



ESCUELA DE DOCTORADO
INTERNACIONAL DE LA USC

Julián
García Fernández

Tesis doctoral

Estudio estadístico de
fluctuaciones en transistores
MOSFET multipuerta
ultraescalados

Santiago de Compostela, 2023

TESIS DOCTORAL

ESTUDIO ESTADÍSTICO DE FLUCTUACIONES EN TRANSISTORES MOSFET MULTIPUERTA ULTRAESCALADOS

Autor

Julián García Fernández

Directores: Antonio Jesús García Loureiro y Natalia Seoane Iglesias

Tutor: Antonio Jesús García Loureiro

PROGRAMA DE DOCTORADO EN INVESTIGACIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN

SANTIAGO DE COMPOSTELA



DECLARACIÓN DEL AUTOR DE LA TESIS

Don Julián García Fernández

Título de la tesis: Estudio estadístico de fluctuaciones en transistores MOSFET multipuerta ultraescalados

Presento mi tesis, siguiendo el procedimiento adecuado al Reglamento, y declaro que:

- 1. La tesis abarca los resultados de la elaboración de mi trabajo.*
- 2. De ser el caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.*
- 3. Confirmando que la tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por mí para la obtención de otros títulos.*
- 4. La tesis es la versión definitiva presentada para su defensa y coincide la versión impresa con la presentada en formato electrónico.*

Y me comprometo a presentar el Compromiso Documental de Supervisión en caso de que el original no esté en la Escuela.

En Santiago de Compostela, Agosto 2023

Fdo. Julián García Fernández



AUTORIZACIÓN DEL DIRECTOR/TUTOR DE LA TESIS
Estudio estadístico de fluctuaciones en transistores MOSFET multipuerta
ultraescalados

Don Antonio Jesús García Loureiro, Catedrático del Área de Electrónica de la
Universidade de Santiago de Compostela

Doña Natalia Seoane Iglesias, Investigadora Post-doctoral Ramón y Cajal del Área de
Arquitectura y Tecnología de Ordenadores de la Universidade de Santiago de
Compostela

INFORMAN:

*Que la presente tesis, se corresponde con el trabajo realizado por **Don Julián García Fernández**, bajo nuestra dirección/tutorización, y autorizamos su presentación, considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC, y que como directores de esta no incurre en las causas de abstención establecidas en la Ley 40/2015.*

*De acuerdo con lo indicado en el Reglamento de Estudios de Doctorado, declaramos también que la presente tesis doctoral es idónea para ser defendida en base a la modalidad de **COMPENDIO DE PUBLICACIONES**, en los que la participación del doctorando fue decisiva para su elaboración y las publicaciones se ajustan al Plan de Investigación.*

En Santiago de Compostela, Agosto 2023

Fdo. Antonio Jesús García Loureiro
Director/a tesis

Fdo. Natalia Seoane Iglesias
Director/a tesis

*Ningún cañón borrará el surco de tu
arrozal.*

Víctor Jara

Agradecimientos

Me gustaría expresar mi agradecimiento a todas las personas que, aunque no aparezca su nombre explícitamente en los posteriores párrafos, me han ayudado en esta etapa.

En primer lugar, deseo extender mi más sincero agradecimiento a Natalia y Antonio por su inquebrantable apoyo a lo largo de esta travesía. Desde el punto de vista profesional, valoro enormemente su dedicación para formarme, brindarme orientación y enfrentar junto a mí cada desafío que se presentó. En el ámbito personal, no puedo expresar suficiente gratitud por el privilegio de haber conocido a dos personas tan excepcionales. La confianza que depositaron en mí y la puerta que abrieron para que pudiera adentrarme en el fascinante mundo de la investigación son regalos que atesoro con cariño. Además, su incansable esfuerzo por asegurarse de que me sintiera acogido y parte integral de su proyecto habla de su calidad humana. Mirando hacia adelante, estoy seguro de que no solo llevo conmigo aprendizajes valiosos, sino también la certeza de haber ganado dos buenos amigos en este camino.

Quiero agradecer a Enrique toda la ayuda que me ha ofrecido, así como las enseñanzas tan valiosas que me ha brindado. Nunca nadie encarnó mejor a Winston Wolfe de *Pulp Fiction*: “Soy el Señor Lobo, soluciono problemas”.

Igualmente, quiero expresar mi agradecimiento a Fernando, Alvarito, Rubén y Ferre por su valiosa ayuda. Sin sus consejos, sus cafés y su amistad, este trabajo tampoco habría sido posible.

Por último, y no menos importante, quería dar las gracias a mi madre por su ejemplo, su lucha diaria y su perseverancia, a María por estar a mi lado y apoyarme siempre, y a mi hermano Carlos del cual estoy orgulloso. También gracias a todos los buenos amigos que no he nombrado (ya saben quienes son) y que hacen mi día a día más feliz.

Agosto 2023

Otros agradecimientos

Gracias a Marc Bescond, a Guéric Ettisse y a todo el grupo NQS del IM2NP de la Aix-Marseille Univesité por su calurosa acogida.

Agradecer al CíTIUS por proporcionarme el soporte necesario para la realización del trabajo y también al CESGA por posibilitar el acceso a sus recursos computacionales.

Este trabajo ha recibido financiación de la *Consellería de Cultura, Educación e Universidade da Xunta de Galicia* del *Fondo Europeo de Desarrollo Regional (FEDER)* y del *Ministerio de España de Ciencia e Innovación* (acreditaciones RYC-2017-23312, PID2019-104834GB-I00, ED431F 2020/08, ED431C 2022/16).

Índice general

Resumo	1
Summary	9
1 Introducción	17
1.1. Motivación	18
1.2. Hipótesis y objetivos generales	20
2 Metodología	23
2.1. Modelado 3D	23
2.2. Simulación del dispositivo	25
2.3. <i>Machine Learning Figure of Merit Python extraction tool</i> (MLFoMpy)	34
3 Discusión	37
3.1. Calibración y descripción de los dispositivos	38
3.2. Estudio y comparación del impacto de diferentes fuentes de variabilidad	42
3.3. Desarrollo de modelos empíricos para predecir el impacto de la variabilidad	48
Conclusiones	69
Estudio de variabilidad para las tres arquitecturas	69
Desarrollo de modelos empíricos para la reducción de coste computacional	70
Trabajo futuro	72
Conclusions	75
Variability study for the three architectures	75
Development of empirical model for the reduction of the computational cost	76

Future work	78
Bibliografía	79
A Trabajos publicados	87
A.1. Artículos en esta tesis	87
A.2. Trabajos complementarios relacionados con esta tesis	106
Índice de figuras	107
Índice de tablas	111
Siglas	113

Resumo

O descubrimento dos transistores no século XX foi o detonante para o avance das tecnoloxías da información ata os nosos días, xa que estes son a base de tódolos sistemas de comunicación e computación actuais. A partir dese punto, os transistores de efecto campo (*Field Effect Transistor* (FET)) experimentaron unha redución considerable do seu tamaño, chegando a integrar nun microprocesador billóns destes (e.g. 2.5 billóns nun Intel Core i9-12900). O escalado destes dispositivos seguiu a chamada lei de Moore, ate que, numerosas dificultades de carácter físico apareceron por mor da súa pequena dimensionalidade. Para a superación destas dificultades ou limitacións o *IEEE International Roadmap for Devices and Systems* (IRDS) expón diversas arquitecturas FET, como poden ser os FinFET, os *nanowire FET* (NWFET) ou os *nanosheet FET* (NSFET). Entón, para o avance deste sector tecnolóxico é imprescindible o estudo destas arquitecturas ou de novos materiais para superar as limitacións derivadas do escalado.

Tamén é importante estudar de xeito preciso cales son as causas que producen estas limitacións. Por mor dos procesos de fabricación, aparecen non idealidades que se magnifican polas pequenas dimensións que acadan algunhas rexións destes dispositivos. Estas desviacións do funcionamento ideal do transistor son derivadas da aparición de diferentes fontes de variabilidade procedentes dos procesos de deposición metálicas, litográficas, etc. O estudo destas fontes de variabilidade é imprescindible para avaliar as arquitecturas máis inmunes a estas, co fin de garantir a homoxeneidade na resposta dos dispositivos e aproveitar ás súas heteroxeneidades para aplicacións específicas como a criptografía ou a xeración de números aleatorios.

Debido a que os estudos estatísticos de variabilidade precisan da avaliación dun gran número de dispositivos (≥ 300) a metodoloxía clásica de fabricación e análise é tremendamente custosa. Ademais a implementación de arquitecturas que se atopan máis alá do estado da arte é

un elevado reto desta metodoloxía debido á dificultade ou imposibilidades da súa fabricación. Por iso, para estudar o impacto da variabilidade úsase xeralmente o deseño asistido por ordenador (*Technology Computer Aided Design (TCAD)*). Aínda co uso de esta metodoloxía, o estudo do impacto da variabilidade ten un elevado custo computacional por dous motivos principais:

- i) A necesidade de simulacións tridimensionais que teñan en conta os efectos do confinamento cuántico para estudar e comprender mellor as propiedades e comportamentos destes dispositivos de dimensión nanométricas.
- ii) O requirimento da simulación de grandes conxuntos de dispositivos con diferentes configuracións para obter resultados estatisticamente significativos. Ao simular unha gran variedade de configuracións, pódese obter unha comprensión máis completa do seu funcionamento medio e as posibles variacións.

Por estes motivos, aparece o interese do uso de técnicas complementarias ao TCAD que permitan estima-lo impacto da variabilidade dun xeito máis eficiente, reducindo o tempo de cálculo necesario. Estas técnicas complementarias poden nutrirse da gran cantidade de datos e información que xeran as simulacións TCAD, e que precisan ser correctamente post-procesados. Co obxectivo de realizar un correcto post-procesado de gran cantidade de datos nesta tese desenvólvese o *software* MLFoMpy.

Co contexto mostrado previamente, este traballo tratará de dar resposta aos diferentes retos actuais coas seguintes hipóteses:

- H1.** *A través dunha análise do impacto de diversas fontes de variabilidade nas arquitecturas propostas, esperase que o NSFET exhiba unha maior robustez en comparación co NWFET, sendo este o candidato preferente para substituír ao FinFET nos futuros nodos tecnolóxicos.*
- H2.** *O desenvolvemento de modelos empíricos pode reducir os custos computacionais vinculados á realización de estudos do impacto da variabilidade mediante a metodoloxía TCAD, o que potencialmente contribuirá a unha maior eficiencia na investigación do sector tecnolóxico ao reducir o tempo e os recursos necesarios.*

Para desenvolver estas hipóteses co fin de asegurar a reproducibilidade, mostramos unha descrición exhaustiva das diferentes metodoloxías empregadas para a simulación destes dispo-

sitivos semicondutores no Capítulo 2. Comezarse pola descrición do modelado tridimensional, que inclúe o deseño, mallado e asignación de características físicas. O proceso de deseño e mallado lévase a cabo co *software* Gmsh, que permite crear unha malla de elementos finitos adaptada á xeometría do dispositivo. A ferramenta de pre-procesado pVENDES encargase da asignación das características físicas aos diferentes elementos da malla, o que permite simular o comportamento do dispositivo de xeito realista.

Unha vez o dispositivo estea modelado, procédese á simulación dos procesos físicos e fenómenos de transporte de portadores. Para iso emprégase o *software* propio *Variability Enabled Nanometric DEvice Simulator* (VENDES), que é un simulador de dispositivos semicondutores baseado nos métodos de arrastre e difusión (*Drift-diffusion* (DD)) e no método de transporte de Monte Carlo (MC). O método DD baséase no potencial electrostático calculado pola ecuación de Poisson e emprega as ecuacións de continuidade para calcular a densidade de corrente. Este método é adecuado para describir o funcionamento transistor na rexión sublimiar debido á boa reprodución das dinámicas de portadores a baixos campos eléctricos, pero presenta problemas para representar os efectos fóra do equilibrio propios da rexión de acendido. O outro método de transporte baséase na resolución da ecuación de transporte de Boltzmann a través do método de MC acoplado coa ecuación de Poisson. En esencia este método consiste na simulación do movemento dun conxunto de portadores que voan libremente seguindo as ecuacións de Newton e sufrindo procesos ou eventos de dispersión cunha probabilidade de suceso determinada pola mecánica cuántica en función de cada material. Este método semiclásico de transporte é adecuado para describir o funcionamento do transistor na rexión de acendido debido á súa capacidade para representar os efectos fóra do equilibrio. Por tanto, ámbolos dous métodos de transporte teñen vantaxes e desvantaxes, sendo precisos en diferentes rexións. Tamén é importante comentar que MC é computacionalmente máis custoso có DD.

Por mor das necesidades da industria e ao escalado dos dispositivos, requírese a simulación de dispositivos de tamaño nanométrico, onde as súas dimensións son comparable coa lonxitude de onda dos electróns. Isto implica ter en conta diferentes efectos cuánticos que afectan ao transistor e polo tanto, xurde a necesidade de aplicar correccións cuánticas (*Quantum Corrections* (QC)). En VENDES pódense aplicar dous tipos de correccións cuánticas: gradiente de densidade (*Density Gradient* (DG)) e a ecuación 2D de Schrödinger (2D-SCH). DG corrixe cuánticamente o potencial a partir de diferentes parámetros de axuste como por exemplo o tensor de masas electrónicas efectivas. O principal problema do DG é a necesidade de calibración

fronte a datos experimentais ou fronte a técnicas máis complexas como a *Non-Equilibrium Green's Functions* (NEGF). Por outra banda, no caso de 2D-SCH, resólvese a ecuación 2D de Schrödinger en láminas transversais do dispositivo para calcular a densidade cuántica eléctrica e interpólase ao dominio 3D para logo calcular o potencial corrixido cuánticamente.

Unha vez explicados os diferentes modelos e correccións cuánticas utilizados para a simulación dos dispositivos, explícase a implementación das fontes de variabilidade estudadas. Neste traballo estudaranse as seguintes: granularidade metálica (*Metal Grain Granularity* (MGG)), rugosidade superficial (*Line Edge Roughness* (LER)), rugosidade de porta (*Gate Edge Roughness* (GER)) e fluctuacións por dopantes aleatorios (*Random Discret Dopants* (RDD)). O MGG débese ao proceso de deposición metálica da porta, onde aparecen diferentes orientacións metálicas (grans) con diferentes funcións de traballo. Para reproducir o MGG, xéranse perfiles de granularidade aleatoria usando diagramas de Poisson-Voronoi dependendo do tamaño medio do gran (*Grain Size* (GS)). Tanto o LER coma o GER son causa dos procesos litográficos na fabricación dos transistores. O LER refírese ás rugosidades que aparecen na contorna do semiconductor estrangulando ou ensanchando a canle de condución. Por outra banda, o GER é unha rugosidade que se produce sobre a porta, nas fronteiras entre fonte/porta e porta/drenador. En ámbolos dous casos, os perfiles de deformación aleatoria xéranse a través da función de autocorrelación, que é definida a partir da transformada de Fourier do espectro gaussiano. Estas deformacións descríbense con dúas magnitudes, a raíz cuadrática media da deformación (*Root Mean Square height* (RMS)) e a lonxitude de correlación (*Correlation Length* (CL)). A última fonte de variabilidade implementada neste traballo é a RDD, que consiste na aparición de dopantes aleatorios que xeran fluctuacións no potencial provocando unha diferenza no comportamento ideal do transistor.

Unha vez explicada a metodoloxía, móstranse os resultados obtidos nesta tese no Capítulo 3. A discusión consistirá primeiro na explicación da calibración e descrición dos dispositivos, seguido do estudo e comparativa do impacto das diferentes fontes de variabilidade en relación coa hipótese **H1**, e por último o desenvolvemento de modelos empíricos para predicir o impacto da variabilidade co fin de constatar a hipótese **H2**.

O proceso de calibración é esencial para garantir a veracidade das simulacións. Neste traballo a calibración realizarase para as tres arquitecturas estudadas (FinFET, NWFET e NSFET), tanto usadas na actualidade como proxectadas polo IRDS para a súa utilización nos próximos nodos tecnolóxicos. A idea é calibrar as características I_D - V_G (sendo I_D a corrente de drenaxe e V_G a tensión na porta) dos dispositivos ideais fronte a datos experimentais

para logo poder realizar os estudos do impacto da variabilidade. Co fin de realizar unha comparación xusta entre as tres arquitecturas, fixéronse diferentes parámetros de deseño dos transistores: mesma lonxitude de porta (L_g), mesma lonxitude total (L), mesma anchura efectiva do óxido (t_{ox}) e mesmo perfil de dopado gaussiano tipo-n. Todos os dispositivos foron calibrados respecto a datos experimentais ou escalados directamente de transistores experimentais calibrados en traballos previos. Móstrase nesta sección a incapacidade do modelo de transporte MC para reproducir con exactitude o comportamento do transistor na rexión sublimiar debido ao elevado ruído estatístico derivado dos pequenos valores de I_D a baixos valores de V_G . Pola contra, froito do exhaustivo proceso de calibración dos diferentes parámetros do modelo DD, conséguense reproducir as características I_D - V_G dos transistores tanto na rexión sublimiar como na rexión de acendido.

