3D Nanosheet Field-Effect Transistors: Optimization and Analysis

Sawyer Brundage
Department of Electrical Engineering
Stanford University

***Abstract*—Nanosheet field-effect transistors (NSFETs) are a promising device architecture for new technology nodes due to their superior electrostatic control and scalability. In this work, we build 3D models and simulations using Sentaurus TCAD and simulation tools to optimize and assess their performance in a 31-stage ring oscillator (RO). We optimize the device geometry, specifically the gate length, spacer length, and contact lengths, while maintaining a Contacted Gate Pitch (CGP) of 42 nm. The optimization metric is initially based on $C_{gg,on}/I_{eff}$ (ratio of gate capacitance in the ON state to the effective current), but we extend our analysis to account for when wire interconnect capacitance dominates in a RO circuit. We evaluate delay and energy efficiency for 1-5 nanosheets in a device. The results demonstrate that, while attempting balance a small $C_{gg,on}$ and a large I_{eff} is suitable when considered a negligible wire capacitance, any attempts to reduce $C_{gg,on}$ are futile when the wire capacitance dominates. These insights provide an optimization roadmap for HVT, SVT, and LVT applications.**

***Index Terms*—NSFET, Ring Oscillator, Sentaurus, Nanosheet**

I. Introduction

For decades, planar field-effect transistors (FETs) served as the cornerstone of analog and digital integrated circuits. This has traditionally been enabled by Moore's Law, which postulates that the number of transistors that can be fit in a set area doubles approximately every two years due to constant device scaling. However, as these devices approached the nanometer-scale, short-channel effects, high leakage currents, and variability due to process limitations rendered them infeasible for use in integrated circuits. To address these challenges, FinFETs emerged as the dominant transistor architecture, offering improved electrostatic control through a three-dimensional fin structure. That said, FinFETs also face scalability limitations along with a quantized, inflexible effective gate width. As a result, there remains a need for a transistor architecture that scales well at the nanometer-scale while also providing the flexibility of a continuous gate width.

To overcome these limitations, three-dimensional nanosheet FETs (NSFET) have been introduced as a next-generation transistor architecture. Although the concept has existed longer than FinFETS, the process technology needed to manufacture them with precision has only recently become available to chip manufacturers. NSFETs enhance gate control by surrounding the channel on all sides, suppressing short channel effects and reducing off-state leakage currents. Compared to FinFETs. NSFETs offer a larger effective width per unit footprint and allow for better width scaling without increasing device height. The greater design flexibility provided by NSFETs make them a promising candidate for both low-power and high-performance applications.

This paper explores the physical structure of 3D nanosheets while also assessing their effective and off-state currents at varying threshold voltages. NMOS and PMOS NSFET structures are modeled and simulated using Sentaurus TCAD, from which the capacitance and current data is gathered. Then, we will conduct a performance analysis using a 31-stage ring oscillator (RO) with a fanout (FO) of 4.

II. NSFET Structure

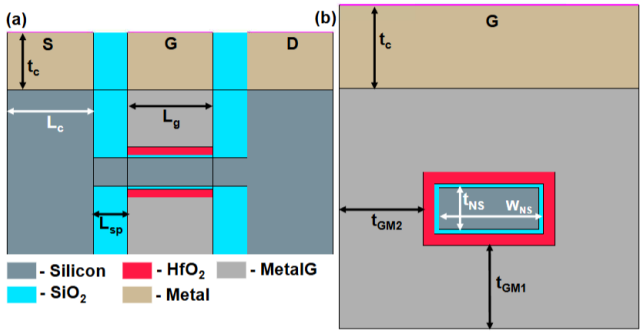


Fig. 1: (a) Longitudinal cross section of a NSFET with one sheet, (b) corresponding lateral cross-section

The basic structure of a NSFET is illustrated in Figure 1. As the illustration shows, the gate metal wraps around the silicon channel(s) between the source and drain, of which an NSFET device can have multiple (one sheet, two sheets, etc.). For our simulations, we limit the number of simulations to 1-5 for reasons that will be revealed in the next section. The various dimensions of the NSFET that we use for our simulation model are also labeled in Figure 1, and the default values we use in our initial, baseline model are shown in Table I.

Our NSFET model is split into two regions in terms of how meshing is handled: the area around the silicon channels (Region II) and the rest of the structure (Region I). Figure 2 provides an illustration of the two regions of a 5-NS device, while Table II shows the mesh spacings for them. This meshing scheme was chosen as a compromise between simulation accuracy and runtime. The workfunction of the gate metal is also tuned depending on the application. For our simulations,

NSFET parameters	Description	Default Values
L_g	Gate Length	15 nm
L_{sp}	Spacer Length	6 nm
CGP	Contacted Gate Pitch	42 nm
t_{NS}	Si NS Thickness	5 nm
t_{GM1}, t_{GM2}	Gate Metal Thickness	10 nm
t_c	S/D/G Contact Metal Thickness	10 nm
t_{HfO2}	HfO_2 Thickness	1.4 nm
t_{SiO2}	SiO_2 Thickness	0.5 nm
W_{NS}	Si NS Width	12 nm
N_{SD}	S/D Doping	10^{20} cm^{-3}
N_{ch}	Channel Doping	0 (intrinsic Si)
$\#NS$	Number of Si NS	Vary from 1 to 5

TABLE I: Default Parameter Values

we test high-threshold (HVT), standard threshold (SVT), and low-threshold (LVT) devices that each have specifically tuned gate workfunctions.

