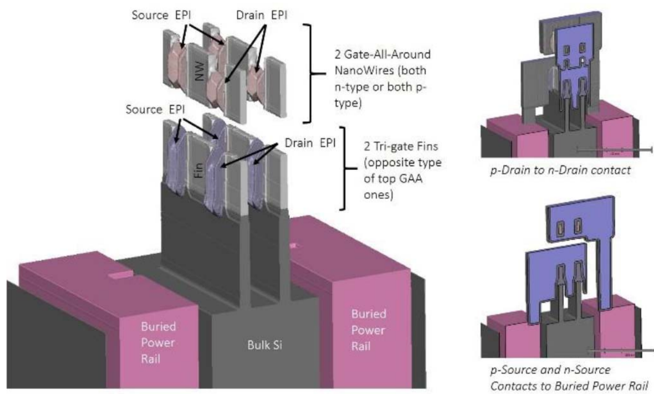

FIGURE 1. Representation of CFET inverter architecture (with all dielectrics as invisible in the 3D model).

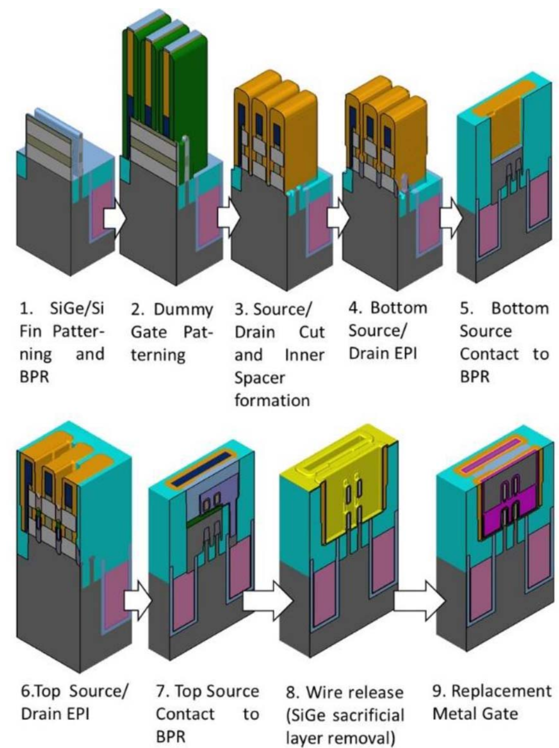

FIGURE 2. Si bulk (starting substrate) CFET process flow description.

The first part of this paper details the fabrication of the legacy CFET architecture described in [2]. It consists of a Gate-All-Around (GAA) nanowire channel stacked on top of a tri-Gate FinFET channel fabricated on an Si bulk substrate. We will highlight the inherent failure risks associated with the integration approach considered. In the second part, we will consider two new process flows, starting on SOI and DSOI substrates, respectively. We will show the benefits of those substrates (especially SOI) to reduce failure rate. In the third and last part of the paper, we will consider a fourth CFET architecture, a GAA Nanosheet stacked on top of a GAA Nanosheet. We will benchmark this design with the other three.

II. GAA NANOWIRE ON SI BULK TRI-GATE FIN CFET

A. PROCESS FLOW DESCRIPTION

Figure 2 depicts a process flow starting on an Si bulk substrate. Substrate dopant type will be opposite of the one of the bottom transistor to allow an isolation junction formation. A stack comprising a SiGe cap/Si top channel layer/SiGe sacrificial interlayer/Si bottom channel layer is epitaxially grown on the Si substrate. Fins are patterned down to the substrate and Buried Power Rails are formed (Step #1 in Figure 1). Next, dummy gates are patterned, the SiGe/Si stack in the future source/drain area is etched, and an inner spacer is formed between the two channels. The top transistor is then protected during Source/Drain epitaxial growth on the bottom channel – this is performed by depositing a Silicon-On-Carbide protective layer followed by spacer formation on the sides of the top transistor (not shown in Figure 1). Next, the bottom source/drain connection to BPR lines is generated by successive contact trench and via patterning. The trench and via are then metallized by Cobalt filling. An isolation layer is selectively grown on the bottom metal to enable Source/Drain epitaxial growth on the top transistor (Step #6). Like the bottom Source/Drain, the top ones are connected to BPR through contact trench and via patterning and filling. Finally, the replacement of the metal gate module is performed by combining the removal of sacrificial SiGe and gate-all-round deposition for the top transistor, along with a top cladding gate for the bottom transistor. The gate replacement introduces two work function metals for threshold voltage adjustment of both p- and n-channels: the bottom WF metal is first deposited on both gates, then removed from the top transistor. This is followed by top WF metal deposition on the top gate and on top of the n-WF metal of the bottom gate. Then, the gate trench is filled with Tungsten.