Unha vez están calibrados os transistores, procederemos ao estudo e comparación do impacto das diferentes fontes de variabilidade sobre eles, co fin de observar cal das diferentes arquitecturas é máis resistente a elas. Para poder realizar a comparación entre elas, empregáronse as principais figuras de mérito (*Figures of Merit* (FoM)) definidas a continuación: a tensión limiar (V_{Th}), que indica o potencial necesario para establecer unha canle entre fonte e drenador, cambiando así o estado do transistor de modo apagado a modo acendido; a pendente sublimiar (SS), que defínese como a inversa da pendente na rexión de crecemento lineal das características I_D - V_G do transistor (correspondente á rexión sublimiar); a corrente de apagado (I_{off}), que é a corrente na cal o potencial da porta é nulo ($V_G=0$ V); e a corrente de acendido (I_{on}), que corresponde coa corrente a $V_G = V_D + V_{Th}$, sendo V_D a tensión de alimentación.

Nun primeiro punto, estudamos unicamente o impacto das diferentes fontes de variabilidade sobre a rexión sublimiar das características I_D - V_G dos dispositivos, centrándonos principalmente nas desviacións estándar da tensión limiar (σV_{Th}) e da pendente sublimiar (σSS), xa que estas magnitudes darnos información sobre o impacto destas fontes. Para cada conxunto de parámetros de variabilidade simuláronse conxuntos de máis de 300 dispositivos afectados polos perfís de aleatorios mencionados previamente para cada arquitectura.

Para avaliar o impacto do MGG estudáronse catro valores diferentes de tamaño medio de gran (GS) para cada arquitectura. Con isto, obsérvase que para σV_{Th} tanto o NSFET coma o FinFET son máis resistentes a esta fonte de variabilidade que o NWFET. Isto é debido a que esperase que os dispositivos con menor área de porta (A_G) presenten maior impacto ao MGG e polo tanto maior σV_{Th} co mesmo GS. Neste caso, para un GS de 3 nm σV_{Th} é un 59% maior para o NWFET ($A_G \approx 395$ nm²) que a σV_{Th} do FinFET e do NSFET ($A_G > 1000$ nm²).

Cabe destacar que tanto o FinFET como o NSFET exhiben unha resposta practicamente idéntica, sendo polo tanto igualmente susceptibles á variabilidade MGG. Tamén obsérvase unha dependencia lineal da σV_{Th} co GS. Outro efecto observado é a saturación da σV_{Th} para o NWFET cando o GS é grande, producida pola escaseza de grans presentes na porta metálica, debido a pequena A_G que ten este dispositivo. Esta aparición de poucos grans metálicos na porta provoca que a distribución estatística de V_{Th} siga un patrón bimodal con picos ao redor dos valores extremos de función de traballo en vez dunha distribución normal. Esta saturación é aínda máis evidente cando se avalía o σSS . No caso do FinFET e o NSFET as distribucións de V_{Th} e SS seguen un patrón normal para todos os valores de GS.

Para estudar o impacto do LER nas tres arquitecturas realizouse a simulación de conxuntos de dispositivos con perfíles de deformación aleatoria para tres valores de RMS e un valor de CL. A influencia desta fonte de variabilidade sobre o NSFET é insignificante sendo a σV_{Th} da orde de 12 e 19 veces máis pequeno que os valores mostrados por FinFET e NWFET respectivamente. Como xa explicamos previamente, o LER indúcese polo proceso litográfico e debido aos procesos de fabricación, esta deformación ten lugar nas dimensións críticas do FinFET e o NWFET, non sendo así no caso do NSFET, o que fai que esta arquitectura sexa tremendamente resistente a esta fonte de variabilidade. Por outro lado, aínda que a dimensión crítica é un pouco menor no caso do FinFET, esta arquitectura é máis resistente que o NWFET, debido a que a súa maior altura implica maior área de condución, reducindo de ese modo o impacto potencial do LER e explicando os menores valores de σV_{Th} e σSS . Cabe destacar que a σSS é maior que a producida pola variabilidade MGG, debido a que o LER actúa sobre a rexión de crecemento lineal das características I_D-V_G mentres que o MGG despraza a curva I_D-V_G dos dispositivos proporcionalmente á función de traballo media da porta.

A variabilidade inducida polo GER é menor se a comparamos co MGG e LER. Neste caso o FinFET é o dispositivo menos perturbado por ista fonte de variabilidade sendo σV_{Th} da orde dun 20 % menor que para NSFET e NWFET. Despois de analizar por qué o FinFET presenta menor variabilidade ca o NSFET cando son dispositivos con área e perímetro de porta similar, viuse que as deformacións localizadas nas rexións inferiores laterais da porta do FinFET (que non envolven ao semiconductor) non teñen impacto sobre o canal de condución, sendo esta parte da porta o 9 % do perímetro total.

Os resultados do impacto do RDD mostran que ao ter tódolos dispositivos o mesmo perfil de dopado, presentan á mesma resistencia a esta fonte de variabilidade sendo $\sigma V_{Th} \sim 8$ mV.

De xeito análogo ao estudo da rexión sublimiar, tamén realizouse unha comparativa na

rexión de acendido a través da avaliación da σI_{on} inducida pola dúas fontes de variabilidade que máis impacto tiñan no comportamento sublimiar, o MGG e o LER. Na rexión de acendido o impacto do MGG sobre σI_{on} é moi similar, sendo o NWFET a arquitectura máis resistente e presentando unha sensibilidade similar o NSFET e o FinFET. Para o LER tamén observase o caso de influencia mínima da variabilidade sobre a σI_{on} do NSFET. O NWFET segue sendo a arquitectura menos resistente cunha σI_{on} un 15 % maior que para o FinFET, unha diferenza menor que no caso sublimiar.

Unha vez avaliado o impacto das fontes de variabilidade nas arquitecturas que se atopan no estado da arte ou son candidatas para os futuros nodos tecnolóxicos, podemos concluír tendo en conta unicamente a resistencia de cada arquitectura, que o NSFET sería a arquitectura máis adecuada para substituír aos FinFETs, principalmente pola súa elevada resistencia ao LER. Con todo isto, abordaría-se todo o correspondente á hipótese **H1** deste traballo.

Co obxectivo de reducir o custo computacional da metodoloxía TCAD nos estudos de variabilidade e tratar deste modo a hipótese **H2**, desenvolveuse no contexto desta tese un modelo empírico baseado na lei de Pelgrom (*Pelgrom-based predictive model (PBP)*) para a predición das variacións inducidas polas dúas fontes de variabilidade que maior impacto teñen no funcionamento dos dispositivos deste traballo (MGG e LER). Estes modelos son validados cunha combinación de resultados extraídos de publicación de diferentes autores e simulacións propias realizadas en VENDES. O modelo PBP desenvólvese por separado para a predición de σV_{Th} e σI_{on} que corresponden coas simulación na rexión sublimiar (QC DD) e na rexión de acendido (QC MC), respectivamente, sendo en ambos os dous casos aplicado para MGG e LER.

Ao aplicar o modelo PBP sobre a rexión sublimiar obtemos que os erros de predición son menores ou iguais ao 10 % respecto aos resultados dos estudos TCAD, sempre que os dispositivos se atopan dentro dos límites de validez do modelo. Estes límites no modelo PBP sublimiar débense á saturación da σV_{Th} producida pola variabilidade MGG cando existe escaseza de grans na porta (grandes valores de GS con dispositivos con área de porta pequena). Este efecto provoca a ruptura da linearidade asumida no modelo entre a σV_{Th} e o GS, destacar tamén que no caso do modelo aplicado ao LER non existe ningún tipo de limitación. Os resultados para o modelo PBP sublimiar aplicado ao MGG obtemos que un 91 % dos erros relativos son menores iguais do 8 %, mentres que no modelo aplicado ao LER tódalas estimacións están dentro da marxe do 10 %, estando o 94 % das predicións por debaixo do 6 %. Estes resultados demostran a adaptabilidade do modelo PBP sublimiar para unha

gran variedade de arquitecturas, dimensións, seccións transversais (tres formas diferentes de FinFETs e dúas diferentes para o NWFET dependentes da variedade de deseños reportados na literatura polos diversos autores) e parámetros de variabilidade (GS para MGG, RMS e CL para LER).

No caso da utilización do modelo PBP aplicado sobre a rexión de acendido sobre as tres arquitecturas estudadas, obtemos erros relativos para a σI_{on} menores do 8 % no 92 % dos casos combinando as predicións realizadas para a estimación do impacto do MGG e do LER. Do mesmo modo que no modelo sublimiar, as limitacións do modelo PBP prodúcense no caso da variabilidade MGG cando existe escaseza de grans na porta metálica. De novo tódalas estimación da σI_{on} debida ao LER están por debaixo do requirido 10 %.

Á hora de avaliar a redución do custo computacional, observase un beneficio claro na aplicación do modelo PBP. Por exemplo, na rexión de acendido para a realización de 9 estudos do impacto da variabilidade MGG sobre un NWFET, aforraríase ao redor dun 90 % respecto ao custo da metodoloxía TCAD. Esta estimación de aforro calcúlase asumindo como estudo da variabilidade MGG as simulacións necesarias para un dispositivo específico con 4 valores GS diferentes.

Co mostrado previamente demostrase que o modelo PBP aplicado á rexión sublimiar e á rexión de acendido proporciona unha estimación simple, rápida e precisa do impacto de dúas das fontes que máis contribúen á variabilidade de transistores con diferentes arquitecturas e dimensións. Polo tanto, o modelo PBP é unha nova estratexia que axuda á redución do custo computacional dos estudos de variabilidade TCAD abordando deste xeito a hipótese **H2** deste traballo.

En vista dos resultados obtidos, podemos concluír que se acadaron tódolos obxectivos desta tese. Desenvolvéronse as dúas hipóteses, abordando problemas relevantes do estado de arte como son a determinación da arquitectura do transistor de efecto campo baseada en silicio máis prometedora para os futuros nodos tecnolóxicos a través do estudo do impacto da variabilidade, e o desenvolvemento de estratexias precisas para a redución do custo computacional da tecnoloxía TCAD.

Summary

The discovery of transistors in the 20th century was the catalyst for the advancement of information technologies to our present day, as they form the basis of all modern communication and computing systems. From that point, *Field Effect Transistor* (FET) experienced a significant reduction in size, with billions of them integrated into a microprocessor (e.g. 2.5 billion in an Intel Core i9-12900). The scaling of these devices followed the Moore's Law until numerous physical challenges emerged due to their small dimensions. To overcome these difficulties or limitations, the *IEEE International Roadmap for Devices and Systems* (IRDS) proposes various FET architectures, such as the FinFET, the *nanowire FET* (NWFET), or the *nanosheet FET* (NSFET). Therefore, for the advancement of this technological sector, studying these architectures or new materials to overcome scaling limitations is essential.

It is also important to accurately study the underlying causes of these limitations. Due to the manufacturing processes, non-idealities arise and are amplified by the small dimensions reached in some regions of these devices. These deviations from the expected transistor behaviour appear from a range of variability sources that result from processes like metal deposition, lithography, and more. Investigating these sources of variability is essential to assess which architectures are more immune to them, aiming to ensure uniformity in device response, or to use their heterogeneities for specific applications as cryptography or random number generation.

As statistical variability studies require the assessment of a considerable number of devices (≥ 300), the traditional approach to fabrication and analysis becomes financially prohibitive. Furthermore, implementing architectures beyond the state-of-the-art is a significant challenge within this methodology due to the difficulties or impossibilities of their manufacturing. Therefore, to investigate the impact of variability, *Technology Computer Aided Design* (TCAD) is generally employed. Even with the use of this methodology, studying variability's impact

carries a significant computational cost for two main reasons:

- i) The requirement for three-dimensional simulations that consider quantum confinement effects is essential to thoroughly investigate and comprehend the properties and behaviours of these nanoscale devices.
- ii) The simulation of large sets of devices with different configurations is crucial to obtain statistically meaningful results. By simulating a wide range of configurations, a more comprehensive understanding of their average performance and potential variations can be achieved.

For these reasons, there is an interest in using complementary techniques to TCAD to efficiently estimate the impact of variability, thus reducing the required computation time. These supplementary techniques can take advantage of the vast amount of data and information generated by TCAD simulations, which need to be properly post-processed. With the aim of performing accurate post-processing of extensive data sets, this thesis introduces the software *Machine Learning Figure of Merit Python extraction tool* (MLFoMpy).

With the context previously presented, this work aims to address the current challenges with the following hypotheses:

- H1.** *Through an analysis of the impact of various sources of variability on the proposed architectures, it is expected that the NSFET will exhibit greater robustness compared to the NWFET, making it the most suitable candidate for replacing FinFET in future technological nodes.*
- H2.** *The development of empirical models can decrease the computational costs associated with performing variability studies using the TCAD methodology, potentially enhancing efficiency in technological sector research by reducing the required time and resources.*

To develop these hypotheses and ensure reproducibility, in Chapter 2 we provide a comprehensive description of the different methodologies employed for simulating these semiconductor devices. We will start by describing the three-dimensional modelling, which involves design, meshing, and allocation of physical characteristics. The design and meshing process is carried out using the software Gmsh, which enables the creation of a finite element mesh tailored to the device geometry. The pre-processing tool pVENDES is responsible for assigning physical characteristics to the different elements of the mesh, enabling the realistic simulation of the device behaviour.

Once the device is modelled, the simulation of the relevant physical processes is carried out. For this purpose, the software *Variability Enabled Nanometric DEvice Simulator* (VENDES) is employed, which is a semiconductor device simulator based on the *Drift-diffusion* (DD) and Monte Carlo (MC) transport methods. The DD method solves the Poisson equation to obtain the electrostatic potential and the continuity equations to obtain the current density. This method is suitable for describing transistor operation in the subthreshold region because it effectively captures the behaviour of carriers under low electric fields. However, it faces challenges in representing the non-equilibrium effects characteristic of the transistor's on region.

The other transport method is based on solving the Boltzmann transport equation using the Monte Carlo (MC) method coupled with the Poisson equation. In essence, this method entails simulating the motion of a group of carriers that experience free flights governed by Newton equations, while also encountering scattering events with probabilities influenced by the quantum mechanics inherent to each material. This semi-classical transport method is suitable for describing transistor operation in the on region due to its capability to represent non-equilibrium effects. Therefore, both transport methods have advantages and disadvantages, being accurate in different regions of the transistor. Additionally, it is worth highlighting that MC is computationally more expensive than DD.

With the demands of the industry and the ongoing device miniaturization, it appears the necessity to simulate nanoscale devices, characterized by dimensions that are in the order of the electron's wavelength. This implies the inclusion of the different quantum effects that affect the transistor's behaviour, via quantum corrections (QC). Two types of quantum corrections can be applied in VENDES: density gradient (DG) and the 2D Schrödinger equation (2D-SCH). DG quantum corrects the potential using various fitting parameters such as the effective mass tensor. The main challenge of DG is the requirement for calibration against experimental data or more complex techniques like *Non-Equilibrium Green's Functions* (NEGF). On the other hand, in the case of 2D-SCH, the 2D Schrödinger equation is solved in transverse slices to the transport direction of the device to calculate the electrical quantum density, which is then interpolated into the 3D domain to calculate the quantum corrected potential.

After explaining the different models and quantum corrections used for device simulation, the implementation of the studied sources of variability is described. In this work, the following sources will be studied: *Metal Grain Granularity* (MGG), *Line Edge Roughness* (LER), *Gate Edge Roughness* (GER), and *Random Discret Dopants* (RDD). The MGG is due

to the gate metal deposition process, where different metal orientations (grains) with different work functions appear. To reproduce the MGG, random granularity profiles are generated using Poisson-Voronoi diagrams depending on the average grain size (*Grain Size (GS)*). Both LER and GER are caused by lithographic processes during transistor manufacture. LER refers to the roughness that appears at the semiconductor contour, constricting or widening the channel. On the other hand, GER is a roughness that occurs on the gate, at the boundaries between source/gate and gate/drain. In both cases, the random deformation profiles are generated through the autocorrelation function, which is defined from the Fourier transform of the gaussian spectrum. These deformations are characterized by two quantities, the *Root Mean Square height (RMS)* and the *Correlation Length (CL)*. The last source of variability implemented in this work is RDD, which consists of the appearance of random dopants that generate fluctuations in the potential, causing a difference in the ideal behaviour of the transistor.

Once the methodology has been explained, the results obtained in this thesis are presented in Chapter 3. The discussion will first involve the calibration and description of the devices, followed by the study and comparison of the impact of different sources of variability in relation to hypothesis **H1**. Finally, empirical models will be developed to predict the impact of variability, aiming to confirm hypothesis **H2**.

The calibration process is crucial to ensure the accuracy of the simulations. In this work, the calibration will be carried out for the three studied architectures (FinFET, NWFET, and NSFET), which are currently used as well as projected by the IRDS for upcoming technological nodes. The goal is to calibrate the I_D - V_G characteristics (where I_D represents drain current and V_G is gate voltage) of ideal devices against experimental data in order to perform variability impact studies. To ensure a fair comparison among the three architectures, different transistor design parameters have been set: same gate length (L_g), same total length (L), same effective oxide width (t_{ox}), and the same profile of n-type gaussian doping. All devices were calibrated based on experimental data or were scaled directly from previously calibrated experimental transistors. Also, it is shown that the MC transport model might struggle to accurately reproduce the transistor behaviour in the subthreshold region due to the high statistical noise derived from the small values of I_D at low values of V_G . In contrast, as a result of the exhaustive calibration process of the different parameters of the DD model, the I_D - V_G characteristics of the transistors are reproduced both in the subthreshold region and in the on region.