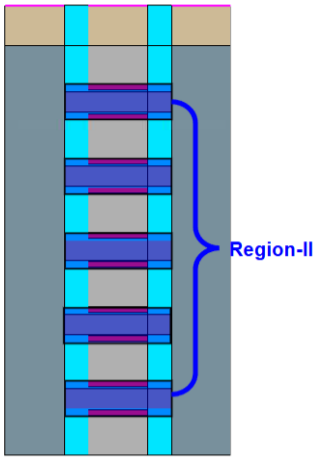


Fig. 2: Region II is highlighted in blue. Region I corresponds to the rest of the structure.

A total of ten models were generated for the simulations: five corresponding to each number of nanosheets (1-5) for both NMOS and PMOS. For NMOS, the source and drain regions were doped with Arsenic. For PMOS, we used Boron. Figures 3 and 4 show 3D models for NMOS and PMOS nanosheet FET devices, respectively.

III. Optimization and Simulations

For each of our simulations, we extracted four key values via gate and drain voltage sweeps: gate-to-channel capacitance C_{gc}, parasitic capacitance C_{par}, effective current I_{eff}, and off-state leakage current I_{off}. The capacitance is extracted using AC small-signal simulations with a frequency of 100 kHz and $V_{ds} = 0$. C_{par} is extracted by measuring the gate capacitance C_{gg} when $V_{gs} = 0$. Then, to find the gate-to-channel capacitance C_{gc}, we can extract the capacitance when the device is ON $V_{gs} = V_{dd} = 0.7$ V and subtract C_{par}. I_{off}

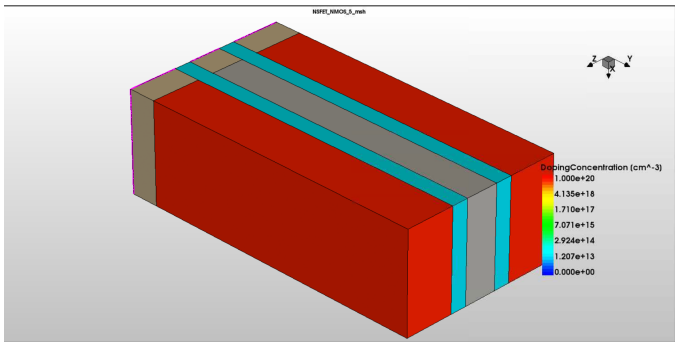


Fig. 3: 3D NMOS NSFET Structure

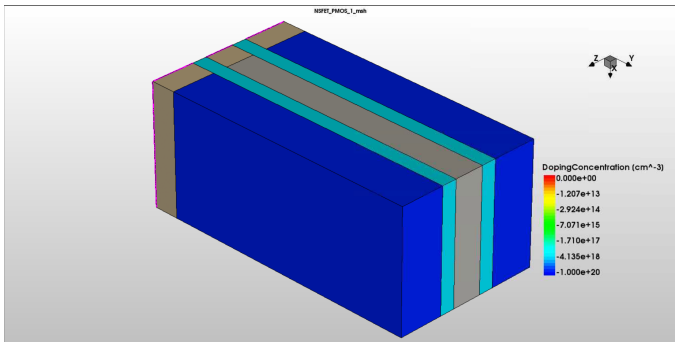


Fig. 4: 3D PMOS NSFET Structure

is extracted by measuring the drain current when $V_{ds} = 0.7$ V and $V_{gs} = 0$. Finally, the effective current I_{eff} is defined as the average of the drain current when $V_{ds} = 0.7$ V, $V_{gs} = 0.35$ V and when $V_{ds} = 0.35$ V, $V_{gs} = 0.7$ V.

In this section, we will provide three designs that have different threshold voltages: LVT (optimized for drive/performance), HVT (optimized for leakage/standby power), and SVT (moderate drive and leakage). In older technology nodes, threshold voltages were tuned primarily by changing the doping concentration of the silicon channel along with altering its length [1]. However, more recent technology nodes alter threshold

Region	Xmax	Ymax	Zmax	Xmin	Ymin	Zmin
I	0.01	0.005	0.005	0.001	0.0005	0.0005
II	0.001	0.0005	0.0005	0.001	0.0001	0.0001

TABLE II: Mesh spacings (in μm) of regions I and II. X-direction is longitudinal (along length of NSFET), whereas Y- and Z-directions are transverse.

voltages by adjusting dielectric thicknesses and permittivity along with tuning gate metal workfunctions [2]. We will be tuning the gate workfunctions to achieve LVT, SVT, and HVT designs.

Our optimization strategy for our NSFET devices is to minimize $C_{gg,on}/I_{eff}$ for our LVT, SVT, and HVT devices by changing L_g, L_{sp}, and L_c while maintaining the CGP given in Table I (CGP = $L_g + 2L_{sp} + L_c$). This metric will be justified for the $C_{wire} = 0$ (no wire capacitance) case in the next section. Since we want to minimize $C_{gg,on}/I_{eff}$, we need to find the right balance of a low $C_{gg,on}$ and a high I_{eff}. As a result, a lower L_g and a higher L_c than the default values will give us an optimal configuration. Table III shows the results of our exploratory simulations with different lengths using a default workfunction of 4.51 eV. The optimal configuration is bold.

For our NMOS device, we used workfunctions of 4.51 eV, 4.4 eV, and 4.27 eV for HVT, SVT, and LVT, respectively. Figures 5 and 6 illustrate how the effective and off-state currents change as a function of the number of nanosheets (#NS). Naturally, we see a somewhat linear increase in both currents as #NS increases. This is due to the increased effective cross-sectional area of the silicon channels (area of one nanosheet * #NS), reducing the total resistance between the source and drain. The LVT device has the highest effective and off-state currents (the latter being orders of magnitude higher than the others) due to having the lowest threshold voltage, while the HVT device has the lowest. Table IV gives the I_{off} constraints for LVt, SVT, and HVT devices that we use for our simulations.