There is a high level of process complexity in this process flow, with some advanced process requirements such as selective dielectric growth on metal, material recess and metal filling in buried structures. As mentioned in the introduction, some assumptions were considered in our model based on the behavioral requirements of advanced processes still under development. For example, although multiple patterning will be considered for fin and gate patterning on Silicon, this simulation considered single patterning only for all process options. Pitch walking induced by multiple patterning is introduced through CD variation of the simulated single exposure.

B. INHERENT FAILURE RISKS

Using virtual fabrication, we discovered two key integration failure risks associated with the flow depicted in Figure 2. These two failure modes are caused by specific layer elevation control challenges. The two failures are not unique, but they are detrimental (i.e., killing failures). This study focuses on these two failures, since it is necessary to avoid these failures to build a robust process flow.

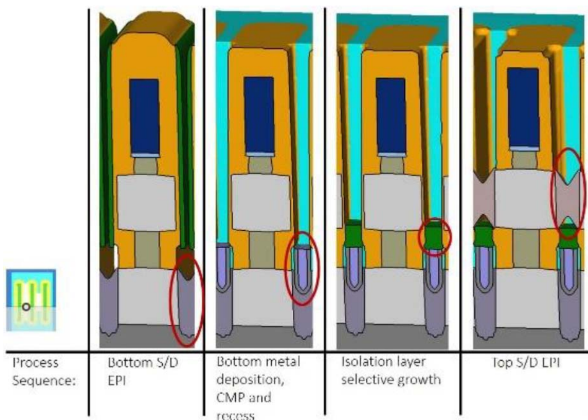

FIGURE 3. Detailed process sequence for introducing the nitride isolation layer between the top and bottom Source/Drain contact metals.

The first failure mode, identified as Failure A, corresponds to an excessively high (or low) elevation for the nitride layer, which is normally used to isolate the top source from the bottom source. Figure 3 details the process sequence for introducing this layer in between the bottom and the top Source/Drain epitaxy steps. This layer must be maintained between the two channel levels to avoid a process failure such as a short between the top metal contact and the bottom Source/Drain or the bottom metal contact and the top Source/Drain. As detailed in the figure, the elevation of the isolation layer will depend on control of the Si/SiGe Fin etch depth, the bottom metal CMP (post deposition), the metal recess depth, and the thickness of the nitride layer itself. The elevation of the bottom and the top channels will mainly depend on control of the SiGe/Si layer thicknesses during epitaxial growth.

The second failure mode, identified as Failure B, corresponds to shorting between the top and the bottom sources, occurring accidentally during the contact formation between the top source metal and the BPR. Figure 4 details the process sequence for this contact formation and highlights the risk of shorting in the location marked with a black arrow. Various sources of variation (in the nitride layer thickness or in the trench opening depth) can easily introduce such shorts.

III. TRI-GATE NANOWIRE ON SOI/DSOI DUAL-GATE FIN CFET

A. PROCESS FLOW DIFFERENCES

Two alternative process flows are also considered in this study, starting on SOI and DSOI substrates. Along with these different starting substrates, these process flows strongly differ from the legacy Si bulk flow in the processes used for the Source/Drain cut and the inner spacer formation.

The Fin patterning sequence in the SOI/DSOI options is done in multiple Si and oxide etch steps, to take advantage of Si etch selectivity to oxide and oxide etch selectivity to Si. In the Si bulk case, the Si/SiGe stack is etched in one single timed-etch step. The degree of timed etch is then

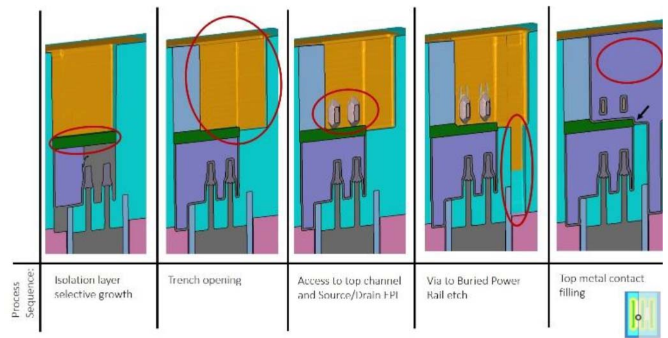

FIGURE 4. Detailed process sequence for connecting the top Source contact metal to the Buried Power Rail.