Once the transistors are calibrated, we will proceed with the study and comparison of the

impact of different sources of variability on them, aiming to observe which of the different architectures is more resilient to them. In order to make the comparison, we use the following *Figures of Merit* (FoM): the threshold voltage (V_{Th}), which indicates the potential required to establish a channel between the source and drain, thereby transitioning the transistor from an off-state to an on-state; the subthreshold slope (SS), defined as the inverse of the slope in the linear growth region of the I_D - V_G characteristics (corresponding to the subthreshold region); the off current (I_{off}), which is the current when the gate potential is zero ($V_G=0$ V); and the on current (I_{on}), which corresponds to the current at $V_G = V_D + V_{Th}$, being the V_D the supply voltage.

In a first step, we study only the impact of the different sources of variability on the subthreshold region of the I_D - V_G characteristics of the devices, focusing mainly on the standard deviations of the threshold voltage (σV_{Th}) and the subthreshold slope (σSS). For each variability parameter, sets of over 300 devices were simulated, taking into account the previously mentioned random profiles for each architecture.

To assess the impact of MGG, four different mean grain sizes (GS) will be studied for each architecture. From this analysis, it can be observed that both NSFET and FinFET are more resistant to this source of variability than the NWFET in terms of σV_{Th} . This is attributed to the idea that devices with smaller gate area (A_G) will exhibit a greater impact from MGG, resulting in higher σV_{Th} for the same GS. For instance, at a GS of 3 nm, σV_{Th} is approximately 59 % higher for the NWFET ($A_G \approx 395 \text{ nm}^2$) than for the FinFET and the NSFET ($A_G > 1000 \text{ nm}^2$). It is remarkable that both the FinFET and NSFET exhibit nearly identical responses, making them equally susceptible to MGG variability. Additionally, a linear dependence of σV_{Th} on GS is observed. Another effect observed is the saturation of σV_{Th} for the NWFET as GS becomes larger, produced by the lack of grains present in the metal gate of the NWFET, due to the small A_G . The presence of only a few metal grains in the gate leads to a bimodal distribution of V_{Th} with peaks centered around the extreme work function values, instead of a V_{Th} normal distribution. This saturation becomes more pronounced when evaluating σSS . In the case of the FinFET and NSFET, both V_{Th} and SS follow normal distributions for all GS values.

To examine the impact of LER across the three architectures, we simulated sets of devices with randomly generated deformation profiles at three RMS values and a constant CL value. The influence of this variability source on the NSFET is negligible, with σV_{Th} being approximately 12 and 19 times smaller than the values exhibited by the FinFET and NWFET,

respectively. As previously explained, LER is induced by the lithography process during manufacturing. The deformation occurs in the critical dimensions of the FinFET and the NWFET, but not in the case of the NSFET. This characteristic makes the NSFET highly resilient to the effects of LER. On the other hand, although the critical dimension is a little smaller in the case of the FinFET, this architecture is more resistant than the NWFET. This is attributed to the FinFET's greater height, which results in a larger conducting area, thereby reducing the potential impact of LER. This accounts for the lower values of σV_{Th} and σSS for the FinFET. Note that, σSS is higher than that induced by MGG variability, as LER affects the linear growth region of the I_D-V_G characteristics, while MGG shifts the I_D-V_G curve proportionally to the average gate work function.

The variability induced by GER is in comparison lower than that of MGG and LER. The FinFET is the least disturbed architecture from this source of variability, with σV_{Th} values being about 20 % lower than for the NSFET and the NWFET. Upon analysing the reason for the FinFET's lower variability compared to the NSFET, despite their similar gate area and perimeter, it was demonstrated that the deformations localized in the lower lateral regions of the FinFET gate (which do not involve the semiconductor) have no influence on channel conduction. This portion of the gate constitutes the 9 % of the total perimeter of the studied FinFET.

The results of the RDD impact reveal that since all devices have the same doping profile, they exhibit equal resistance to this source of variability with $\sigma V_{Th} \sim 8$ mV.

Analogous to the study of the subthreshold region, a comparison was also made in the on region, through the evaluation of the σI_{on} influenced by the two sources of variability that had impacted the subthreshold behaviour the most: MGG and LER. In the on region, the impact of MGG on σI_{on} is quite similar to that observed in the subthreshold region, with NWFET emerging as the most resilient architecture, while the NSFET and the FinFET exhibit similar levels of sensitivity. It is also observed a minimal impact of LER on the σI_{on} of the NSFET. For this source of variability, the NWFET remains the less resilient architecture, with an σI_{on} approximately 15 % higher than that of the FinFET, a difference smaller than in the subthreshold case.

After evaluating the impact of variability sources on architectures that are in the state-of-the-art or are candidates for future technological nodes, we can conclude, considering only the resilience of each architecture, that the NSFET would be the most suitable choice to replace FinFETs, mainly due to its high resistance to LER. With this analysis, we would address all

aspects related to the **H1** hypothesis of this study.

In order to reduce the computational cost of the TCAD methodology in variability studies and thus address the **H2** hypothesis, an empirical model based on the Pelgrom law (*Pelgrom-based predictive model* (PBP)) was developed within the scope of this thesis. This model aims to predict the variations induced by the two most harmful sources of variability in the operation of the devices studied in this work (MGG and LER). These models are validated through a combination of results extracted from publications by several authors and our own simulations conducted in VENDES. The PBP model is developed separately for predicting σV_{Th} and σI_{on} , which correspond to simulations in the subthreshold region (QC DD) and the on region (QC MC), respectively.

When applying the PBP model to the subthreshold region, the prediction errors are equal or lower than the 10 % when compared to the TCAD results, as long as the devices fall within the validity limits of the model. These limits in the subthreshold PBP are due to the saturation of σV_{Th} caused by MGG variability when there is a lack of grains in the gate (large values of GS with devices having a small gate area). This effect breaks the assumed linearity in the model between σV_{Th} and GS. Notably, in the case of the LER application, there are no such limitations. For the subthreshold PBP model applied to MGG, about the 91 % of the relative errors are equal or lower than 8 %, while for the LER, all the predictions are within the 10 % margin, with the 94 % of predictions being below 6 %. These results demonstrate the adaptability of the subthreshold PBP model to a wide variety of architectures, dimensions, cross-sections (including three different FinFET shapes and two different ones for NWFET, considering various transistor designs reported in the literature by different authors), and variability parameters (GS for MGG, RMS and CL for LER).

When applying the PBP model to the on region for the three studied architectures, the relative errors for σI_{on} are below 8 % in the 92 % of the cases, combining predictions made for estimating the impact of MGG and LER. Similarly to the subthreshold case, limitations of the PBP model arise in the case of MGG variability when there is a lack of metal grains in the gate area. Once again, all estimations of σI_{on} due to LER remain below the required 10 %.

In terms of assessing the computational cost reduction, a clear benefit is observed in the application of the PBP model. For instance, in the on region, employing the PBP model for investigating the effects of MGG variability on a NWFET would lead to a substantial time reduction of around 90 % when compared to the computational expenses associated with the TCAD methodology. This estimation of savings is based on the assumption that the study of

MGG variability involves the simulations required for a specific device with 4 different GS values.

As demonstrated earlier, it is evident that the application of the PBP model to both the subthreshold and on regions offers a simple, fast, and accurate estimation of the impact caused by two of the most significant sources contributing to the variability of transistors across various architectures and dimensions. Therefore, the PBP model serves as a novel strategy that aids in reducing the computational cost of TCAD variability studies, thus addressing the hypothesis **H2** of this work.

Considering the obtained results, we can conclude that all the objectives of this thesis have been achieved. Both hypotheses were addressed, tackling relevant issues in the current state-of-the-art. These achievements include the identification of the most promising silicon-based field-effect transistor architecture for future technological nodes in terms of resilience to variability and the development of accurate strategies to reduce the computational cost of the TCAD methodology.

CAPÍTULO 1

INTRODUCCIÓN

Esta tesis está organizada en cuatro capítulos. El capítulo 1 se centra en contextualizar el trabajo realizado, además de mostrar las principales hipótesis y objetivos.

El capítulo 2 muestra la metodología empleada para la consecución de esta tesis y la descripción de las diferentes herramientas usadas y desarrolladas. En primer lugar, se describe el modelado tridimensional de los dispositivos, continuando por la exposición de todos los modelos de variabilidad aplicados, las correcciones cuánticas y los modelos de transporte empleados para simular los dispositivos.

En el capítulo 3 se llevará a cabo la discusión de los contenidos correspondientes a las publicaciones troncales que componen este trabajo. Primeramente, se explicará el proceso de calibración de los diferentes dispositivos estudiados, seguido del estudio y comparación del impacto de diferentes fuentes de variabilidad. Por último, se presentarán los modelos empíricos desarrollados para predecir el impacto de la variabilidad sobre diferentes transistores basados en silicio con el fin de reducir el costo computacional.

Finalmente, en las conclusiones se destacan las principales contribuciones de este trabajo, presentando los resultados más relevantes y proponiendo posibles líneas futuras de investigación.

Además, en el apéndice A se mostrarán los artículos que componen esta tesis, junto con trabajos complementarios como artículos o participaciones en congresos tanto internacionales como nacionales.

1.1. Motivación

No se puede entender el mundo actual sin el descubrimiento por parte de William Bradford Shockley, John Bardeen y Walter Houser Brattain de los transistores [1]. Este hallazgo, que fue merecedor del premio Nobel de Física en 1956, produjo un cambio de paradigma mundial en el siglo XX constituyendo la base de todos los sistemas de comunicación y computación actuales. La tecnología en la cual se fundamenta el desarrollo electrónico es la tecnología basada en la combinación de metal, óxido y semiconductor (*Metal Oxide Semiconductor* (MOS)), base de las unidades de memoria y procesamiento computacionales.

Las tecnologías basadas en transistores de efecto campo (*Field Effect Transistor* (FET)) han experimentado una reducción considerable de su tamaño, llegando incluso a integrar en un microprocesador miles de millones de estos (e.g. 2.5 billones en un Intel Core i9-12900K). El escalado de los dispositivos ha seguido la denominada Ley de Moore [2], hasta que numerosas dificultades de carácter físico han aparecido debido a la pequeña dimensionalidad de estos. Del mismo modo, siguiendo las directrices del *IEEE International Roadmap for Devices and Systems* (IRDS) [3], se plantean nuevas arquitecturas y formas de hacer frente a estas limitaciones. Ejemplos de estas arquitecturas son los FinFET [4], *nanowire FET* (NWFET) [5] o los *nanosheet FET* (NSFET) [6]. Estas estructuras FET han sido reportadas experimentalmente por IBM en colaboración con otras instituciones y empresas como: GlobalFoundries, Toshiba America Electronic Components, NEC Electronics o Samsung Electronics. Por tanto, para el avance del sector tecnológico es tremendamente necesario el estudio de nuevos materiales y diseños de transistores para la superación de las limitaciones en el escalado de los mismos. En consecuencia, la modelización de dispositivos que vayan más allá del estado del arte es de gran importancia.

También se convierte en indispensable la minimización de errores en el funcionamiento de los transistores, siendo muchos de estos errores derivados de los procesos de fabricación. Por causa de estos procesos, aparecen no-idealidades que se magnifican por causa de las pequeñas dimensiones que reclama la electrónica contemporánea. Estas desviaciones del funcionamiento ideal del transistor son derivadas de la aparición de diferentes fuentes de variabilidad, las cuales pueden ser consecuencia de los procesos de deposición, litográficos, etc. El estudio del impacto de estas fuentes de variabilidad es de vital importancia, ya que nos permite predecir y acotar las desviaciones respecto al comportamiento ideal de los dispositivos. A su vez, entender las diferentes fuentes de variabilidad nos permite evaluar que diseños de transistores son más inmunes a estas, con el fin de garantizar la homogeneidad en la respuesta de los dispositivos o

aprovechar su heterogeneidad para aplicaciones específicas (e.g. criptografía). A tan pequeña escala, existen fenómenos de carácter cuántico inherentes a la estructura de la materia como pueden ser la naturaleza discreta del átomo, el efecto túnel o el confinamiento cuántico. Todos estos fenómenos cuánticos producen también desviaciones respecto al comportamiento ideal de los transistores, degradando su rendimiento.

La metodología clásica de prueba y error se fundamenta en la fabricación de los dispositivos y su estudio experimental, y dado que para los estudios estadísticos de variabilidad se necesita un gran número de dispositivos (≥ 300), esta práctica es muy costosa. A mayores surgen otros problemas derivados de esta metodología, entre ellos la fabricación de arquitecturas más allá del estado del arte, ya que los instrumentos de fabricación están optimizados para materiales y diseños concretos. Por lo tanto, es imprescindible el uso de nuevas herramientas para su desarrollo. Aquí es donde entra el modelado de transistores a través del diseño asistido por ordenador (*Technology Computer Aided Design* (TCAD)), que se ha convertido en una herramienta indispensable para el sector. El TCAD nos permite explorar y predecir futuros diseños, arquitecturas o materiales óptimos para el control de las fuentes de variabilidad y el aprovechamiento de los efectos cuánticos previamente descritos.

Aun teniendo costos más bajos que la clásica metodología de ensayo y error, los estudios TCAD de variabilidad requieren de una gran cantidad de simulaciones para tener significancia estadística, además, la reducida dimensión de los dispositivos hace imprescindible la aplicación de correcciones cuánticas (gradiente de densidad, Schrödinger) a los modelos de transporte clásicos (arrastre difusión, Monte Carlo). Las simulaciones de dispositivos tridimensionales, la aplicación de correcciones cuánticas y la gran cantidad de simulaciones necesarias para realizar estudios estadísticos elevan en gran medida el costo computacional. Por tanto, es necesario encontrar otras alternativas más eficientes computacionalmente para la estimación del impacto de la variabilidad en los dispositivos actuales y en los futuros nodos tecnológicos.

Las simulaciones TCAD generan grandes cantidades de datos e información, por ello, es tremendamente importante tener herramientas de post-procesado para poder realizar el análisis de los resultados de manera rápida y coherente. La correcta gestión de gran cantidad de datos (*Big Data Management*) pasa a tener un carácter crucial en los estudios del impacto de la variabilidad. Una vez los datos e información están correctamente gestionados, adquiere relevancia la creación de repositorios públicos siguiendo estándares internacionales que garanticen su veracidad y reproducibilidad, con el fin de transferir y comparar los resultados con el resto de la comunidad científica e industria.

Con esta gran cantidad de datos y conociendo los parámetros característicos de las fuentes de variabilidad, se pueden desarrollar modelos predictivos empíricos o emplear diferentes técnicas como las basadas en *Machine Learning* (ML), para provocar un gran avance en la reducción de costos computacionales frente al TCAD sin pérdida de precisión en las estimaciones.

Este trabajo centrado en el estudio del impacto de las fuentes de variabilidad sobre dispositivos electrónicos de última generación, propone nuevas soluciones en el contexto de la nanoelectrónica, indispensables para el progreso del sector tecnológico. En esta tesis se comentarán ordenadamente las diferentes etapas de la metodología TCAD, desde el modelado de tridimensional de los dispositivos, hasta el estudio estadístico de la variabilidad. También, se realizará la descripción de una *software* propio creado para el post-procesado de la gran cantidad de datos generados en estos estudios y se desarrollaran nuevos modelos empíricos que permitirán la reducción del costo computacional de las simulaciones TCAD.