The capacitance plots as a function of #NS for the LVT, SVT, and HVT NMOS devices are shown in Figure 7. Similar to the currents, the capacitance values increase linearly with the number of nanosheets. This follows from what we know about the structure of the NSFET and how capacitance adds linearly in parallel. If the gate capacitance is C_{gg} for each nanosheet, then a device with n nanosheets should have a total gate capacitance of nC_{gg}.

Figure 8 illustrates how our performance metric, $C_{gg,on}/I_{eff}$, changes as the number of nanosheets increases. For the HVT, SVT, and LVT cases, the ratio decreases until it reaches 3 NS. Afterwards, the plot shows negligible increase or decrease with more nanosheets. This serves to illustrate the infeasibility of increasing the number of nanosheets past 3, since there will be diminishing returns in terms of performance. Thus, we can conclude that adding more nanosheets beyond 3 is generally not worth the negligible performance increase.

The same simulations were also run for our PMOS NSFET model, which was configured to have approximately the same effective and off-state current values as the NMOS for all #NS. The channel width W_{NS} was doubled from 12 nm to 24 nm to account for the lower mobility of holes in silicon. We used workfunctions of 4.04 eV, 4.15 eV, and 4.29 eV for HVT, SVT, and LVT, respectively. The plots shown in Figures 9 through 12 show largely the same trends as the NMOS case. The magnitude of the capacitance values is approximately doubled compared to the NMOS as a result of the wider nanosheets, but the $C_{gg,on}/I_{eff}$ ratio still plateaus at about 3 NS.

In an attempt to achieve a better $C_{gg,on}/I_{eff}$ ratio, we tested a thinner nanosheet of 3 nm (as opposed to the default 5 nm). This would ideally reduce the capacitance enough to decrease the ratio and thus achieve better performance for LVT, SVT, and HVT devices. The results for the NMOS case are shown in Table V. The workfunctions did not need to be altered to meet the LVT, SVT, and HVT requirements in the $t_{NS} = 3$ nm case.

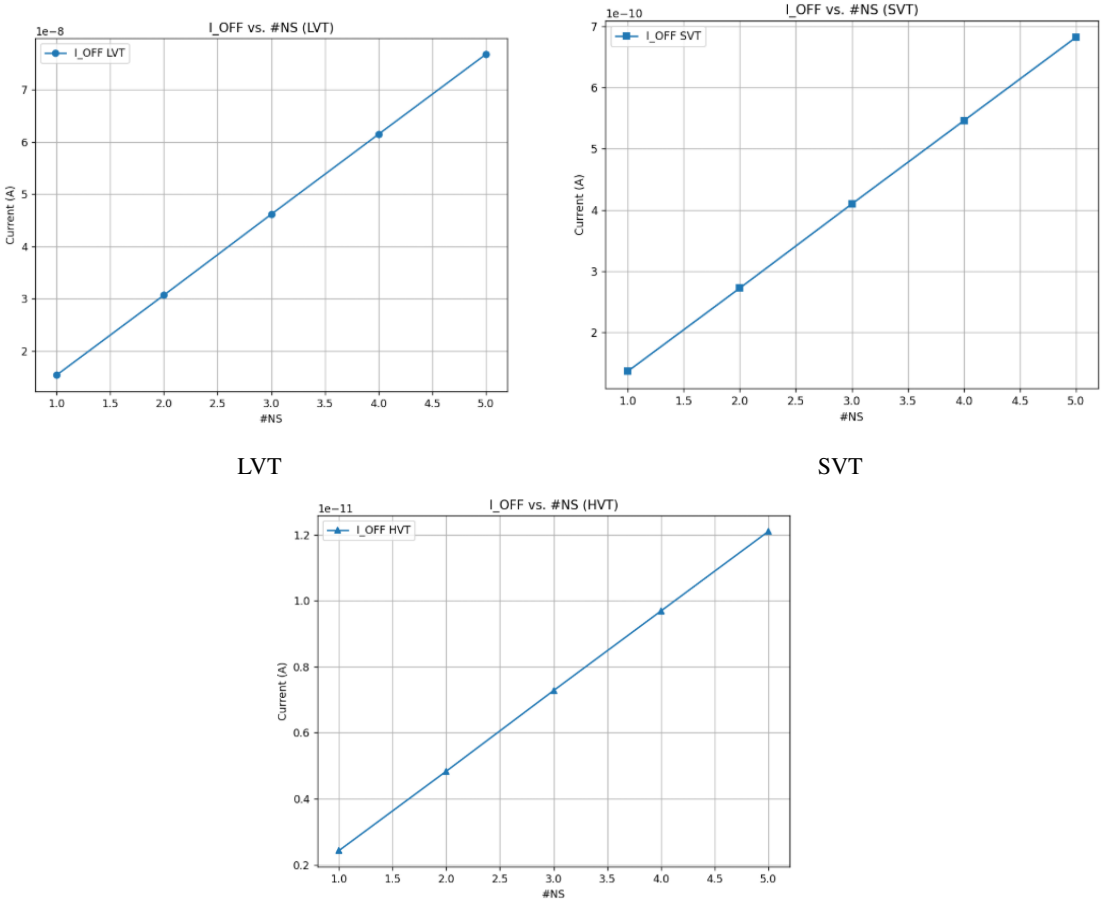
As the simulation results show, a thinner nanosheet does not result in better performance. Although the capacitance decreases due to less gate metal needed to wrap around the thinner nanosheet, the effective current also decreases drastically. Since the cross-sectional area of the channel (nanosheet) decreases, the channel resistance must increase. In our simulation, the effective current decreases by a larger factor than the capacitance, resulting in a larger $C_{gg,on}/I_{eff}$ with a thinner nanosheet. In a real scenario, very thin nanosheets bring limitations that are not captured by the current sdevice physics model in our simulations. Due to the high surface-to-volume ratio of a thin nanosheet, the effects of Coulomb scattering and surface roughness scattering on mobility would need to be considered for accurate simulations. Additionally, Schrödinger's equation also needs to be solved along with Poisson's equation (which is already included in the physics model) to accurately model quantum confinement that occurs in ultrathin channels. This would more accurately model the density of states in the nanosheets, providing a better estimation of carrier distribution.