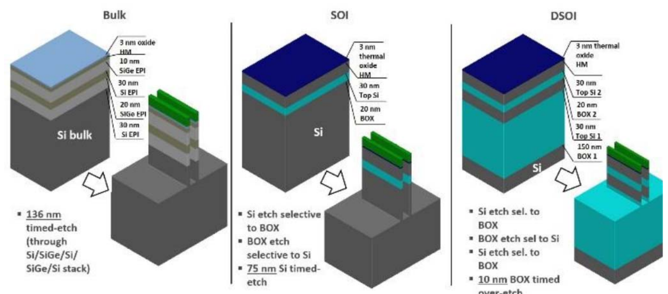

FIGURE 5. Comparison of Fin patterning etch processes for Si bulk, SOI, and DSOI cases. The amount of timed etch is strongly reduce in the SOI and DSOI cases.

strongly reduced on the SOI/DSOI case, leading to much better control as depicted in Figure 5. The same difference is also present for the Source/Drain Cut process.

Dummy gates on SOI/DSOI are also much shorter than in the bulk case, since there are no long etch requirements for inner spacer formation. The inner spacer formation in the Si bulk case, introduced after a lateral SiGe recess, is indeed not considered for the SOI/DSOI cases due to the Buried Oxide already present, as illustrated in Figure 6. A shorter dummy gate will induce shorter recess depth in the rest of the flow, which is also favorable for better uniformity control.

The DSOI option differs from the SOI case by using one additional BOX layer to isolate the bottom transistor from the substrate. Bottom device in the bulk and SOI flows are isolated from the substrate with isolation junction. Note that this additional DSOI BOX layer makes a clear difference with respect to the transistor's elevation control for both the SOI and DSOI cases. The bottom transistor in the SOI case is built within the bulk Si, whereas the bottom transistor in the DSOI case is built in the first SOI layer. The elevation of the bottom transistor in the DSOI case has more sources of variation (first BOX and first Si layer thicknesses) than for the SOI case.

B. BENCHMARK WITH RESPECT TO FAILURE RATE

In order to compare the three first process options (starting on Si, SOI, or DSOI) in terms of failure rate, we executed virtual Design of Experiments (DOEs) using the

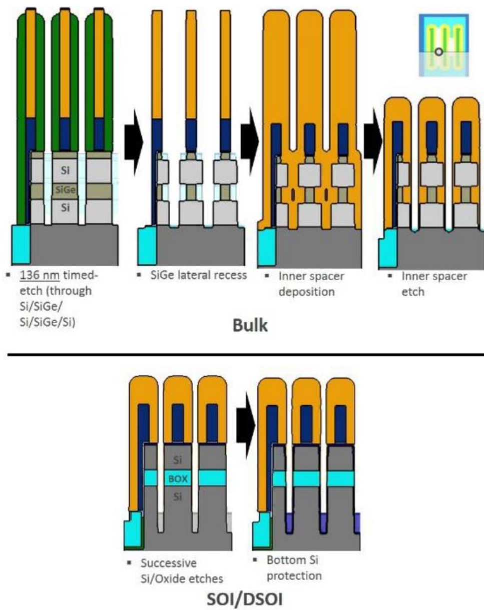

FIGURE 6. **Difference between bulk and SOI/DSOI flows for the Source/Drain Cut and Inner spacer formation.**

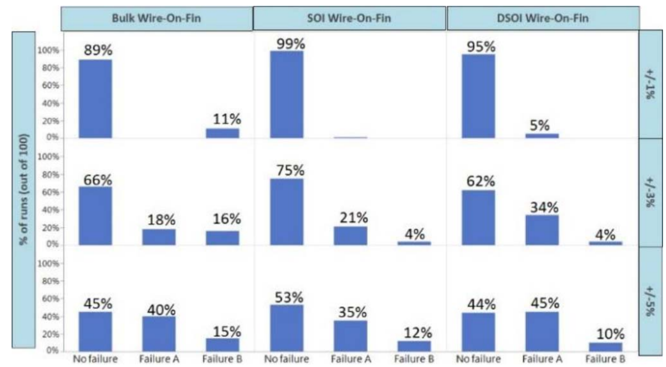

FIGURE 7. **Percentage of passed runs and failed runs for the three flows and for three process variation assumptions (+/ − 1, 3, 5%).**