1.2. Hipótesis y objetivos generales

Las líneas principales de esta tesis se fundamentan en el análisis del rendimiento de dispositivos semiconductores candidatos a ser utilizados en los próximos nodos tecnológicos, el estudio del impacto que diferentes fuentes de variabilidad tienen en la fiabilidad y eficiencia de estos dispositivos, y el desarrollo de técnicas que optimicen el coste computacional y permitan la gestión eficiente de los datos. En este contexto y de manera más esquemática, este trabajo tratará de dar respuesta a los diferentes actuales con las siguientes hipótesis:

- H1.** *A través de un análisis detallado del impacto de diversas fuentes de variabilidad en las arquitecturas propuestas, se espera que el NSFET exhiba una mayor robustez en comparación con el NWFET como candidato preferente para reemplazar al FinFET en los futuros nodos tecnológicos.*
- H2.** *El desarrollo de modelos empíricos reducirá los costes computacionales vinculados a la realización de estudios del impacto de la variabilidad mediante la metodología TCAD, lo que potencialmente contribuirá a una mayor eficiencia en la investigación del sector tecnológico al reducir el tiempo y los recursos necesarios.*

Para llevar a cabo estas hipótesis, se enmarcan los diferentes objetivos generales mostrados a continuación:

- O1.** *Modelado y correcta calibración de dispositivos semiconductores que se encuentran más allá del estado del arte para su uso en los próximos nodos tecnológicos.*
- O2.** *Estudio estadístico del impacto de diferentes fuentes de variabilidad y sus implicaciones en las tecnologías actuales.*
- O3.** *Desarrollo de un software de post-procesado para la gestión de grandes cantidades de datos producidos por la metodología TCAD.*
- O4.** *Desarrollo de modelos para la predicción del impacto de las fuentes de variabilidad y reducción del costo computacional.*

Para la consecución de los objetivos O1, O2, y O3, realizamos los siguientes trabajos publicados en revistas internacionales o en contribuciones a congresos:

- N. Seoane, J. G. Fernandez, K. Kalna, E. Comesaña, and A. García-Loureiro. Simulations of statistical variability in n-type FinFET, nanowire, and nanosheet FETs. *IEEE Electron Device Letters*, 42(10):1416–1419, 2021
- J. G. Fernandez, N. Seoane, E. Comesaña, K. Kalna and A. García-Loureiro. Impact of metal grain granularity on three gate-all-around advanced architectures. *2021 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 201-205, 2021
- J. G. Fernandez, N. Seoane, K. Kalna, and A. García-Loureiro. Threshold voltage variability study in a 12 nm gate length nanosheet FET. *13th Spanish Conference on Electron Devices (CDE)*, 2021
- J. G. Fernandez, N. Seoane, E. Comesaña, and A. García-Loureiro. Analysis of the statistical impact of variability in a 12 nm nanosheet FET. *14th Spanish Conference on Electron Devices (CDE)*, 2023

El software *Machine Learning Figure of Merit Python extraction tool* (MLFoMpy) desarrollado para la consecución del objetivo O3, está disponible en <https://pypi.org/project/mlfompy/>. También se puede ver un ejemplo de gestión de los datos resultantes de esta herramienta de post-procesado en <https://doi.org/10.5281/zenodo.7674909>. En lo referente al objetivo O3, también se ha participado en el siguiente workshop:

- N. Seoane, J. G. Fernandez, E. Comesaña and A. Garcia-Loureiro. Semiconductor devices variability: modelling, characterization and data management. *18th MOS-AK ESSDERC/ESSCIRC Workshop*, 2021

En lo relacionado con los objetivos O2 y O4 se aportan las siguientes publicaciones en revistas internacionales:

- J. G. Fernandez, N. Seoane, E. Comesaña, and A. García-Loureiro. Pelgrom-based predictive model to estimate metal grain granularity and line edge roughness in advanced multigate MOSFETs. *IEEE Journal of the Electron Devices Society*, 10:953-959, 2022
- J. G. Fernandez, N. Seoane, E. Comesaña, and A. García-Loureiro. A comprehensive Pelgrom-based on-current variability model for FinFET, NWFET and NSFET. *Solid-State Electronics*, 199:108492, 2023
- J. G. Fernandez, N. Seoane, E. Comesaña, J. C. Pichel, and A. García-Loureiro. An accurate machine learning model to study the impact of realistic metal grain granularity on Nanosheet FETs. *Solid-State Electronics*, 207:108710, 2023
- A. García-Loureiro, N. Seoane, J. G. Fernandez, E. Comesaña, and J. C. Pichel. A machine learning approach to model the impact of line edge roughness on gate-all-around nanowire FETs while reducing the carbon footprint. *PLOS ONE*, 18(7) 2023

CAPÍTULO 2

METODOLOGÍA

En este capítulo se realizará una descripción de las herramientas empleadas para la simulación de diferentes arquitecturas y dispositivos semiconductores.

La Figura 2.1 muestra un esquema básico de la metodología seguida en esta tesis. El primer paso, que será tratado en la sección 2.1, corresponde con el modelado 3D de los dispositivos que incluye el diseño, mallado y asignación de características físicas, empleando para ello el software de código abierto Gmsh [7] y la herramienta de pre-procesado pVENDES. Una vez el dispositivo está modelado, se procede a la simulación realista de los procesos físicos y fenómenos de transporte de portadores, cuya descripción se encuentra en la sección 2.2. Para el proceso de simulación se empleará el software desarrollado en la Universidad de Santiago de Compostela, *Variability Enabled Nanometric DEvice Simulator* (VENDES) [8], el cual permite implementar diferentes modelos de variabilidad, aplicar diferentes correcciones cuánticas y simular varios modelos de transporte. Por último, una vez obtenidos los resultados de las simulaciones, estos son procesados para poder ser estudiados de manera correcta y sacar conclusiones relevantes. En este apartado es donde aparece la necesidad de una herramienta de post-procesado de gran cantidad de datos. Por tanto, en el contexto de esta tesis también se ha desarrollado un software de post-procesado llamado MLFoMpy, cuya descripción y funcionalidades serán comentadas en la sección 2.3.

2.1. Modelado 3D

El modelado 3D es la primera etapa del diseño tecnológico asistido por computador TCAD, que es la técnica comúnmente utilizada para realizar estudios estadísticos de variabilidad.

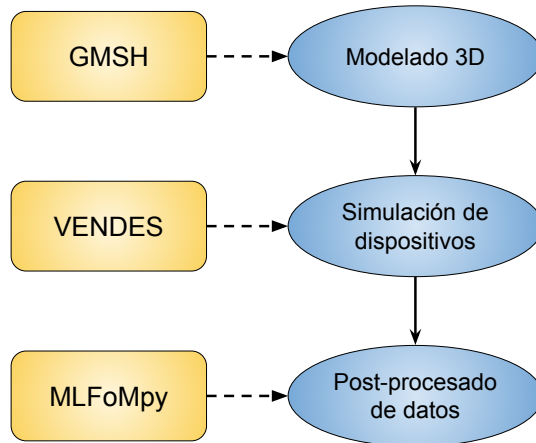


Figura 2.1: Esquema de la metodología aplicada.

Esta etapa tiene dos procesos diferenciados, primero el diseño y mallado del dispositivo y posteriormente la asignación de propiedades físicas a los diferentes materiales, regiones y nodos de este.

El inicio del procedimiento consiste en diseñar la geometría de nuestro dispositivo, para posteriormente generar una malla de elementos finitos que en nuestros casos de estudio será tridimensional con elementos tetraédricos, todo ello se realizará utilizando el software de código abierto Gmsh [7]. Las características geométricas y la forma del dispositivo vendrá determinada por la arquitectura escogida (FinFET, NWFET o NSFET), siendo consistentes en todo momento con transistores experimentales [6, 4, 5] o transistores escalados de los experimentales siguiendo las directrices marcadas por el IRDS [3]. La generación de la malla por elementos finitos, permite describir de manera precisa los diferentes dominios de simulación (los cuales pueden ser muy complejos) y además posibilita la introducción de deformaciones realistas en las dimensiones de los dispositivos (punto importante para estudiar determinadas fuentes de variabilidad). Esta funcionalidad es fundamental para el estudio de los efectos de la variabilidad, ya que la correcta distribución de potenciales y portadores es necesaria para poder predecir el comportamiento observado experimentalmente. En la Figura 2.2, se muestra un ejemplo de un NSFET modelado y mallado con Gmsh [7]. Una vez está realizado el diseño y mallado, es necesario el pre-procesado del dispositivo, proceso que consiste en otorgarle a los nodos de cada región sus propiedades físicas. En este caso, las regiones a definir corresponderán al metal, óxido y semiconductor (diferenciando la zona de canal, fuente y drenador). Las

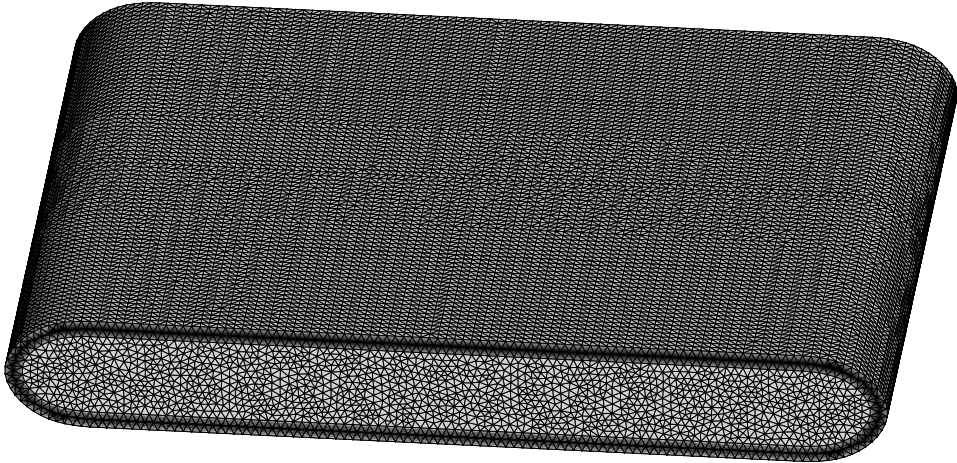


Figura 2.2: Ejemplo de modelado y mallado por elementos finitos de un NSFET.

propiedades físicas de cada nodo son otorgadas con la herramienta de pre-procesado pVENDES, la cual dotará a los nodos de cada región de magnitudes como dopados, permitividades, afinidades, movilidades, tiempos de vida de electrones y huecos, entre otras. Una vez se ha realizado el correcto modelado 3D del dispositivo, siguiendo las etapas de diseño, mallado y pre-procesado, se puede avanzar a la siguiente etapa del diseño tecnológico asistido por computador correspondiente a la simulación del dispositivo.

2.2. Simulación del dispositivo

La simulación del dispositivo es la parte central de la metodología, y aspira a la descripción realista de los diferentes procesos físicos que tienen lugar en el seno del dispositivo. Con este fin, se presentará el simulador VENDES (*Variability Enabled Nanometric DEvice Simulator*) [8], herramienta que permite implementar diferentes fuentes de variabilidad, mientras se aplican varios modelos de transporte, ya sean clásicos o semi-clásicos (con correcciones cuánticas). El diagrama con las funcionalidades del software VENDES es presentado en la Figura 2.3. El punto de partida para la realización de la simulación es la malla 3D correctamente pre-procesada del dispositivo semiconductor, sobre la cual se pueden introducir diferentes modelos de variabilidad o trabajar directamente sobre el dispositivo ideal (sin ningún modelo de

variabilidad implementado). El dispositivo ideal es primordial para la correcta calibración de los modelos de transporte y correcciones cuánticas empleadas en el simulador frente a los dispositivos experimentales o frente a complejas metodologías de simulación como el *Non-Equilibrium Green's Functions* (NEGF). El proceso de calibración será exhaustivamente comentado en la sección 3.1. Como este trabajo se fundamenta en el impacto de la variabilidad sobre los dispositivos nanoelectrónicos, a continuación será descrita la implementación de los modelos de variabilidad en VENDES.

2.2.1. Modelos de variabilidad

Se define como fuente de variabilidad a todos aquellos efectos físicos realistas que alejan a los transistores experimentales del comportamiento ideal, algunos de ellos propios de la física que sucede a estas escalas nanométricas y otros derivados de alguno de los procesos de fabricación de estos dispositivos basados en semiconductores.

La caracterización de algunas de las fuentes de variabilidad realizadas en la literatura [9, 10], permiten su implementación en técnicas TCAD. A continuación se desarrollarán las diferentes fuentes de variabilidad implementadas en VENDES, que corresponden con algunas de las que tienen mayor impacto sobre el funcionamiento de los transistores: granularidad metálica (*Metal Grain Granularity* (MGG)), rugosidad superficial (*Line Edge Roughness* (LER)), rugosidad de puerta metálica (*Gate Edge Roughness* (GER)) y fluctuaciones por dopantes aleatorios (*Random Discret Dopants* (RDD)).

Granularidad metálica (MGG)

Durante el proceso de la fabricación de la puerta del transistor, el metal se deposita de tal manera que aparecen diferentes orientaciones superficiales. Estas orientaciones superficiales se generan dado que la celda unidad superficial de la estructura cristalina está intersecada por diferentes planos, descritos por los índices de Miller $\langle h, k, l \rangle$. Estos cortes producen que las orientaciones tengan diferentes densidades superficiales atómicas, que se traduce en diferentes funciones de trabajo (ϕ_m). Estos entornos con varias ϕ_m causan un efecto de potencial no-uniforme a lo largo del canal de conducción.

Para implementar esta fuente de variabilidad en VENDES se generan perfiles de granularidad aleatorios utilizando diagramas de Poisson-Voronoi [11] que dependen del tamaño de grano medio (*Grain Size* (GS)). Este tamaño de grano depende de la temperatura y duración del recocido durante el proceso de deposición.

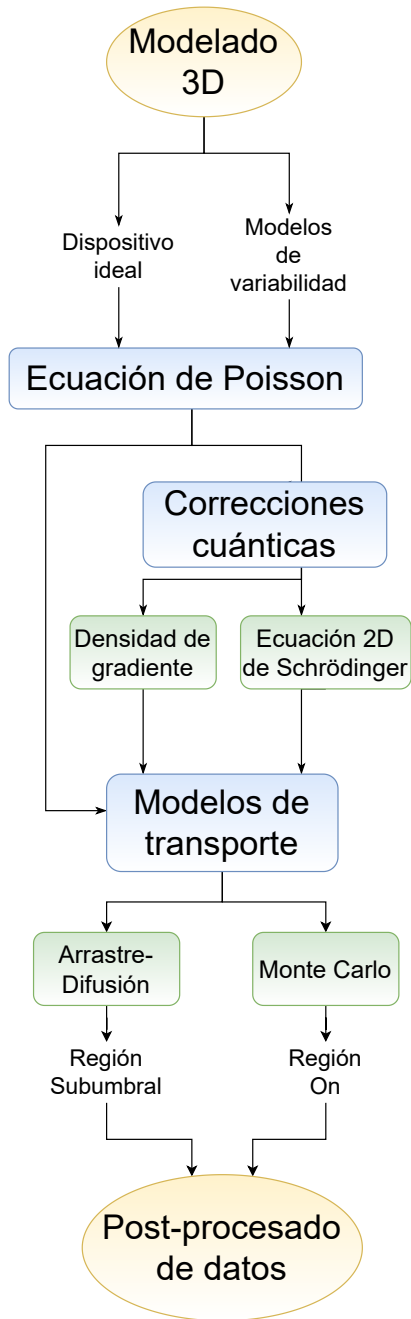


Figura 2.3: Esquema del proceso y funcionalidades del simulador VENDES.

En este trabajo todos los dispositivos estudiados tienen una puerta metálica basada en el nitruro de titanio (TiN), ya que es comúnmente utilizado en la actualidad por diferentes razones: tiene una estabilidad termal muy buena para el proceso de recocido [12], es compatible con el dióxido de silicio (SiO_2) y tiene una capacitancia muy baja para circuitos de muy baja potencia [13]. El TiN tiene dos posibles orientaciones metálicas $\langle 2, 0, 0 \rangle$ y $\langle 1, 1, 1 \rangle$ con $\phi_m = 4.4 \text{ eV}$ y 4.6 eV , en probabilidades de ocurrencia de un 40 % y 60 %, respectivamente [9]. En la Figura 2.4a) se muestra un NWFET con una puerta de TiN, donde cada color representa una orientación de grano diferente y donde el tamaño medio de grano es de 10 nm.

Rugosidad superficial (LER)

Por causa de los procesos litográficos en la fabricación de los transistores, al tener estos tan pequeña escala, es prácticamente imposible asegurar la idealidad de los dispositivos en lo que al contorno superficial se refiere. De manera realista, aparecen rugosidades en la superficie del semiconductor, que tienen un gran impacto en el funcionamiento del transistor, ya que estrangulan o ensanchan el canal de conducción, disminuyendo o aumentando respectivamente, la corriente que circula por él.

Para aplicar esta fuente de variabilidad en el simulador VENDES, se deben generar perfiles aleatorios que sean capaces de reproducir estas deformaciones realistas. Para crear estos perfiles se emplea la función de autocorrelación, que se basa en la transformada de Fourier del espectro gaussiano [10, 14]. Este espectro depende de dos parámetros, la longitud de correlación (*Correlation Length* (CL)) y la raíz cuadrática media de la deformación (*Root Mean Square height* (RMS)). Este primer parámetro, CL, representa cómo la deformación se propaga en su entorno vecino de la superficie y RMS es la profundidad de la deformación en la dirección del vector normal a la superficie. La Figura 2.4b) muestra un ejemplo de un NWFET impactado por la rugosidad superficial. Los valores de CL y RMS utilizados en este trabajo son consistentes con los parametrizados en dispositivos experimentales fabricados con litografía ultravioleta extrema (EUV por sus siglas en inglés) y analizados con microscopio de barrido electrónico (*Scanning Electron Microscope* (SEM)) [15, 16].

Rugosidad de puerta (GER)

La rugosidad de puerta sigue un proceso análogo a la superficial con la diferencia de que la deformación es aplicada sobre la puerta, concretamente, en las fronteras entre fuente/puerta y puerta/drenador, como se puede ver en la Figura 2.4c). Los perfiles aleatorios también se

generan a través de la función de autocorrelación, definida a partir de la transformada de Fourier del espectro Gaussiano [17]. Por tanto, los parámetros CL y RMS describirán también esta deformación con la peculiaridad de que CL se aplicará en la dirección del vector tangencial a la sección transversal del dispositivo en las fronteras entre fuente y drenador con la puerta metálica.