Lengths (nm)	$C_{gg,on}$ (F)	I_{eff} (A)	$C_{gg,on}/I_{eff}$
$L_g = 15, L_{sp} = 6, L_c = 15$	3.72e-17	4.15e-6	8.96e-12
$L_g = 18, L_{sp} = 4, L_c = 16$	4.82e-17	5.061e-6	9.524e-12
$L_g = 10, L_{sp} = 10, L_c = 12$	2.54e-17	3.07e-6	8.27e-12
$\mathbf{L_g = 12, L_{sp} = 6, L_c = 18}$	**3.32e-17**	**4.485e-6**	**7.402e-12**
$L_g = 14, L_{sp} = 4, L_c = 20$	4.27e-17	5.24e-6	8.15e-12
$L_g = 12, L_{sp} = 6, L_c = 20$	3.32e-17	4.445e-6	7.47e-12

TABLE III: Exploratory Simulation Results (WF = 4.51 eV)

VT Type	LVT	SVT	HVT
I_{off} Constraint	50 nA	500 pA	5pA

TABLE IV: HVT, SVT, LVT Criteria



LVT SVT

HVT

Fig. 5: Off-State Leakage Currents for NMOS NSFETs

	$t_{NS} = 5$ nm	$t_{NS} = 3$ nm
LVT $C_{gg,on}/I_{eff}$	3.873e-12	4.90e-12
SVT $C_{gg,on}/I_{eff}$	5.16e-12	6.55e-12
HVT $C_{gg,on}/I_{eff}$	7.402e-12	9.75e-12

TABLE V: Default vs. Thinner Nanosheet

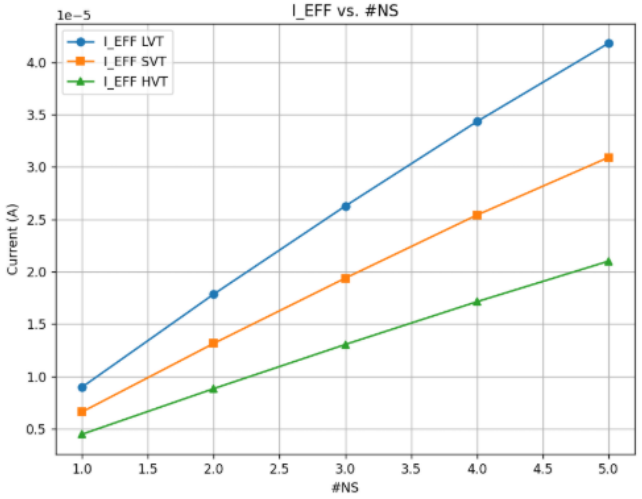


Fig. 6: Effective Currents for NMOS NSFETs

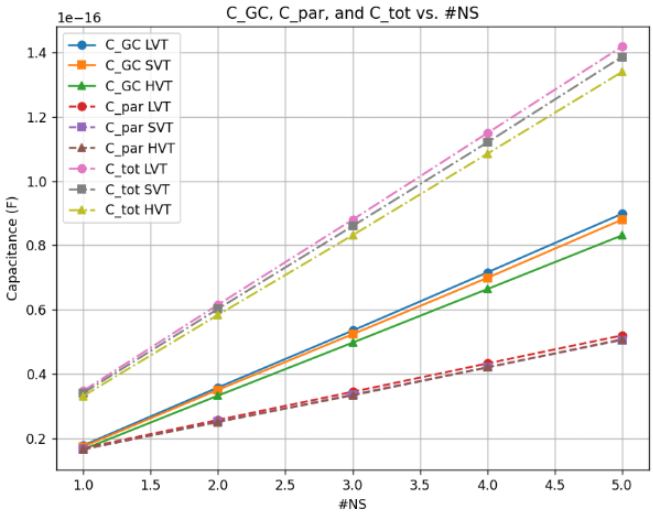


Fig. 7: Capacitance Breakdown for NMOS NSFETs

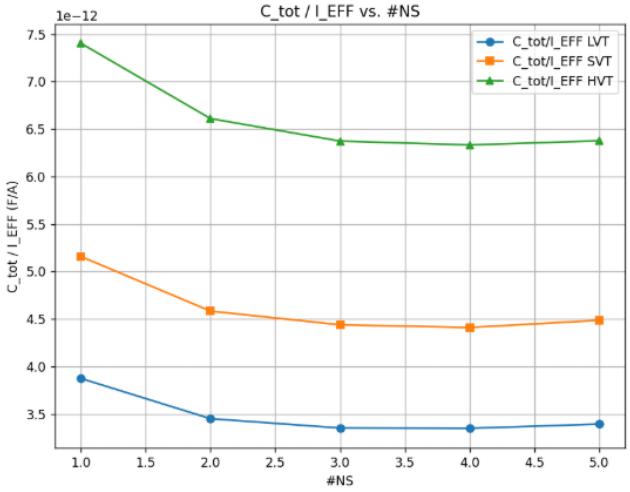


Fig. 8: $C_{GG,on}/I_{eff}$ for NMOS NSFETs

IV. RO Performance

To assess the performance of our optimized NSFETs, we will use a 31-stage ring oscillator (RO) with a fanout (FO) of 4. We will calculate the overall delay and energy consumption of the RO using NSFETs with #NS = 1-5, with and without considering wire capacitance. The following equations are used to calculate these performance metrics:

$$C_{tot} = FO * (C_{GC} + C_{par}) + C_{wire}$$

$$Delay = 31 * C_{tot} * \frac{V_{dd}}{I_{eff}}$$

$$E_{tot} = E_{dyn} + E_{stat}$$

$$E_{dyn} = 31 * 0.5 * C_{tot} * V_{dd}^2$$

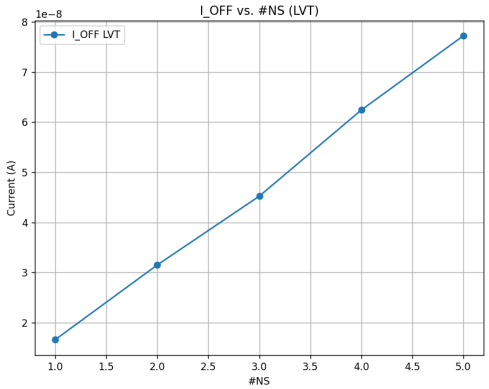
$$E_{stat} = 31 * I_{off} * V_{dd} * Delay$$

Without considering the wire capacitance, the RO delay and energy (and breakdown between dynamic and static) consumption for 1-5 nanosheet NSFETs are shown in Figure 13 through Figure 16. The shape of the delay curve illustrated in Figure 13 hearkens back to the $C_{GG,on}/I_{eff}$ curves in Figures 8 and 12, considering they all reach a minimum/plateau at around 3 or 4 nanosheets (depending on the threshold voltage). Although energy consumption increases somewhat linearly with the number of nanosheets, the $C_{GG,on}/I_{eff}$ curves in the last section serve as a fair delay optimization metric when examining the ideal case of zero wire capacitance.

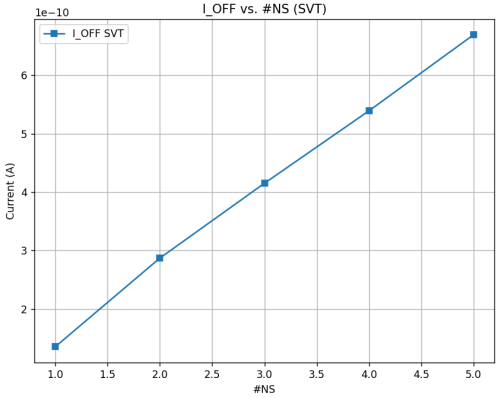
For an RO with wire capacitance considered, we will assume a metal wire capacitance per unit length of 278 aF/μm and a 3.9 μm long metal wire. This gives us a wire capacitance of $C_{wire} \approx 1.084$ fF, which, albeit an unrealistic value, will be used to analyze the case where total capacitance is dominated by the wire. The results of the RO delay and energy consumption calculations with wire capacitance are given in Figures 17 through 20.

The delay curves for the RO with wire capacitance (Figure 17) is markedly different than the delay without it (Figure 13). The $C_{wire} \neq 0$ curves are monotonically decreasing instead of plateauing, as we see in the $C_{wire} = 0$ curves and the $C_{gg,on}/I_{eff}$ plots (Figures 8 and 12). As a result, the $C_{gg,on}/I_{eff}$ curves are no longer a fair delay optimization metric as they were before. Since C_{tot} is dominated by C_{wire}, the optimization strategy should focus on maximizing I_{eff} instead of reducing C_{tot} to minimize the ratio $C_{tot}/I_{eff} \approx C_{wire}/I_{eff}$.

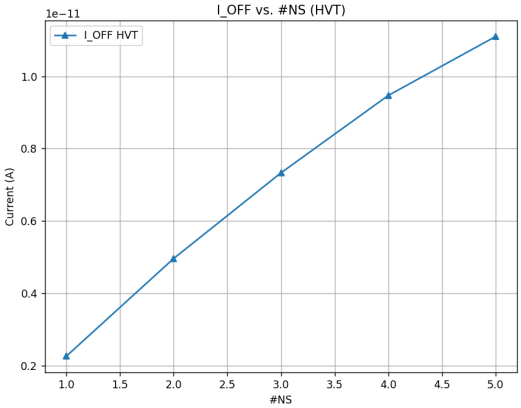
Under these assumptions, the optimal lengths (L_g, L_{sp}, L_c) would change. Since we want to maximize I_{eff} and no longer consider gate-to-channel and parasitic capacitance, we would choose a larger L_c to reduce contact resistance. We would also choose a smaller L_g and L_{sp} to reduce the resistance of the silicon channels (nanosheets). These alterations to the NSFET dimensions would serve to maximize I_{eff} without any regard for $C_{gg,on}$, allowing us to minimize C_{wire}/I_{eff}.