	Bulk	SOI	DSOI
Top process (with p-value <1e-5) parameters impacting elevation of isolation layer (obtained from sensitivity analysis)	Gate HM deposition thickness PCE	Gate HM deposition thickness PCE	Gate HM polish depth PCE
	Gate HM polish depth PCE	Spacer etch depth PCE	Gate HM deposition thickness PCE
	Spacer etch depth PCE	Gate HM polish depth PCE	Spacer etch depth PCE
	Co recess PCE	Co recess PCE	Co recess PCE
	Channel thickness (30nm +/- 1nm)	Top contact trench etch PCE	BOX1 thickness (150nm +/-1nm)
	Dummy Gox thk (2nm +/-0.4nm)	BOX thickness (20nm +/-0.5nm	
		Channel thickness (30nm +/- 0.5nm)	
Top process (with p-value <1e-5) parameters impacting elevation of the bottom of the top channel (obtained from sensitivity analysis)	Channel thickness (30nm +/- 1nm)	BOX thickness (20nm +/-0.5nm)	BOX1 thickness (150nm +/-1nm)
	SiGe sacrificial layer thickness (20nm +/-1nm)		Channel thickness (30nm +/- 0.5nm)
			BOX thickness (20nm +/-0.5nm)

FIGURE 8. **Ranking of variation sources by their impact on Failure A rate for Bulk, SOI, and DSOI flows.**

SEMulator3D virtual fabrication platform. For each of the 3 process flows considered, process parameters were selected and varied within an Analytics study using a Monte Carlo uniform distribution. In the deposition layers, variations to the target were established at +/−2 nm for CVD and +/−1 nm for ALD and EPI. For SOI/DSOI substrates, Si and BOX thickness target variations were set at +/−0.5 nm. Finally, specific timed etch, timed CMP, and thickness targeted deposition steps were varied within +/− x% of the target with x = 1, 3 or 5. These variation magnitudes were selected based upon typical variability controls available during advanced manufacturing, as provided by process experts.

A set of DOEs (with each DOE having 100 virtual runs) was established by introducing all variations. Those 3 DOEs were used to compare flows with respect to the occurrence of failures A and B, reported earlier. Figure 7 displays the percentage of runs without failures, runs with Failure A, and runs with Failure B for the three flows and for the three variation assumptions. We note that the SOI option outperforms both bulk Si and DSOI, by about a 10% higher pass yield.

Regression and sensitivity analyses were performed on the data collected by these DOEs in order to identify the most important source of variation impacting the elevations of the isolation layer and the channels (related to the source of Failure A). Those most critical sources of variation are listed in Figure 8 for the three flows. PCE, indicated in most parameter names, stands for Percentage Error to Target. In this table we highlight key sources of variation that are present for Si bulk and DSOI cases but are not present for SOI. The introduced variations (channel thickness of 30 nm +/−1 nm, SiGe sacrificial thickness of 20 nm +/−1 nm, DSOI BOX1 layer thickness of 150 nm +/−1 nm) are then defined as the root cause for the observation of higher failure rates in both bulk Si and DSOI as opposed to SOI. Specifically, SOI outperforms DSOI – the contribution of BOX1 thickness and channel thickness on the bottom transistor elevation for the DSOI case (not existing for the SOI case) are pointed as the main root cause for DSOI lower success rate. BOX1 thickness has actually the exact same impact on the elevation of the isolation layer and on the elevation of the bottom of the top transistor and then no impact on Failure A events. Channel thickness variation is then the key root cause for SOI outperforming DSOI in terms of success rate.

IV. GAA NANOSHEET ON SI BULK GAA NANOSHEET CFET

A. DEVICE DESCRIPTION

This section describes a fourth type of CFET architecture consisting of stacked GAA Nanosheets, as depicted in Figure 9. A key benefit of this architecture is the ability to relax the spacing between the p- and n- transistors, which is constrained by a maximum gate height in the earlier designs. The constraint on earlier designs is driven by the gate height, which must not exceed a limit on the maximum high aspect ratio that it is feasible to fabricate during later process steps in the Replacement Metal Gate process. The Nanosheet-On-Nanosheet considers use of a 40 nm spacing between the two channels, whereas previous designs had only 20 nm. This wider spacing is expected to provide greater tolerance for elevation control and for failures reported earlier.