Fluctuaciones de dopantes aleatorios (RDD)

A la hora de dopar las regiones de fuente y drenador del transistor, debido a la naturaleza discreta de los dopantes, se generan potenciales que fluctúan el comportamiento ideal del transistor.

Por tanto, se describirá primero la implementación del dopado en nuestro dispositivo y posteriormente la aplicación de los dopantes aleatorios que producen variabilidad sobre el funcionamiento ideal del transistor. Para el perfil de dopado dentro del semiconductor, se asume un dopado tipo-p intrínseco en el canal de conducción y un dopado Gaussiano tipo-n en las regiones de fuente y drenador. Más detalles del proceso de ingeniería inversa para calibrar los perfiles de dopado con los resultados experimentales se pueden ver en [18, 19]. Para modelar las fluctuaciones de dopantes aleatorios se utiliza una técnica de exclusión sobre el perfil de dopaje previamente descrito en las zonas de fuente y drenador. Para cada transistor se mapea la malla tetraédrica tridimensional y se introducen dopantes aleatorios que se extienden a sus nodos próximos empleando el método de *cloud-in-cell* [20, 21]. Una visualización del efecto de esta fuente de variabilidad se puede observar en la Figura 2.4d), dónde se muestra un mapa de los potenciales en el canal de un FinFET y se puede apreciar el efecto en el potencial (V) de los dopantes aleatorios (zonas rojas).

2.2.2. Ecuación de Poisson

Partiendo del dispositivo ideal o del dispositivo afectado por las fuentes de variabilidad, el siguiente paso es la simulación de los procesos que ocurren en su interior. Para simular la física del dispositivo semiconductor se debe conocer el potencial de cada nodo del transistor. La ecuación de Poisson (2.1) permite calcular el potencial electrostático clásico ($V_{cl}(\mathbf{r})$) en cada uno de los nodos de la malla 3D tetraédrica introducida en el simulador:

$$\nabla \cdot (\epsilon(\mathbf{r})\nabla V_{cl}(\mathbf{r})) = q(p(\mathbf{r}) - n(\mathbf{r}) + N_D^+(\mathbf{r}) - N_A^-(\mathbf{r})), \quad (2.1)$$

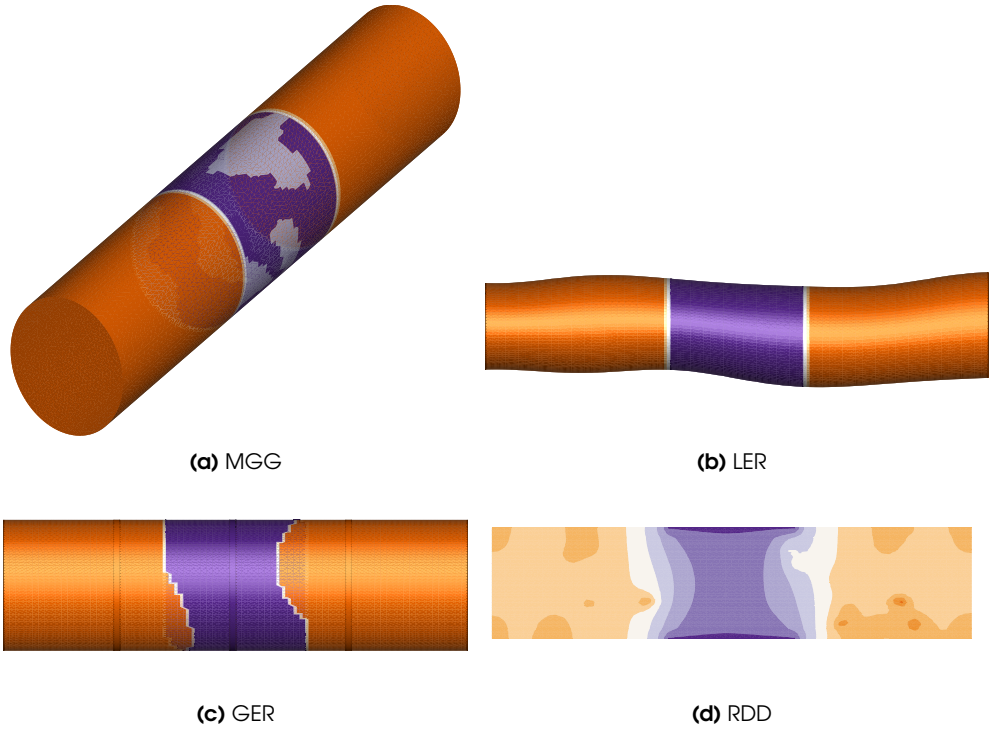


Figura 2.4: Esquema de las diferentes fuentes de variabilidad estudiadas en este trabajo, aplicadas sobre un NW-FET. (a) Muestra los granos en la puerta metálica que se corresponden a las diferentes funciones de trabajo. En (b) se observa la rugosidad superficial del dispositivo debida a los procesos litográficos. (c) Muestra la rugosidad producida sobre la puerta, haciendo que el contorno de esta sea no-uniforme. (d) Corte transversal del NW-FET dónde se puede observar el efecto de la aparición de dopantes aleatorios sobre el potencial interno del dispositivo.

siendo $\mathbf{r} \equiv (x, y, z)$ las coordenadas espaciales de cada nodo de la malla de elementos finitos, $\epsilon(\mathbf{r})$ la constante dieléctrica del material, q la carga del portador, $n(\mathbf{r})$ y $p(\mathbf{r})$ son las densidades de electrones y huecos y $N_D^+(\mathbf{r})$ y $N_A^-(\mathbf{r})$ la concentración de dopado efectivo de donadores y aceptores respectivamente.

2.2.3. Correcciones cuánticas

Una vez conocido el potencial electrostático clásico es necesario tener en cuenta diferentes efectos cuánticos relevantes. Dado el tamaño actual de los dispositivos en el estado del arte y de los futuros nodos tecnológico, los mecanismos cuánticos tienen un enorme impacto, y

por ello hay que incluirlos de manera obligada en la simulación realista del dispositivo. En esta sección se plantean dos tipos de correcciones cuánticas (*Quantum Corrections* (QC)) sobre el potencial clásico implementadas en VENDES, para tener una descripción realista del comportamiento de los transistores: el gradiente de densidad y la ecuación bidimensional de Schrödinger.

Gradiente de densidad

La primera de ellas, la ecuación 3D de la gradiente de densidad (*Density Gradient* (DG)), corrige cuánticamente el potencial para los electrones. En este caso, el potencial cuántico (V_{dg}) se calculará:

$$V_{dg}(\mathbf{r}) = 2[b_n] \frac{\nabla^2 \sqrt{n(\mathbf{r})}}{\sqrt{n(\mathbf{r})}} = \phi_n(\mathbf{r}) - V_{cl}(\mathbf{r}) + \frac{k_B T}{q} \ln \left(\frac{n(\mathbf{r})}{n_i(\mathbf{r})} \right) \quad (2.2)$$

donde,

$$[b_n] = \frac{\hbar^2}{4qr_n} \begin{pmatrix} 1/m_x & 0 & 0 \\ 0 & 1/m_y & 0 \\ 0 & 0 & 1/m_z \end{pmatrix} \quad (2.3)$$

En estas ecuaciones, $\phi_n(\mathbf{r})$ representa el cuasi-potencial de Fermi para electrones, $n_i(\mathbf{r})$ es la concentración de portadores intrínsecos, k_B la constante de Boltzmann y T la temperatura de red. También aparece el tensor $[b_n]$, que depende de la constante de Planck reducida \hbar , de un parámetro de ajuste r_n y de m_x , m_y , m_z , las masas electrónicas efectivas DG, que actuarán como parámetros de calibración de nuestros dispositivos [22], no teniendo relación con las masas efectivas de transporte en el material. El principal problema de este modelo es el requerimiento de su calibración frente a datos experimentales (si existen en la literatura) o frente a técnicas más complejas como el NEGF.

Ecuación 2D de Schrödinger

Por otra parte, la aplicación de la ecuación 2D de Schrödinger (2D-SCH) para la corrección del potencial electrostático en cada nodo, es libre de cualquier calibración. Esta ecuación es resuelta en láminas bidimensionales transversales al canal de conducción colocadas a lo largo dispositivo, como se puede ver en la Figura 2.5. A partir de la ecuación 2D-SCH se calcula la densidad cuántica electrónica 2D ($n_{sch}(y, z)$), la cual es interpolada al dominio 3D para ser

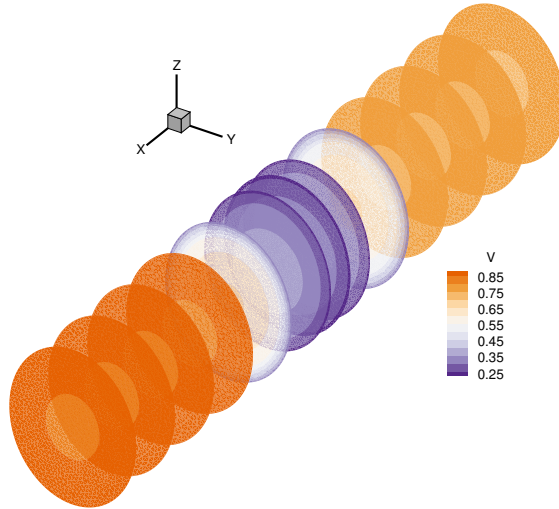


Figura 2.5: Láminas transversales de un NWFET dónde se resuelve la ecuación 2D de Schrödinger bajo una polarización en el drenador (V_D) de 0.1 V.

consecuentes con las características tridimensionales del dispositivo ($n_{sch}(\mathbf{r})$). Finalmente, la corrección cuántica del potencial será tal que:

$$V_{sch}(\mathbf{r}) = \frac{k_B T}{q} \log \left(\frac{n_{sch}(\mathbf{r})}{n_i(\mathbf{r})} \right) - V_{cl}(\mathbf{r}) \quad (2.4)$$

2.2.4. Modelos de transporte

Para simular el transporte a lo largo del canal de conducción VENDES incluye dos modelos de transporte diferentes: i) arrastre y difusión (*Drift-diffusion* (DD)) y ii) Monte Carlo (MC), como se puede ver en el esquema de la Figura 2.3.

Arrastre-Difusión

El método de transporte de arrastre y difusión [23] parte del potencial electrostático calculado por la ecuación de Poisson y que puede ser corregido cuánticamente (por DG o 2D-SCH) y de las ecuaciones de continuidad de los electrones, para obtener la densidad de

corriente de electrones, $J_n(\mathbf{r})$:

$$J_n(\mathbf{r}) = -q\mu_n(\mathbf{r})n(\mathbf{r})\nabla(\phi_n(\mathbf{r})), \quad (2.5)$$

$$\nabla \cdot (J_n(\mathbf{r})) = qR(\mathbf{r}), \quad (2.6)$$

definiendo $\mu_n(\mathbf{r})$ como la movilidad electrónica y $R(\mathbf{r})$ como el término de recombinación. Este método, que solo tiene en cuenta la relación local entre velocidad y campo eléctrico, no es capaz de representar los fenómenos de transporte fuera del equilibrio. Aun así, VENDES aplica diferentes correcciones (e.g. modelo de Caughly-Thomas) para describir mejor el transporte de carga a campos eléctricos elevados. La aplicación de estos modelos, con el fin de evitar el problema de la localidad, depende de la calibración de diferentes parámetros como: la movilidad de portadores a campos bajos, una velocidad de saturación v_{sat} o la aparición de un campo eléctrico crítico E_{cn} . Este método de transporte necesita de calibración previa frente a dispositivos experimentales o frente a complejas metodologías de simulación como el NEGF y presenta bajo ruido en todo el rango de simulación. Por tanto, este método de transporte será adecuado para describir el funcionamiento del transistor en la región subumbral (Subthreshold region) debido a la buena reproducción de las dinámicas de portadores a bajos campos eléctricos, pero presenta problemas para representar los efectos fuera del equilibrio propios de la región de encendido (On region) [24].

Monte Carlo

El otro método de transporte es el MC [25] que tiene características de un método semi-clásico de transporte, fundamentándose en la resolución de la ecuación de transporte de Boltzmann a través del método de Monte Carlo acoplado con la ecuación de Poisson. En esencia, este método consiste en la simulación del movimiento de un conjunto de portadores que se desplazan en régimen de vuelo libre, gobernados por las ecuaciones de Newton, sufriendo procesos o eventos de dispersión (o *scattering*) con una probabilidad de suceso determinada por la mecánica cuántica en función de cada material.

Para inicializar la distribución de las cargas en el espacio real partimos del potencial calculado en la resolución de la ecuación de Poisson corregido cuánticamente. Los procesos de dispersión de portadores incluidos en la simulación MC son: i) interacción electrón-fonón (que tiene diferentes tipos), ii) interacción electrón-impurezas ionizadas y apantallamiento, y iii) dispersión de electrones por rugosidades superficiales. La metodología MC describe de manera correcta los efectos fuera del equilibrio y por tanto, será más exacto para reproducir

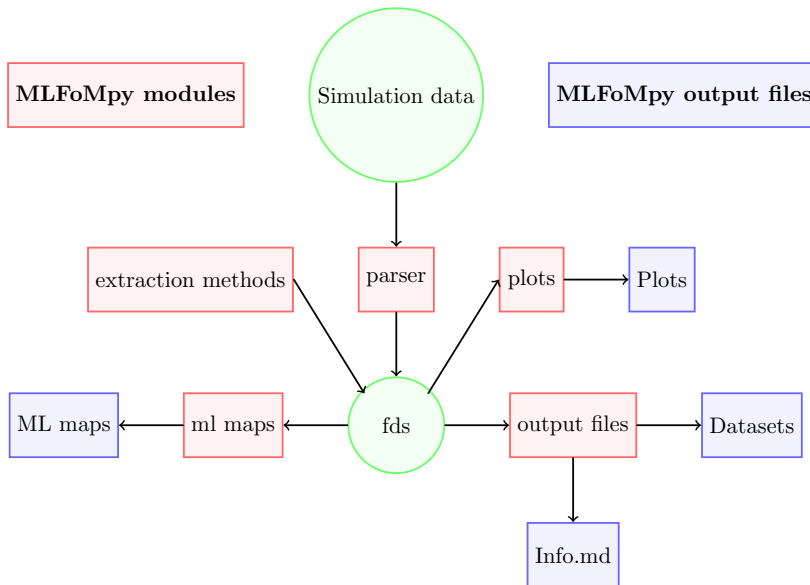


Figura 2.6: Esquema y funcionalidades del MLFoMpy.

el comportamiento de los transistores en la región en encendido del dispositivo, pero con el inconveniente de producir un mayor ruido a bajos campos eléctricos en la región subumbral que el DD.

2.3. MLFoMpy

Después de la aplicación de la ecuación de Poisson con las correspondientes correcciones cuánticas y modelos de transporte seleccionados al dispositivo, se tendrán los resultados de la simulación como curvas intensidad-voltaje. Al realizar estudios de variabilidad con cientos o miles de simulaciones, la cantidad de datos generada es inmensa, ya que las salidas serán cientos o miles de curvas I-V necesarias para tener significancia estadística, por ello es necesario realizar un correcto post-procesado de los datos para poder extraer resultados. En esta última etapa de la metodología se empleará el software de post-procesado MLFoMpy, cuyas funcionalidades son mostradas en la Figura 2.6 y cuyo desarrollo también ha sido realizado en el contexto de esta tesis.

La función principal de este software es la extracción de figuras de mérito (*Figures of Merit* (FoM)) a partir de las curvas características corriente-voltaje (I-V) de los dispositivos

simulados, pero también tiene otras funciones como la generación de conjuntos de datos (*datasets*) o la creación de archivos específicos para la realización de estudios basados en *Machine Learning*, aprovechando así la gran cantidad de datos generados en estudios de variabilidad. La extracción de FoM es relevante, porque permite caracterizar, comparar numéricamente y realizar estudios estadísticos de diferentes variables que determinan el funcionamiento de los dispositivos.

El primer paso del software es importar los resultados de las simulaciones a través del módulo *parser* en el objeto *fds* (*fompy dataset*). Este módulo permite incorporar datos desde diferentes formatos de curvas I-V (.csv, .txt, curvas I-V sin normalizar, ...). Una vez los datos están en el objeto *fds*, se aplican las operaciones y tratamientos de interés sobre este objeto. El siguiente módulo es el correspondiente a *extraction methods* cuya función es extraer las FoM de las curvas I-V importadas previamente. Las FoM que extrae esta herramienta son: el voltaje umbral (*threshold voltage, V_{Th}*) que indica el voltaje al cual el transistor pasa de estar en un estado apagado a otro encendido; la corriente en off (*off current, I_{off}*) que es una corriente de fuga que circula por el canal en el estado en apagado del dispositivo, debida a fenómenos cuánticos como el efecto túnel; la corriente subumbral (*subthreshold slope, SS*) la cual muestra cómo transiciona el transistor de un estado de apagado al de encendido; o la corriente en on (*on current, I_{on}*), que es la corriente de funcionamiento en el estado en encendido del dispositivo, además el ratio I_{on}/I_{off} da información sobre el rendimiento del dispositivo.