LVT



SVT



HVT

Fig. 9: Off-State Leakage Currents for PMOS NSFETs

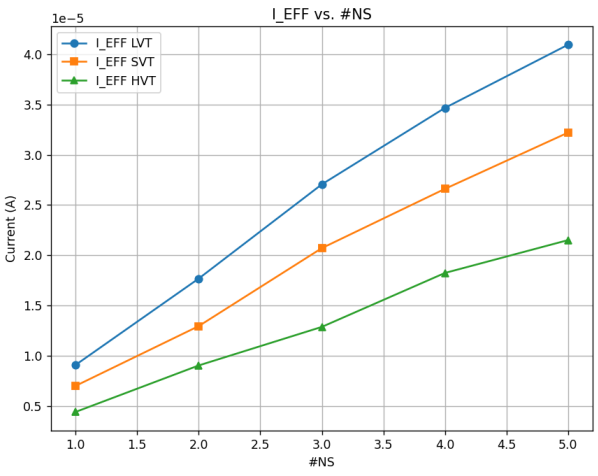


Fig. 10: Effective Currents for PMOS NSFETs

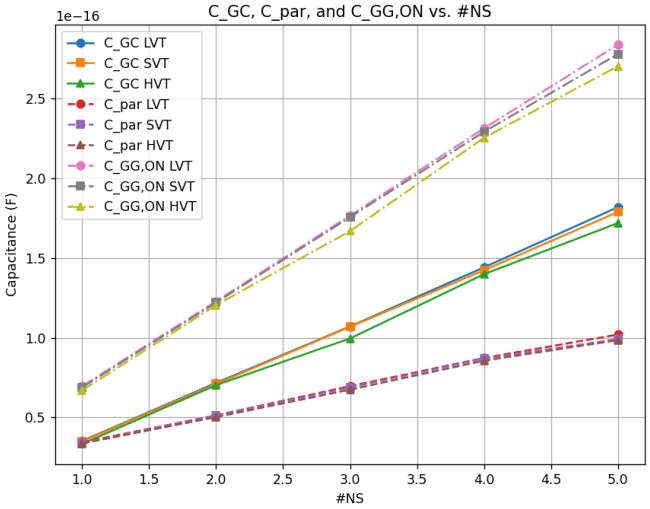


Fig. 11: Capacitance Breakdown for PMOS NSFETs

As previously discussed, the presence of a significant and overpowering wire capacitance drastically changes the optimal configuration of the NSFETs. Without the wire capacitance, it is necessary to balance a relatively small gate capacitance (parasitic and gate-to-channel) with a large effective current to maximize performance. However, a dominating wire capacitance renders attempts to minimize the gate capacitance

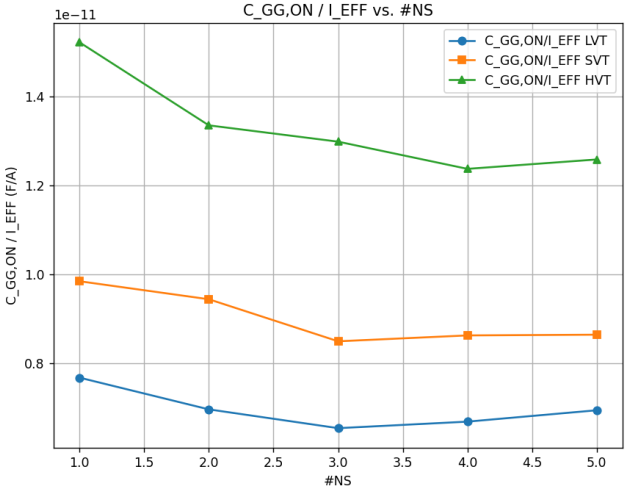


Fig. 12: $C_{GG,on}/I_{eff}$ for PMOS NSFETs

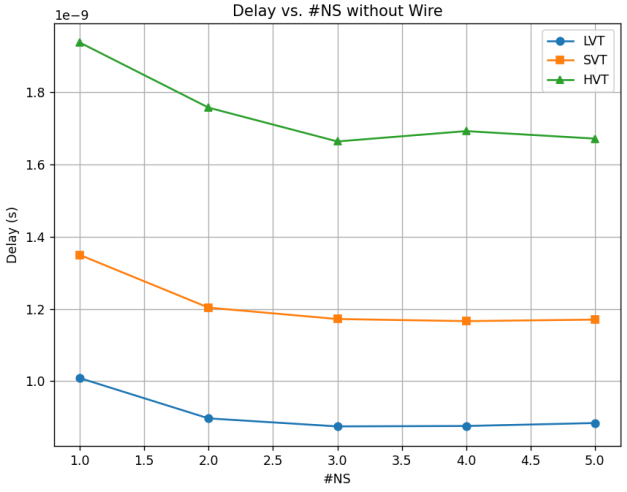


Fig. 13: RO Delay without Wire Capacitance

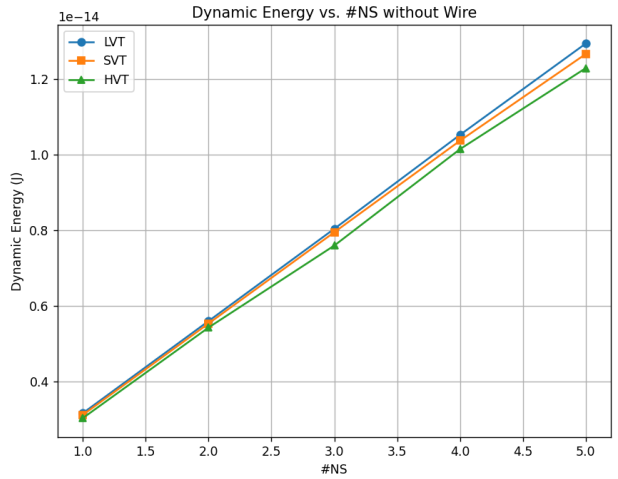


Fig. 14: Dynamic Energy of RO without Wire Capacitance

futile. Instead, the effective current of the NSFETs should be maximized to achieve optimal performance. Our analysis serves to highlight the importance of circuit-design context when optimizing the physical structure of NSFETs.