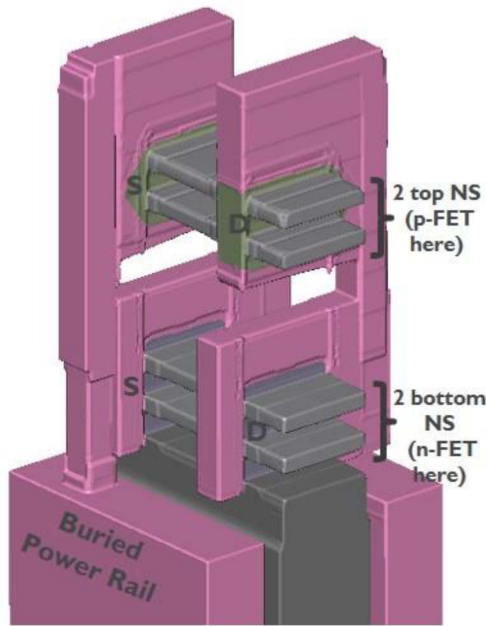

FIGURE 9. **3D model of a CFET device comprising a p-FET transistor with two nanosheets, on top of an n-FET transistor with two nanosheets.**

The fabrication process of this nanosheet-on-nanosheet design is somewhat similar to the nanowire-on-Fin design, although some process specifications differ. In the Nanowire-On-Fin approach, the channel height (targeted at 30 nm) is defined by epitaxial growth and the channel width (targeted at 5 nm) is defined by Fin patterning. In the Nanosheet-On-Nanosheet case, the horizontal channel height (still targeted at 30 nm) is defined by Fin patterning epitaxial growth, and the channel width (still targeted at 5 nm) is defined by epitaxial growth.

B. BENCHMARK WITH RESPECT TO FAILURE RATE

In this section, we present a first benchmark of the stacked Nanosheet design against the earlier design, based on robustness with respect to failures A and B reported earlier when process variations are introduced.

We consider only five variation sources (the most important factors determined in the previous section): the delta-to-target values for channel thickness, channel separation layer thickness, metal recess depth prior to nitride isolation growth, gate hardmask polish endpoint, and oxide recess depth for top contact. In this section, the DSOI

Nanowire-On-Fin design is not considered, since it demonstrated no specific advantage from the results shown in Figure 7. For the three remaining design options, 100 virtual wafers were processed for each flow with the five sources of variation set at either $+/-3\%$ or $+/-5\%$. Similar to the methodology developed earlier, we report in Figure 10 the percentage of runs for each failure type, for each of the three flows and for each variation assumption. We note that the Nanosheet-On-Nanosheet option strongly reduces (for large

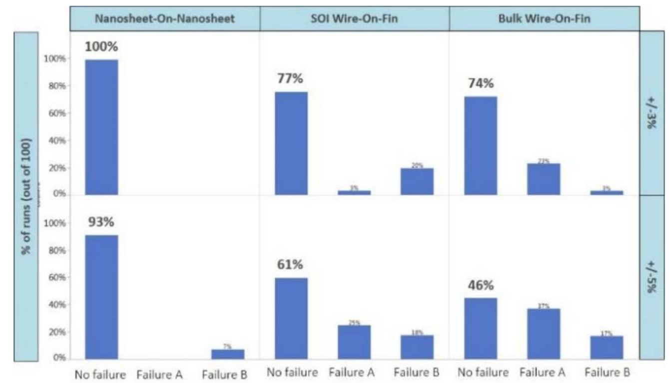

FIGURE 10. **Percentage of passed runs and failed runs for the three flows and for two process variation assumptions (+/−3, 5%).**

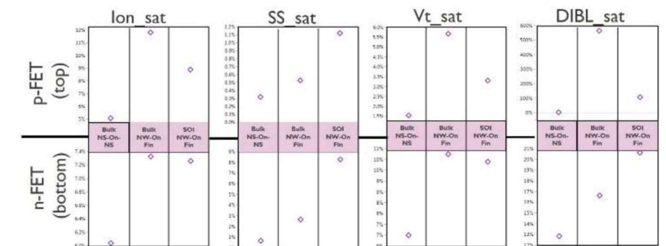

FIGURE 11. **(Standard Deviation/Mean) ratio of Ion, SS, Vt and DIBL in the saturation regime for p- and n-FET out of 100 simulation runs for each flow.**

variation assumptions) the failure rate thanks to the larger spacing between the two transistors.