Otro módulo interesante es *ml maps* que permite relacionar los datos resultantes de simulaciones de variabilidad con sus perfiles respectivos de variabilidad. De este modo, se puede tener un conocimiento más extenso sobre la manera en la cual diferentes fuentes de variabilidad impactan sobre el funcionamiento de los transistores. Además, se pueden entrenar redes neuronales con esta gran cantidad de datos para realizar estudios que permitan ahorrar tiempo computacional. El resto de módulos dotan al software de diversas funcionalidades de gran utilidad para el usuario, de este modo, *plots* permite extraer gráficas recurrentes para tener una inmediata visualización de los resultados de variabilidad. También el módulo *output files* da la posibilidad de almacenar los resultados de los estudios de variabilidad en *datasets*, guardar la información necesaria para asegurar la reproducibilidad del proceso de simulación, o generar salidas que puedan ser empleadas para otras técnicas complementarias, como por ejemplo, el entrenamiento de redes neuronales.

CAPÍTULO 3

DISCUSIÓN

El trabajo en esta tesis se basa en el estudio del funcionamiento de transistores nanométricos impactados por diferentes fuentes de no-idealidad. Estas fuentes llamadas de variabilidad derivan de la pequeña dimensionalidad de estos dispositivos o son inherentes a los procesos de fabricación. De forma más específica, el objetivo principal es estudiar los futuros nodos tecnológicos y observar que fuentes de variabilidad pueden ser las más perjudiciales en su funcionamiento. Todo esto se realizará para conocer cómo los procesos de fabricación alejan el funcionamiento de los transistores de la idealidad, ya sea con el fin de reducir el impacto de las fuentes de variabilidad o, en casos especiales, aprovecharlo (generación de *Physical Unclonable Functions* (PUFs) [26]). Además, dado el elevado costo computacional de los estudios de variabilidad realizados con técnicas TCAD, en esta tesis también se desarrollan diferentes modelos empíricos para tener una rápida estimación de las características eléctricas de dispositivos futuros, ya sea en la región subumbral o en la región de encendido.

El contenido de la discusión se desarrollará comenzando por la sección 3.1 dónde se realizará la descripción del proceso de calibración de los diferentes dispositivos para tres arquitecturas, tanto usadas en la actualidad como proyectadas por el IRDS [3] para su utilización en próximos nodos tecnológicos. Seguidamente, en la sección 3.2 se mostrará una comparativa entre estas arquitecturas (FinFET, NWFET y NSFET), con el fin de vislumbrar cuál de ellas es más resistente a diferentes fuentes de variabilidad. Finalmente, en la sección 3.3 se procederá a la deducción de modelos empíricos basados en la Ley de Pelgrom que permitirán tener una estimación precisa de las diferentes FoM para dispositivos afectados por fuentes de variabilidad.

3.1. Calibración y descripción de los dispositivos

En esta sección se muestra el proceso de calibración de los dispositivos ideales que serán empleados posteriormente para la realización de los estudios de variabilidad. Para explicar la calibración se describirán sus arquitecturas, sus dimensiones y los diferentes modelos de transporte y correcciones cuánticas empleados en su simulación.

Los dispositivos se presentan en la Figura 3.1 estando basados en arquitecturas que se encuentran en el estado del arte y/o son candidatas a los futuros nodos tecnológicos. En la Figura 3.1(a) se puede ver un NSFET con longitud de puerta (L_g) de 12 nm, este dispositivo ha sido diseñado conforme al experimental publicado en [6]. El FinFET con L_g de 12 nm ha sido desarrollado en este trabajo, como una versión tripuerta del NSFET de 12 nm, como se muestra en la Figura 3.1(b). La elección de la forma de este FinFET ha sido tomada teniendo en cuenta las recomendaciones del IRDS [3], donde comentan este acercamiento en la fabricación del FinFET a la del NSFET para los futuros nodos tecnológicos. Por último, el NWFET de 12 nm (véase su forma en la Figura 3.1(c)) ha sido escalado de manera acorde con las dimensiones propuestas por el IRDS [3] desde el NWFET de 22 nm experimental publicado por [5], el cual ha sido previamente calibrado en [19]. En la Figura 3.1(d) se muestran las dimensiones y datos relevantes de los tres dispositivos a calibrar. Nótese que para la realización de una comparativa justa, todos los dispositivos se han diseñado con una L_g de 12 nm. El resto de parámetros mostrado en la Figura 3.1(d) son: $L_{s/d}$ la longitud de fuente/drenador, L longitud total del dispositivo, w_{ch} ancho del canal, h_{ch} alto del canal, t_{ox} anchura del óxido efectiva, P_s el perímetro del semiconductor, $N_{s/d}$ el dopado tipo-n en fuente drenador, X_m la posición de decaimiento del dopado gaussiano tipo-n y δ la forma en la que decae el dopado gaussiano.

El perfil de dopado es esencial para poder realizar simulaciones realistas del transistor en cuestión, siendo muy difícil acceder a este a partir de un trabajo experimental. Por ello, para conseguir este perfil de dopado es necesaria la realización de un proceso de ingeniería inversa partiendo de las características I_D - V_G (siendo I_D la corriente de drenador y V_G la tensión aplicada sobre la puerta) de la región subumbral del dispositivo experimental [18, 19]. Como se comenta previamente en la metodología, asumiremos dopado tipo-p uniforme a lo largo del canal ($N_{ch} = 1 \times 10^{15} \text{ cm}^{-3}$) y un dopado gaussiano tipo-n dependiente de unos parámetros de ajuste como son, el máximo dopado en la fuente y drenador (en este caso $N_{s/d} = 5 \times 10^{19} \text{ cm}^{-3}$), la posición de decaimiento del dopado gaussiano (X_m , cercana a las fronteras fuente/canal y canal/drenador), o la forma en la que decae el perfil (δ). En la Figura 3.2 se muestra el perfil de dopado gaussiano, con los diferentes parámetros que lo definen, para un NSFET con L_g de

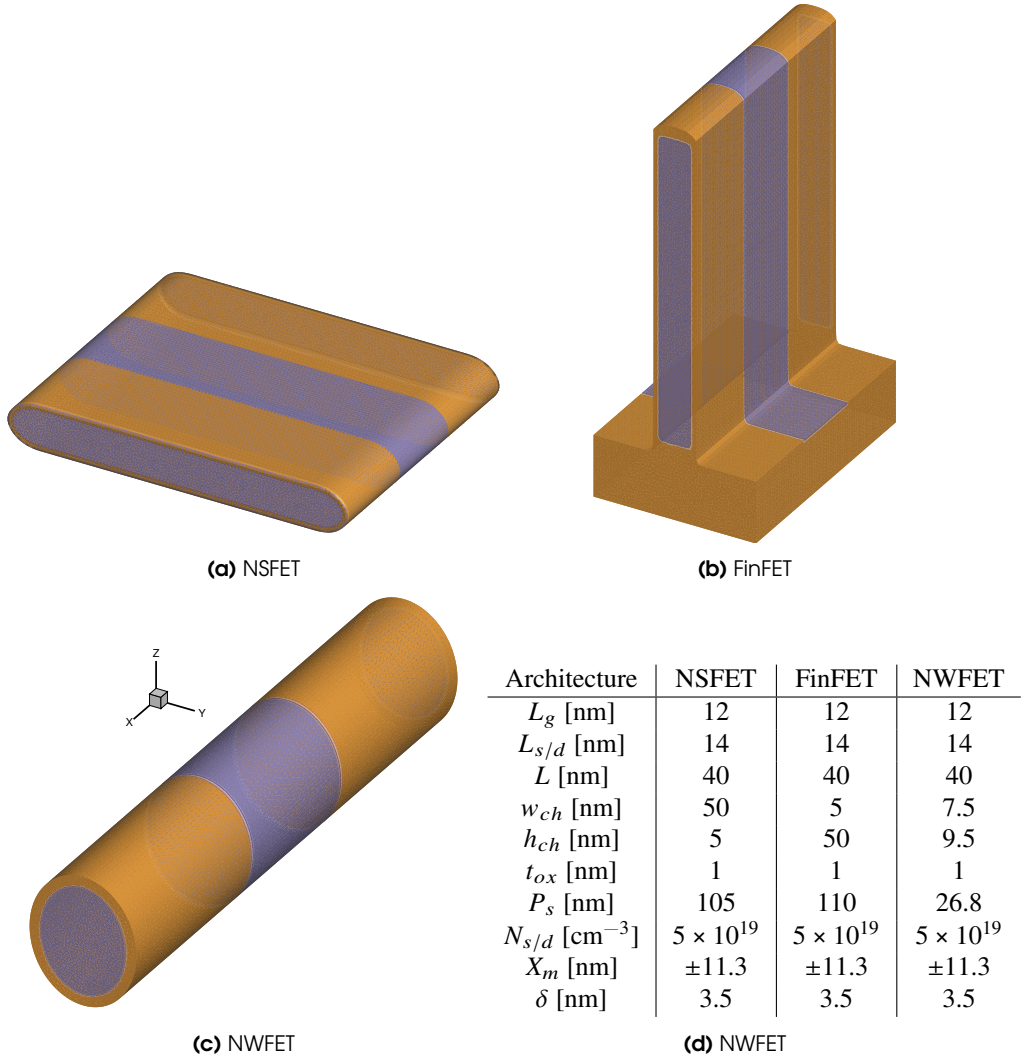


Figura 3.1: Esquema de las diferentes arquitecturas 3D modeladas siendo: (a) NSFET, (b) FinFET y (c) NWFET. También se muestra una tabla con los valores de los dispositivos simulados en (d) siendo: L_g la longitud de puerta, $L_{s/d}$ la longitud de fuente-drenador, L longitud total del dispositivo, w_{ch} ancho del canal, h_{ch} alto del canal, t_{ox} anchura del óxido efectiva, P_s el perímetro del semiconductor, $N_{s/d}$ el dopado tipo-n en fuente-drenador, X_m la posición de decaimiento del dopado gaussiano tipo-n y δ la forma en la que decae el dopado gaussiano.

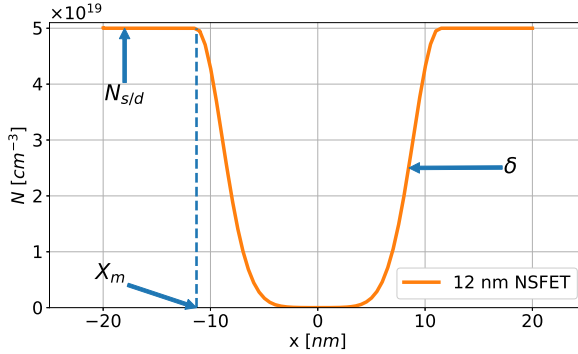


Figura 3.2: Perfil de dopado gaussiano de un NSFET con longitud de puerta de 12 nm. N es la concentración de dopado tipo-n en el dispositivo. $N_{s/d}$ es el máximo dopado en la fuente y drenador, X_m la posición de decaimiento del dopado gaussiano cercana a las fronteras fuente/canal y canal/drenado, y δ la forma con la que decae el perfil de dopado.

12 nm. Estos parámetros son modificados repetidamente hasta que se consigue un buen ajuste entre las características I_D - V_G experimentales y las obtenidas en la región subumbral con la metodología DD corregida cuánticamente con DG o 2D-SCH (descritos en la metodología), más detalles del proceso de ingeniería inversa se muestran en [19].

Una vez ajustado el dopado, se calibrará el modelo de transporte DD con las correcciones cuánticas DG y 2D-SCH. La comparación entre las características I_D - V_G experimentales y las simuladas para el dispositivo ideal, se muestran en la Figura 3.3, donde se puede observar que los modelos implementados (DD/DG, DD/SCH y MC/SCH) reproducen fielmente las propiedades de los diferentes dispositivos: Figura 3.3(a) NSFET, Figura 3.3(b) FinFET y Figura 3.3(c) NWFET, todos con L_g de 12 nm a una tensión de drenador de $V_D = 0.7$ V. Más información sobre el proceso de calibración de estos dispositivos se puede encontrar en [18, 19]. Tal y como se comentó en la metodología, se observa en la Figura 3.3 que el método de transporte DD corregido cuánticamente reproduce perfectamente el comportamiento de los transistores tanto en la región subumbral como en la región de encendido. Esto se debe a la exhaustiva calibración realizada de los diferentes parámetros del modelo DD, como son la movilidad de portadores a campos bajos, la velocidad de saturación v_{sat} o el campo eléctrico crítico normalizado E_{cn} . También se puede vislumbrar la incapacidad del modelo de transporte basado en el MC semiclásico para reproducir con exactitud el comportamiento del transistor en la región subumbral. Esto se debe al elevado ruido estadístico derivado de los pequeños valores de I_D a bajos V_G que es inherente a este modelo de transporte.

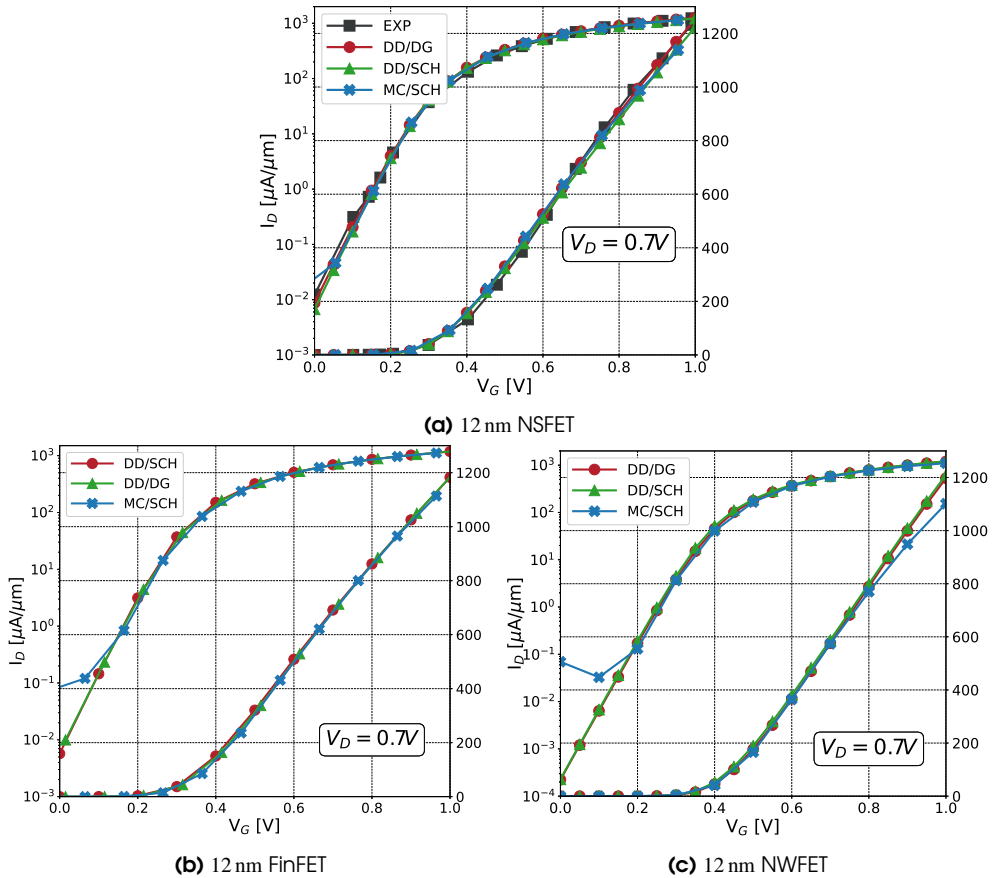


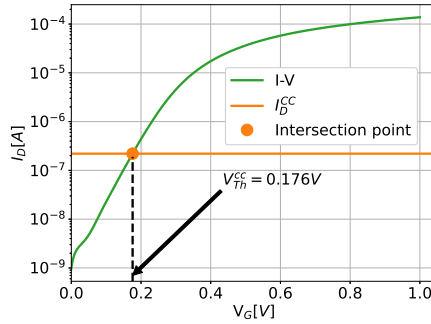
Figura 3.3: Características I_D - V_G de los diferentes dispositivos calibrados para este trabajo a una tensión de drenador V_D de 0.7 V: (a) 12 nm NSFET, (b) 12 nm FinFET y (c) 12 nm NWFET. En las leyendas de las figuras se muestra la combinación método de transporte/corrección cuántica, siendo DD arrastre difusión, MC Monte Carlo, DG gradiente de densidad y SCH la ecuación de Schrödinger. Los datos experimentales se denotan como EXP

3.2. Estudio y comparación del impacto de diferentes fuentes de variabilidad

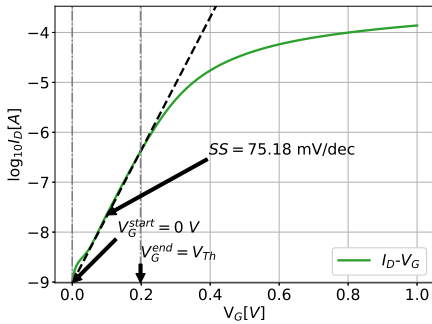
Una vez los dispositivos han sido calibrados, se llevará a cabo un estudio comparativo para analizar el impacto de las diferentes fuentes de variabilidad: MGG, LER, GER y RDD (previamente descritas en la sección 2.2.1). Como se mencionó con anterioridad, la condición de polarización del drenador será $V_D = 0.7\text{ V}$ para las tres arquitecturas. Con el objetivo de obtener la significancia estadística necesaria en los estudios de variabilidad, se han generado conjuntos de 300 dispositivos afectados por cada una de las fuentes y parámetros de variabilidad. En este trabajo la simulación de la región subumbral se ha realizado con la metodología DD 2D-SCH y para el estudio de la región de encendido se ha empleado MC 2D-SCH por lo comentado previamente en la sección 2.2.