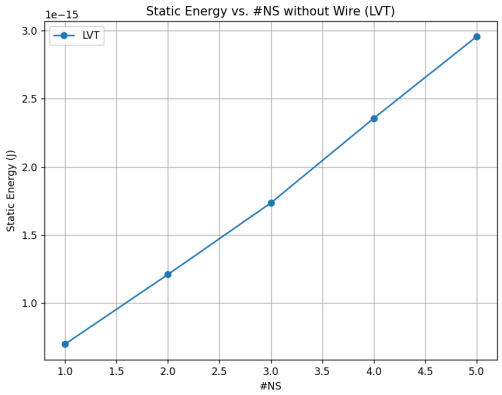
V. Conclusion

Although the simulations performed in this paper were integral to analyzing and optimizing the performance of our NSFET design, there are several limitations in both our simulation tools and our assumptions. As we have mentioned in a previous section, our simulations are limited by the physics models we use. For example, our simulations fail to take into account the effects of surface roughness scattering and Coulomb scattering. Additionally, using a 2D Schrödinger's equation solver would give a more accurate estimation of carrier distributions in the nanosheets at varying gate and drain voltages. Granted, this would drastically increase the simulation runtimes.

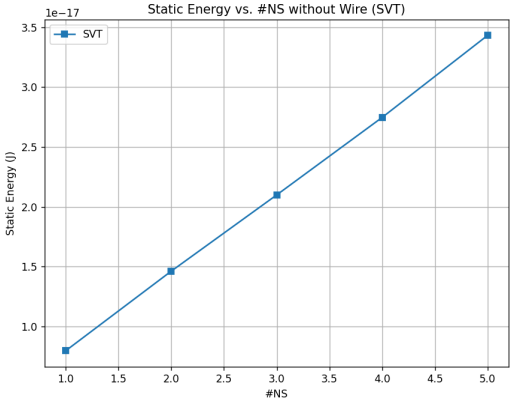
We also make an unrealistic assumption about the magnitude of the wire capacitance in the ring oscillator. In reality, the wire capacitance would not dominate the gate capacitance to the degree in the last section. Our analysis is largely limited by the length of the simulations, with a full simulation of a single nanosheet device taking approximately 1.5 hours. This runtime increases to around 6.5 hours when simulating a 5-NS device. Access to greater computing power would allow us to run more accurate simulations in a reasonable amount of time.

As a result of their lengthy runtimes, the bulk of this project was spent running and analyzing the simulations. However, we feel that there was enough time to run all of the simulations required between when the project was assigned and the final deadline. Having the meshing scheme given rather than found through experimentation made this possible, considering how much longer the simulations can run for with very fine meshes. For future offerings of this course, it would be worthwhile to have project check-ins a few times during the quarter. This would ensure that everyone has TCAD models and simulations, which would save time if someone faces convergence issues or overly long runtimes.

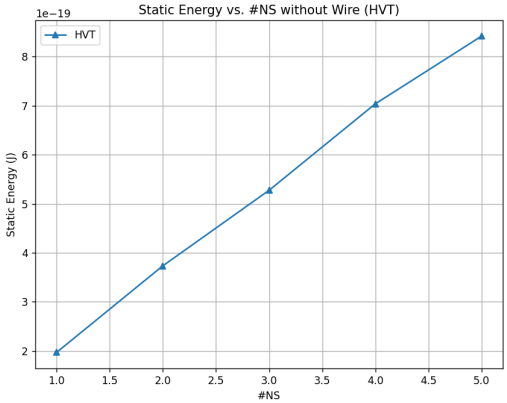
The analysis performed in this paper serves to illustrate how the optimal configuration for NSFETs changes depending on the desired applications and assumptions. It illustrates how the presence of significant wire capacitance drastically affects the optimization metric used to find the optimal dimensions of an NSFET. It also highlights the trade-off of adding nanosheets to increase drive current and reduce delay at the cost of greater energy consumption. Thus, lower-power devices would benefit from fewer nanosheets (1-2) at the cost of performance. Similarly, high-performance devices would benefit from more nanosheets, especially when significant interconnect capacitance is considered.