C. BENCHMARK WITH RESPECT TO DEVICE PERFORMANCE CONTROL

In this section, we continue the benchmark between the Nanosheet-On-Nanosheet design and the Nanowire-On-Fin design from another perspective. We review the impact of process variation on electrical performance control obtained for each design. Note that we will not discuss the actual device performance for each flow but will only focus on the electrical performance variation with respect to median values, since the scope of this paper is limited to process variation control. CFET electrical performance was measured on a 3D model using the SEMulator3D device analysis capability (including a Poisson equation and Drift Diffusion solver). In this section, four sources of variations were considered: Channel EPI thickness variation of $+/-10\%$ to target, Channel patterning CD variation of $+/-2$ nm to target, Gate patterning CD variation of $+/-2$ nm to target and Gate Spacer thickness variation of $+/-10\%$ to target. Again, the variation levels were selected based upon state-of-the-art variation margins obtained in High Volume Manufacturing. 100 runs per flow were generated considering those sources and magnitudes of variation. For each run simulated, Ion current (Ion), Subthreshold Slope (SS), Threshold Voltage (Vt), and Drain Induced Barrier Lowering (DIBL) were extracted in the saturation regime of the p- and the n- transistor. In

Figure 11, we report the (Standard Deviation/Mean) ratio of each of those collected parameters. We see from the figure that the Nanosheet-On-Nanosheet design consistently offers tighter control on electrical performance. Some additional simulation (not detailed here) showed that the epitaxy thickness variation must be reduced from $+/-10\%$ to $+/-3.3\%$ and the channel patterning CD uniformity from $+/-2$ nm to $+/-1$ nm on a Nanowire-On-Fin architecture in order to obtain similar device performance to what Nanosheet-On-Nanosheet displays in Figure 11. In conclusion, thanks to its specific geometry (with channel height being controlled by patterning and channel width controlled by epitaxy), the Nanosheet-On-Nanosheet architecture presents another key benefit with respect to device performance control. We note that this benefit is generic to the nature of the nanosheet orientation and is not specific to nanosheet designs implemented in a CFET architecture.

V. CONCLUSION

Virtual fabrication was used in this study to perform a deep dive into design and process considerations for future CFET fabrication process flows. In this study, we benchmarked four different options and identified key potential failure modes. Failures due to layer elevation control challenges were more pronounced for flows starting on bulk or DSOI substrates compared to SOI substrates for a Nanowire-On-Fin design. These failure types are not present when considering a Nanosheet-On-Nanosheet design. In addition, thanks to the horizontal channel specificity of the Nanosheet design, tighter control over transistor electrical performance was also demonstrated. In conclusion, the GAA-Nanosheet on Si bulk GAA-Nanosheet integration approach was determined to be the most robust flow for CFET fabrication with respect to all process variations tested during modeling.

REFERENCES

[1] L. Brunet *et al.*, "First demonstration of a CMOS over CMOS 3D VLSI CoolCube integration on 300mm wafers," in *Proc. IEEE Symp. VLSI Technol.*, Honolulu, HI, USA, 2016, pp. 1–2, doi: 10.1109/VLSIT.2016.7573428.

[2] J. Ryckaert *et al.*, "The complementary FET (CFET) for CMOS scaling beyond N3," in *Proc. IEEE Symp. VLSI Technol.*, Honolulu, HI, USA, 2018, pp. 141–142, doi: 10.1109/VLSIT.2018.8510618.

[3] A. Gupta *et al.*, "High-aspect-ratio ruthenium lines for buried power rail," in *Proc. IEEE Int. Interconnect Technol. Conf. (IITC)*, Santa Clara, CA, USA, 2018, pp. 4–6, doi: 10.1109/IITC.2018.8430415.

[4] (2019). *Coventor SEMulator3D*. Accessed: May 11, 2020. [Online]. Available: https://www.coventor.com/products/semulator3d/

[5] J. P. Colinge *et al.*, "Basics of silicon-on-insulator (SOI) technology," in *Wafer Bonding* (Materials Science), vol. 75, M. Alexe and U. Gösele, Eds. Heidelberg, Germany: Springer, 2004.

[6] M. K. Emsley, O. Dosunmu, and M. S. Unlu, "High-speed resonant-cavity-enhanced silicon photodetectors on reflecting silicon-on-insulator substrates," *IEEE Photon. Technol. Lett.*, vol. 14, no. 4, pp. 519–521, Apr. 2002, doi: 10.1109/68.992597.