Para llevar a cabo este estudio, se compararán las principales figuras de mérito (FoM) que ayudan a describir el funcionamiento de los distintos transistores. La extracción de estas FoM se llevará a cabo con el software desarrollado en el contexto de esta tesis MLFoMpy, descrito en la sección 2.3. A continuación presentaremos las diferentes FoM junto con los criterios de extracción empleados:

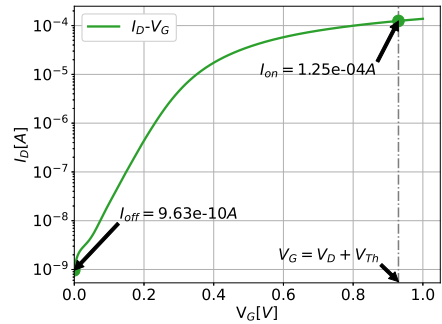
- La tensión umbral (V_{Th}) indica el potencial necesario para establecer el canal de conducción entre la fuente y el drenador, cambiando así el estado del transistor del modo apagado a modo encendido. Para asegurar una comparación justa de V_{Th} entre las diversas arquitecturas, se aplicó el criterio de corriente constante [27]. Para este método, el criterio se fijó a una densidad de corriente de drenador de $2.0\ \mu\text{A}\ \mu\text{m}^{-1}$, próxima a la zona de transición entre la región subumbral y la de encendido para las características I_D-V_G de los tres dispositivos estudiados. Este criterio se puede visualizar en la Figura 3.4(a) aplicado sobre la curva I_D-V_G del 12 nm NSFET, donde la densidad de corriente se multiplica por el perímetro del semiconductor para calcular la corriente a la que se obtiene la tensión umbral (I_D^{CC}), que en este caso será $2.2 \times 10^{-7}\text{ A}$.
- La pendiente subumbral (SS) se define como la inversa de la pendiente en la región de crecimiento lineal (correspondiente a la región subumbral) de las características I_D-V_G del transistor. La región lineal subumbral está definida en el intervalo comprendido entre $V_G^{start} = 0\text{ V}$ y $V_G^{end} = V_{Th}$, como se puede observar en la Figura 3.4(b).



(a) Criterio corriente constante (CC) para la extracción de V_{Th} . La corriente utilizada en este criterio (I_D^{CC}) es 2.2×10^{-7} A, correspondiendo con V_{Th} el punto de corte de ambas curvas



(b) Pendiente subumbral (SS), definida como la inversa de la pendiente de la recta definida en el intervalo entre $V_G^{start} = 0$ V y $V_G^{end} = V_{Th}$



(c) Corrientes de apagado (I_{off}) y de encendido (I_{on}), siendo $I_{off} = I_D(0$ V) y $I_{on} = I_D(V_D + V_{Th})$

Figura 3.4: Métodos de extracción de figuras de mérito empleado en el desarrollo de este trabajo aplicados sobre la curva del NSFET con longitud de puerta 12 nm.

- La corriente de apagado (I_{off}) se define como la corriente en la cual el potencial de la puerta es nulo ($V_G = 0$ V). La extracción de I_{off} está ejemplificada en la Figura 3.4(c).
- La corriente de encendido I_{on} es igual a la corriente correspondiente a $V_G = V_D + V_{Th}$. La extracción de I_{on} se muestra conjuntamente con la extracción de I_{off} en la Figura 3.4(c).

Las Figuras 3.5(a)-(b) presentan las desviaciones estándar (σ) de V_{Th} y SS que corresponden al impacto de la variabilidad MGG para diferentes GS sobre las tres arquitecturas

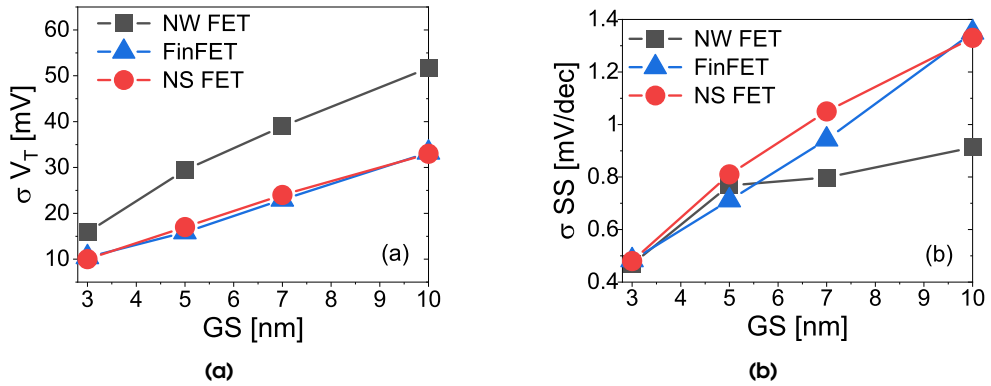


Figura 3.5: Desviaciones estándar de (a) la tensión umbral (σV_{Th}) y de (b) la pendiente subumbral (σSS) debidas a la variabilidad MGG frente al GS para NWFET, FinFET y NSFET.

analizadas. Podemos notar que tanto el NSFET como el FinFET son más resistentes al impacto de esta variabilidad en comparación con el NWFET. Esto se debe a que a medida que el área de la puerta (A_G) disminuye, se esperan mayores desviaciones. En este caso, para un GS de 3 nm, la desviación estándar de V_{Th} para el NWFET ($A_G \approx 395 \text{ nm}^2$) es un 59% mayor que la desviación estándar de V_{Th} para el FinFET y el NSFET ($A_G > 1000 \text{ nm}^2$).

Tanto el FinFET como el NSFET exhiben una respuesta prácticamente idéntica ante la variabilidad MGG (Figura 3.5(a)), sugiriendo que ambas arquitecturas son igualmente susceptibles a MGG. Además, se puede afirmar que σV_{Th} aumenta linealmente con el GS para FinFET y NSFET. El NWFET experimenta un impacto mayor inducido por esta fuente de variabilidad que FinFET y NSFET. También se puede observar que σV_{Th} comienza a saturarse para NWFET cuando el GS es grande, debido a la escasez de granos metálicos presentes en la puerta metálica, ya que este dispositivo tiene una A_G mucho menor que el FinFET y el NSFET (por ejemplo, para un GS de 10 nm solo habría sobre 4 granos en la puerta). La aparición de pocos granos metálicos en la puerta provoca que la distribución estadística de V_{Th} siga un patrón bimodal en vez de normal. Esta saturación se hace aun más evidente en σSS para los NWFET con GS igual o superior a 5 nm (Figura 3.5(b)). Sin embargo, las distribuciones de V_{Th} y SS siguen una distribución normal para todos los valores de GS en el caso del FinFET y NSFET.

Las Figuras 3.6(a)-(b) muestran la σV_{Th} y la σSS , respectivamente, de los dispositivos impactados por LER en función de RMS para un CL de 20 nm. Este valor de CL ha sido elegido basándonos en estudios previos realizados con SEM [10, 16]. Como podemos observar, la

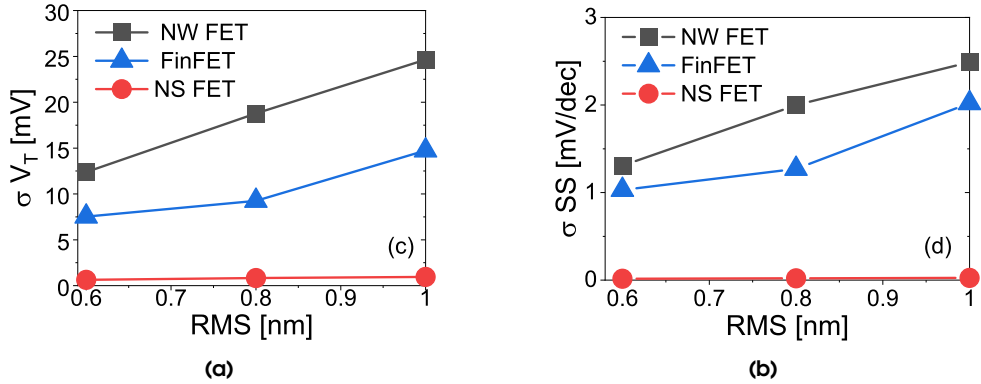


Figura 3.6: Desviaciones estándar de (a) la tensión umbral (σV_{Th}) y (b) la pendiente subumbral (σSS) debidas al impacto del LER frente a la raíz cuadrática media de la deformación (RMS) para una longitud de correlación (CL) de 20 nm para NWFET, FinFET y NSFET.

influencia del LER en el NSFET es insignificante, siendo los valores de σV_{Th} del orden de 12 y 19 veces menores a los mostrados por el FinFET y el NWFET, respectivamente, para un RMS de 0.6 nm, aumentando más esta diferencia con el incremento de RMS.

El LER es inducido por el proceso litográfico [28] y como se mencionó anteriormente, la deformación del LER se aplica en la dirección Y (véase Figura 3.1), correspondiendo a las dimensiones críticas para el FinFET (5.0 nm) y el NWFET (7.5 nm). Sin embargo, la dimensión crítica del NSFET es su altura (dirección Z) que no se verá afectada por el LER, ya que la deformación está definida por el crecimiento epitaxial [6]. La anchura del NSFET (50 nm) es lo suficientemente grande como para que prácticamente no se vea afectada por cualquier deformación de este tipo. Por otro lado, el FinFET aun teniendo la misma anchura que el NWFET, su mayor altura (50 nm) implica mayor área de conducción, lo cual reduce el impacto potencial de las deformaciones LER y explica los valores menores de σV_{Th} y σSS en comparación con las del NWFET. Además, en el caso del FinFET y el NWFET, la σSS inducida por el LER suele ser mayor que la observada para la variabilidad MGG. Esto se puede explicar, ya que la MGG actúa sobre los dispositivos desplazando las características I_D-V_G proporcionalmente a la función de trabajo (ϕ_m) media de la puerta, mientras que el impacto del LER es más relevante en el comportamiento lineal de la región subumbral de los dispositivos.

La Figura 3.7(a) presenta la σV_{Th} de los transistores impactados por la GER, en relación con la altura RMS para un valor CL de 20 nm. En ella, se observa que la influencia de GER

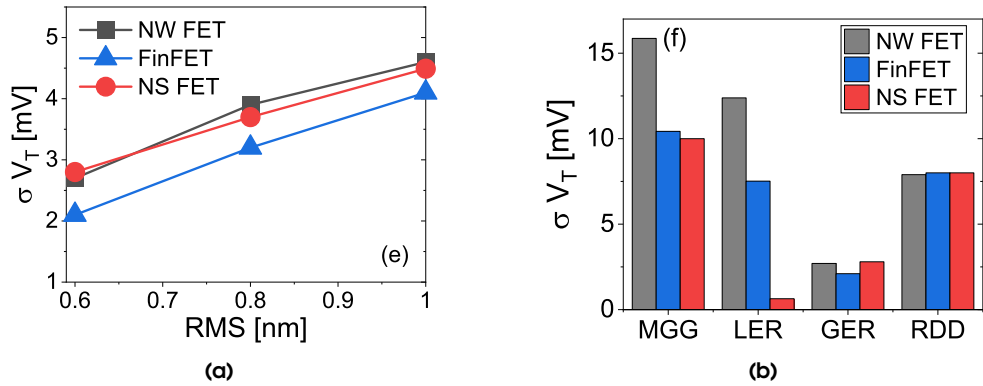


Figura 3.7: (a) σV_{Th} debida al impacto de GER frente al RMS (CL=20 nm) y (b) comparativa del impacto de las diferentes fuentes de variabilidad sobre σV_{Th} para las tres arquitecturas (GS=3 nm y RMS=0.6 nm).

es significativamente menor en comparación con otras fuentes de variabilidad como el MGG o LER. Los resultados revelan que el FinFET se ve menos perturbado por la variabilidad de GER que las arquitecturas NSFET y NWFET. Por ejemplo, para una altura RMS de 0.6 nm, se observa que el FinFET tiene una σV_{Th} aproximadamente un 20% más baja que las arquitecturas NSFET y NWFET (que son muy similares). Para comprender la razón física detrás de este comportamiento, se realizaron simulaciones de FinFETs y NSFETs, donde se mantuvo la longitud total del dispositivo (L) en 40 nm, pero se aumentó/redujo L_g y, por lo tanto, se redujo/incrementó L_s/d en 1 nm para considerar escenarios sintéticos de deformación extrema. Con esta deformación observamos que la reducción/aumento de L_g conlleva una reducción/incremento del valor de V_{Th} (V_{Th}^{def}) con respecto al dispositivo ideal (V_{Th}^{ideal}), esta diferencia la llamaremos $\delta V_{Th} = \|V_{Th}^{ideal} - V_{Th}^{def}\|$. Tanto el FinFET como el NSFET presentaron exactamente los mismos valores de δV_{Th} , lo que indica una sensibilidad similar de ambas arquitecturas a las deformaciones en la puerta. Por lo tanto, la menor variabilidad observada en el FinFET al impacto del GER no puede explicarse por la sensibilidad de su arquitectura, sino que esta diferencia se debe a la parte de la puerta que se encuentra sobre las regiones inferiores laterales (que no envuelven el canal semiconductor), zona que abarca más de 10 nm (vease Figura 3.1(b)). Cualquier deformación GER que afecte a estas zonas laterales no tendrá un impacto en las características I_D-V_G del dispositivo y de tal modo reducirá el impacto del GER en los FinFETs.

Como se muestra en la Figura 3.7(b), las tres arquitecturas exhiben la misma resistencia

($\sigma V_{Th} = 8$ mV) frente a las variaciones por dopantes aleatorios (RDD). Es importante destacar que, como se indica en la Figura 3.1(d), todas las arquitecturas tienen la misma longitud de fuente/drenador y perfil de dopado gaussiano ($N_{s/d}$, X_m , δ). Sin embargo, las variaciones por RDD son más pronunciadas en el FinFET de este trabajo, si lo comparamos con FinFET de trabajos previos con longitud de puerta similar [29], con $\sigma V_{Th} = 5$ mV. Mientras que en [29] informan de casos extremos poco frecuentes con un único dopante en el canal, en nuestros dispositivos el decaimiento del dopaje gaussiano ocurre relativamente cerca de la región de puerta (véase Figura 3.2). Esto aumenta la probabilidad de encontrar un dopante (o incluso más) en esta zona. Si el decaimiento exponencial del dopaje S/D ocurriera a una distancia de 1 nm más allá de la puerta, se produciría una reducción del 50 % de la σV_{Th} observada debido al impacto actual de RDD. Esto resalta la necesidad de una precisión a nanoescala en la implantación iónica.

La Figura 3.7(b) muestra la σV_{Th} para las cuatro fuentes de variabilidad (MGG, LER, GER y RDD) analizadas sobre las tres arquitecturas. Para esta comparativa, se han empleado las configuraciones de parámetros de variabilidad con menos impacto para cada fuente, es decir, un GS de 3 nm para MGG y unos valores CL y RMS de 20 nm y 0.6 nm para LER y GER.

En general, MGG y LER son las dos fuentes de variabilidad que más afectan al rendimiento del dispositivo en la región subumbral. Por lo tanto, el estudio de la variabilidad en la región de encendido se limitará a estas dos fuentes, ya que para obtener resultados fiables necesitamos las simulaciones basadas en el método QC MC. Esta elección de metodología está justificada por lo mostrado en [30] donde se demuestra que la estimación de σI_{on} para MGG y LER realizada por la metodología QC DD sobreestima en hasta un 20 % los valores obtenidos con QC MC.

Las Figuras 3.8(a)-(b) muestran la desviación típica de I_{on} debida al MGG y LER, respectivamente. Para calcular esta desviación, se ha normalizado I_{on} por el perímetro del semiconductor, considerando en el caso particular del FinFET únicamente las regiones que envuelven el semiconductor (paredes laterales y la parte superior, Figura 3.1(d)), ya que la parte inferior no tiene ningún efecto real en la conducción del dispositivo.