LVT



SVT



HVT

Fig. 15: Static Energy of RO without Wire Capacitance

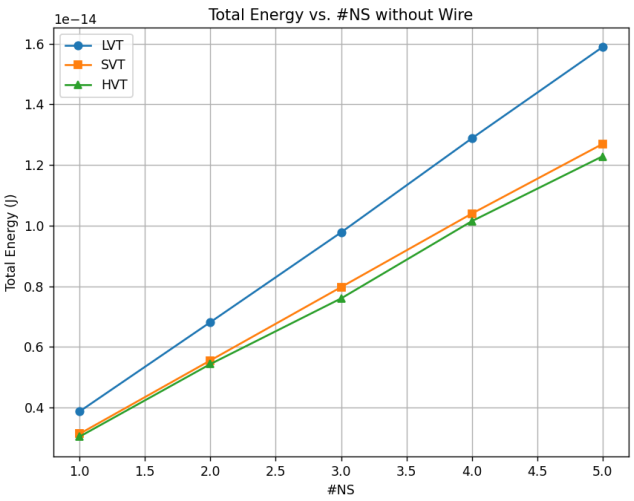


Fig. 16: Total Energy of RO without Wire Capacitance

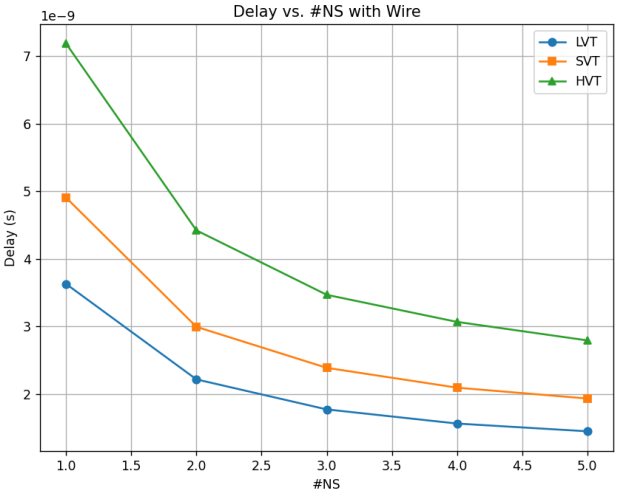
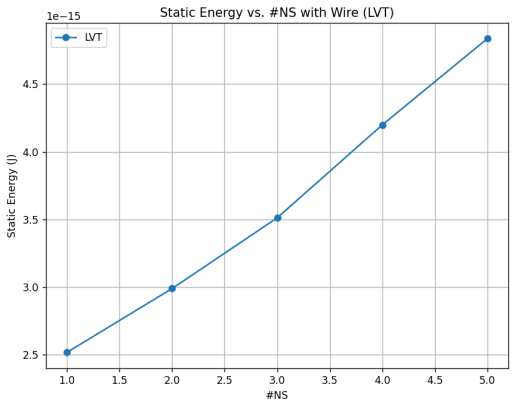


Fig. 17: RO Delay with Wire Capacitance

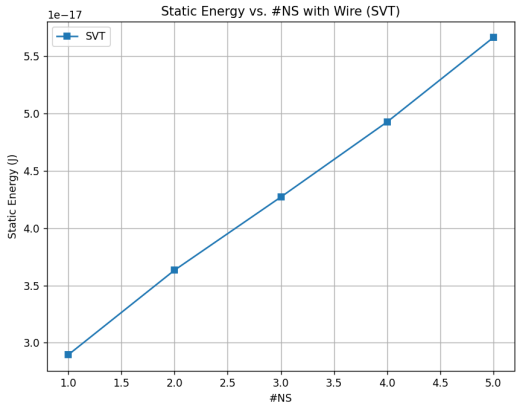
REFERENCES

[1] S. Maitra, "Study of the variation of the threshold voltage with the doping concentration and channel length," 2017 Devices for Integrated Circuit (DevIC), Kalyani, India, 2017, pp. 388-390, doi: 10.1109/DEVIC.2017.8073976.

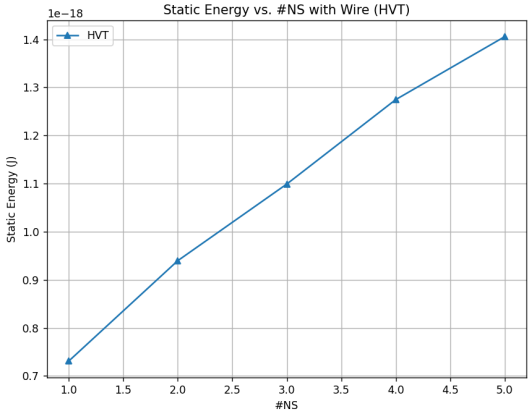
[2] K. Han, J. Lee, S. Tang, H. Maynard, N. Yoshida and A. Brand, "FinFET multi-Vt tuning with metal gate work function modulation by plasma doping," 2014 International Workshop on Junction Technology (IWJT), Shanghai, China, 2014, pp. 1-3, doi: 10.1109/IWJT.2014.6842043.



LVT



SVT



HVT

Fig. 18: Static Energy of RO with Wire Capacitance

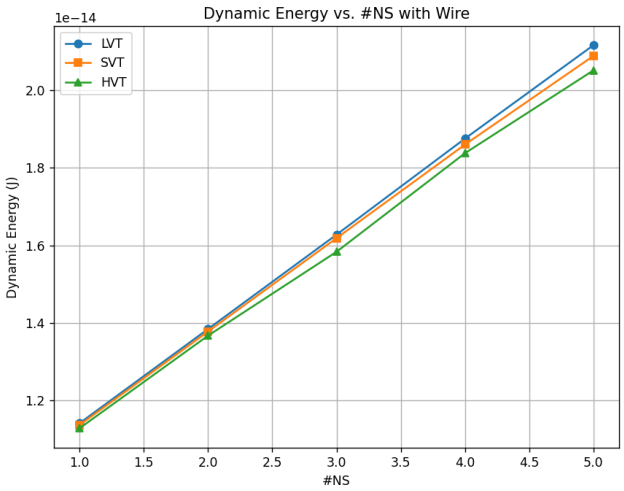


Fig. 19: Dynamic Energy of RO with Wire Capacitance

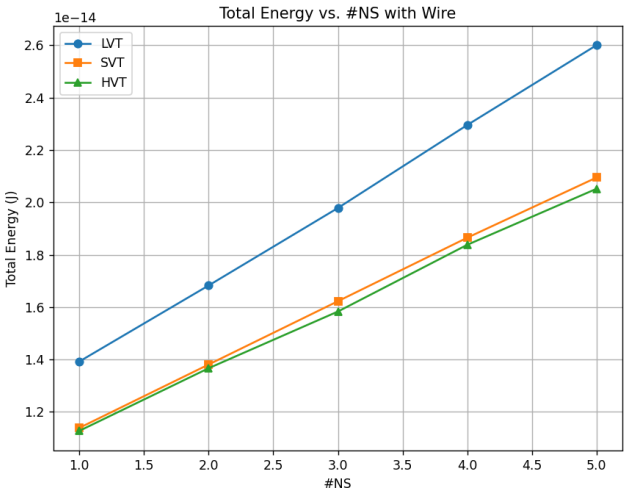


Fig. 20: Total Energy of RO with Wire Capacitance