De forma similar a lo observado para las FoM de la región subumbral, el NSFET y el FinFET se ven igualmente afectados por la variabilidad MGG en la región de encendido, siendo arquitecturas más resistentes que el NWFET. Esto se observa claramente en la Figura 3.8(a), ya que la σI_{on} del NWFET es como mínimo (GS de 3 nm) un 20 % mayor que las correspondientes

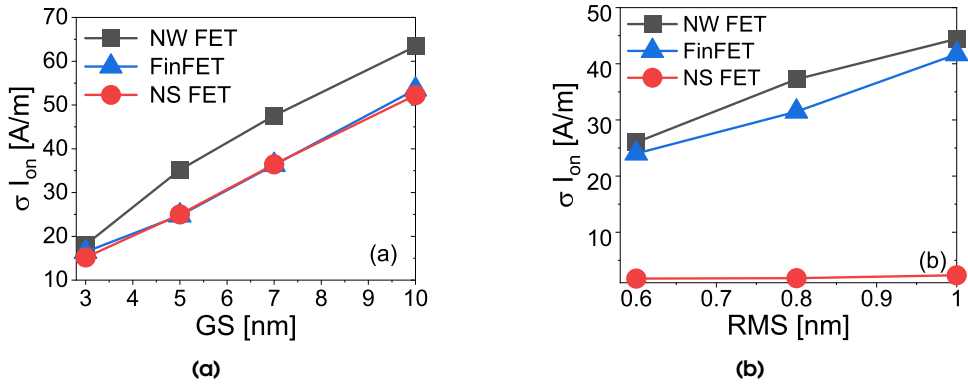


Figura 3.8: Desviación estándar de la corriente de encendido (σI_{on}) debido a las fuentes de variabilidad (a) MGG y (b) LER en función del GS y del RMS, respectivamente, para NWFET, FinFET y NSFET.

al NSFET y el FinFET. La variabilidad inducida por el LER muestra un comportamiento similar tanto en la región subumbral como en la de encendido, con una influencia mínima del LER en σI_{on} para el NSFET (alrededor de 2 A m^{-1}). En el caso del FinFET, observamos en la Figura 3.8(b) que esta arquitectura es ligeramente más resistente que el NWFET al impacto del LER. Por ejemplo, para un valor RMS de 0.8 nm la σI_{on} debido al LER en el FinFET es de 31.5 A m^{-1} , lo cual es un 15 % menor que para el NWFET (37.2 A m^{-1}).

3.3. Desarrollo de modelos empíricos para predecir el impacto de la variabilidad

Los estudios estadísticos del impacto de la variabilidad en el rendimiento de los dispositivos conllevan un elevado coste computacional, ya que es necesario:

- i) Implementar simulaciones 3D que tengan en cuenta los efectos del confinamiento cuántico para estudiar y comprender mejor las propiedades y comportamientos de estos sistemas a escala nanométrica.
- ii) Realizar simulaciones de grandes conjuntos de dispositivos con diferentes configuraciones para obtener resultados estadísticamente significativos. Al simular una gran variedad de configuraciones, se pueden observar las variaciones en las propiedades y comporta-

mientos de los dispositivos, permitiendo obtener una comprensión más completa de su funcionamiento promedio y las posibles variaciones.

Por estos motivos, se han desarrollado técnicas complementarias que permiten predecir la variabilidad de manera más eficiente, reduciendo así el tiempo de cálculo necesario. Algunas de estas técnicas son por ejemplo las basadas en el aprendizaje automático [31, 32, 33, 34, 35], el mapa de sensibilidad de fluctuaciones (FSM) [36], el método de impedancia [37] o el modelo estadístico descrito en [38].

En el caso de los MOSFETs planares clásicos, el impacto de diferentes fuentes de variabilidad puede ser estimado con la Ley de Pelgrom [39] que establece que la desviación estándar de una magnitud ψ ($\sigma\psi$) es proporcional a la inversa de la raíz cuadrada del área efectiva de la puerta (A_G). Esta ley se ha demostrado para diferentes fuentes de variabilidad [40, 41, 42] y puede extenderse a arquitecturas 3D con puerta rodeando el canal de conducción definiendo A_G como el producto de la longitud de puerta y el perímetro efectivo (P^*):

$$\sigma\psi = \frac{A_\psi}{\sqrt{L_g \cdot P^*}}, \quad (3.1)$$

donde A_ψ es el coeficiente de Pelgrom de proporcionalidad de ψ , que está determinado por las contribuciones de todas las posibles fuentes de variación del transistor [43] sobre esa magnitud.

En este trabajo se desarrolla un modelo predictivo basado en la Ley de Pelgrom (*Pelgrom-based predictive model* (PBP)) para estimar el impacto de MGG y LER sobre los transistores previamente calibrados. Este modelo será validado con una combinación de resultados extraídos de la literatura [14, 18, 19, 20, 29, 44, 45, 46, 47, 48] y simulaciones propias realizadas en VENDES [8] en el caso de necesitar completar los datos de la literatura.

El modelo PBP se desarrollará para la predicción de σV_{Th} y σI_{On} que corresponden a simulaciones en la región subumbral (QC DD) y a la región de encendido (QC MC), respectivamente. En ambos casos también será aplicado a las dos fuentes de variabilidad que más impacto tienen sobre los dispositivos estudiados, MGG y LER, como se demostró en la sección 3.2.

3.3.1. Modelo PBP para la predicción de σV_{Th}

Mostramos en la Figura 3.9 las secciones transversales 2D, de todos los dispositivos empleados para este trabajo con la referencia correspondiente a sus autores. Todos los dispositivos

analizados tienen una puerta metálica de nitruro de titanio (TiN). A la hora de desarrollar el modelo PBP para la predicción de σV_{Th} , es importante destacar que aunque se estudian tres arquitecturas por separado (FinFET, NWFET y NSFET), existen diferencias entre las formas de los dispositivos correspondientes a la misma arquitectura entre unos autores y otros. Esto se observa en los FinFETs de las Figuras 3.9(a)-(c) y los NWFETs de las Figuras 3.9(d)-(e).

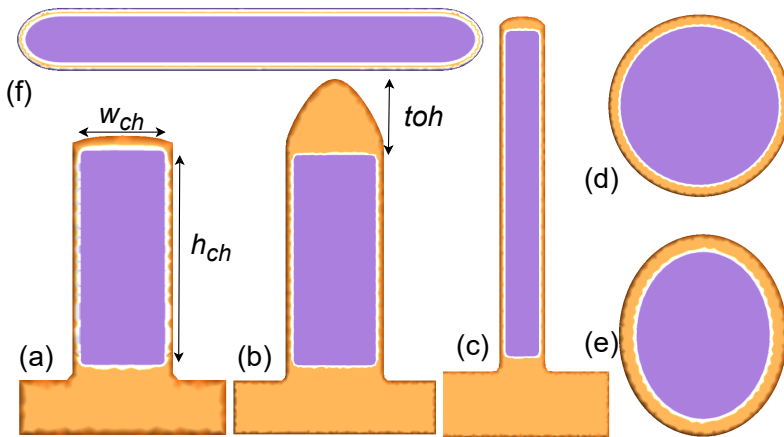


Figura 3.9: Sección transversal de las tres arquitecturas usadas en este trabajo. Los FinFETs: (a) de Wang (29) y Leung (48), (b) de Seoane (44), Indalecio (14) y (c) de Seoane (20). Los NWFETs: (d) de Lü (45), Sung (47) y (e) de Nagy (46) y Elmessary (19). Los NSFETs: (f) de Nagy (18) y simulados para este trabajo. w_{ch} y h_{ch} son el ancho y alto del canal respectivamente. Denotamos como toh la altura del óxido en la cúpula del FinFET tipo-(b).

La mayoría de los FinFETs [14, 44], NWFETs [19, 47] y NSFETs [18] están basados en dispositivos experimentales (FinFET [4], NWFET [5], NSFET [6]) o son versiones escaladas de los mismos. Obsérvese que los FinFET de las Figuras 3.9(a)-(b) tienen una relación anchura/altura del semiconductor similar, $w_{ch}/h_{ch} \sim 0.4$ y difieren en la altura del óxido superior (toh), mientras que el FinFET de la Figura 3.9(c) tiene una relación completamente distinta, $w_{ch}/h_{ch} \sim 0.1$. El FinFET diseñado en [20] (Figura 3.9(c)) es una versión tri-puerta del NSFET [6], basado en las predicciones actuales del IRDS [3]. Además, las secciones transversales de los NWFET estudiados son circulares ($w_{ch} = h_{ch}$, véase la Figura 3.9(d)) o elípticas ($w_{ch} \neq h_{ch}$, ver Figura 3.9(e)). Los NSFET modelados específicamente para este trabajo (Figura 3.9(f)) tienen una relación $w_{ch}/h_{ch} \sim 10$ basada en [6].

Las simulaciones propias realizadas con VENDES correspondientes a la región subumbral (empleadas para estudiar V_{Th}), se han realizado con la metodología QC DD.

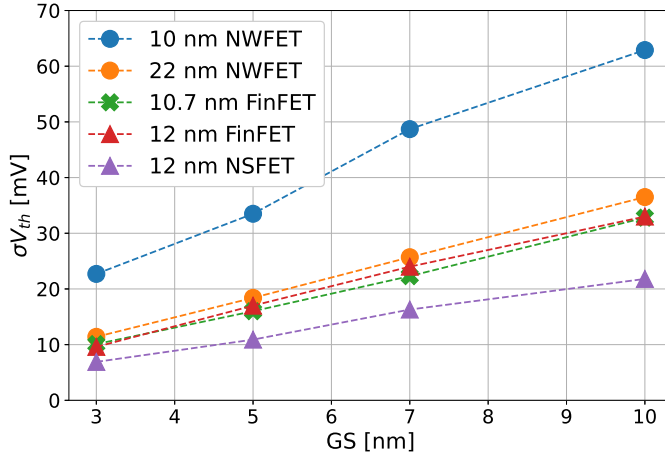


Figura 3.10: Desviación estándar del voltaje umbral σV_{Th} frente al GS para la variabilidad MGG comparando NWFETs, FinFETs y NSFET con diferentes longitudes de puerta.

Modelo PBP subumbral aplicado a MGG

Para desarrollar el modelo PBP subumbral primero sustituimos la variable genérica ψ por V_{Th} y luego consideramos que MGG es la única fuente de variabilidad que contribuye al coeficiente Pelgrom de proporcionalidad (A_v). También asumimos que V_{Th} aumenta linealmente con la GS, como se muestra en la Figura 3.10 para diferentes arquitecturas y longitudes de puerta, y como también se demostró anteriormente en [49], por lo tanto:

$$A_v = \theta_v^{mgs} \cdot GS, \quad (3.2)$$

siendo θ_v^{mgs} el desfase de tensión umbral inducido por la variabilidad MGG. Además, como la MGG es una variabilidad inducida sobre la puerta, en la Ley de Pelgrom (3.1), el perímetro efectivo (P^*) será al perímetro de puerta (P_g). El modelo PBP subumbral aplicado a la predicción del impacto de la variabilidad MGG será:

$$\sigma V_{Th} = \frac{\theta_v^{mgs}}{\sqrt{L_g \cdot P_g}} \cdot GS. \quad (3.3)$$

Dado que el impacto de MGG sobre los transistores no está relacionado con V_D , como se ha demostrado en [44, 46], asumimos que el parámetro θ_v^{mgs} depende exclusivamente de la arquitectura seleccionada y del metal utilizado durante el proceso de deposición de la puerta.

Una vez presentadas las hipótesis del modelo, procedemos a validarlo y comparar los resultados de sus predicciones con los resultados de las simulaciones. Estos estudios computacionales realizados para evaluar el impacto de MGG, tanto en este estudio como en Indalecio [14], Nagy [18], Seoane21 [20], Wang [29], Seoane16 [44], Nagy [46] y Sung [47], se llevaron a cabo utilizando la metodología QC DD. Sin embargo, en Lü [45] se realizaron simulaciones utilizando el método DD clásico.

La Figura 3.11 muestra las gráficas de Pelgrom para σV_{Th} debido a MGG en cada una de las arquitecturas: Figura 3.11(a) para FinFETs, Figura 3.11(b) para NWFETs y Figura 3.11(c) para NSFETs. En estas gráficas se presentan los resultados correspondientes a diferentes valores de GS, así como los valores de A_v . En las Figuras 3.11 se observa que existe una relación lineal entre σV_{Th} y $(L_g P_g)^{-1/2}$, siguiendo el modelo PBP para MGG consecuentemente con la ecuación (3.3). Se espera que cuanto mayor sea el área de la puerta, menor sea la influencia de los granos metálicos. Por lo tanto, desde un punto de vista macroscópico, cuando $(L_g P_g)^{-1/2} \rightarrow 0$, la puerta se comportará como si tuviera un valor de función de trabajo (ϕ_m) uniforme, lo que implica $\sigma V_{Th} \rightarrow 0$ V.

La pendiente de las gráficas de Pelgrom (A_v) aumenta linealmente con el GS para las tres arquitecturas, validando la asunción realizada en la ecuación (3.2). Esta relación lineal para cada arquitectura y metal en la puerta está marcada por un θ_v^{mgs} común. P_g corresponde al perímetro de puerta que rodea al semiconductor para todos los dispositivos presentados en la Figura 3.9, excepto para los FinFET con la forma presentada en la Figura 3.9(b), ya que, como se demostró en [50] los granos situados en la parte superior de la puerta de estos dispositivos no influyen en σV_{Th} . Por lo tanto, para tener en cuenta solo los granos situados en las paredes laterales, para los FinFET con la forma de la Figura 3.9(b) definimos el perímetro efectivo de la puerta como $P_g = 2h_{ch}$. También es relevante destacar que la relación $\theta_v^{mgs} / \sqrt{L_g P_g}$ es un factor que determina la sensibilidad de un determinado dispositivo al impacto generado por MGG.

En la Tabla 3.1, mostramos una comparación de la σV_{Th} predicha por el modelo PBP (σV_{Th}^P), frente a la σV_{Th} procedente de estudios de simulación (σV_{Th}^S) disponibles en la literatura. También se muestran el GS, el error relativo entre los resultados predichos y simulados (σ_r), el área efectiva de puerta ($L_g \times P_g$) y θ_v^{mgs} . En esta tabla, [*] indica los datos simulados específicamente para este estudio. Obsérvese que, como se ha mencionado anteriormente, para una arquitectura de dispositivo y un metal en la puerta concretos, θ_v^{mgs} no cambia. Para determinar este parámetro para una arquitectura es necesaria la simulación de un

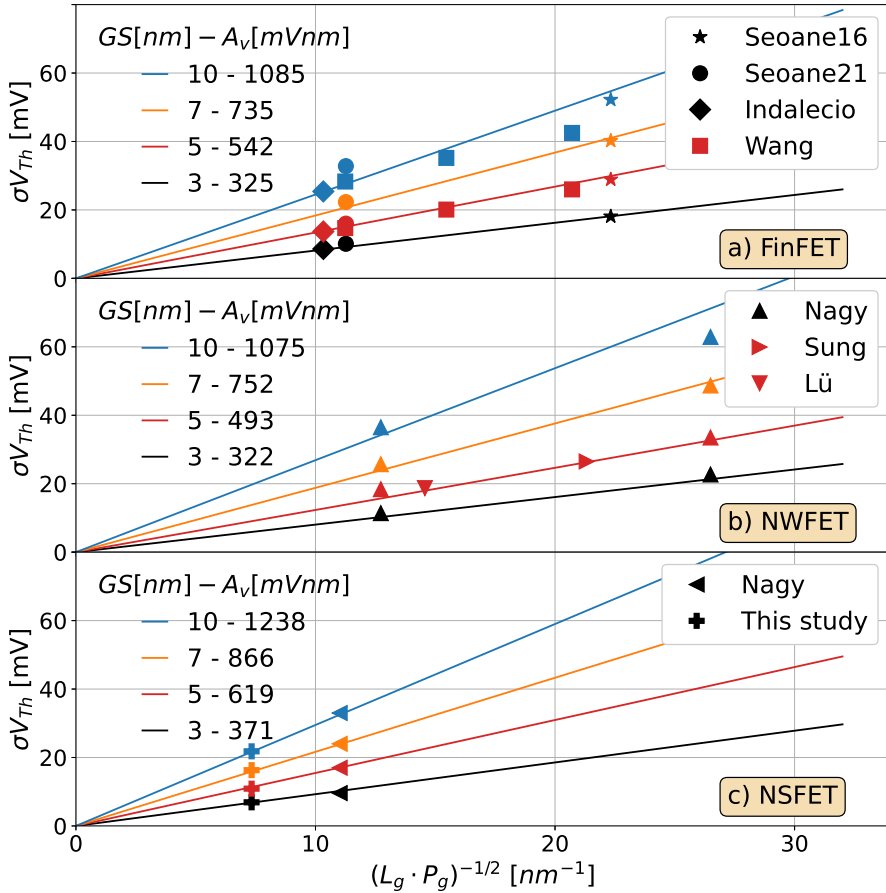


Figura 3.11: Gráficas de Pelgrom evaluando la σV_{Th} debida al impacto de MGG con diferentes GS para: (a) FinFETs extraídos de Seoane16 (44), Seoane21 (20), Indalecio (14), Wang (29), (b) NWFETs reportados por Nagy (46), Sung (47), Lü (45), y (c) NSFETs de Nagy (18) conjuntamente con simulaciones propias realizadas para este estudio. Obsérvese cómo los dispositivos siguen con exactitud la Ley de Pelgrom, además el coeficiente Pelgrom de proporcionalidad para V_{Th} (A_v) aumenta linealmente con el GS.

Tabla 3.1: Comparación de la σV_{Th} predicha por el modelo PBP (σV_{Th}^P), frente a la σV_{Th} procedente de estudios de simulación (σV_{Th}^S) disponibles en la literatura. También se muestran el GS, el error relativo entre los resultados predichos y simulados (σ_r), el área efectiva de puerta ($L_g \times P_g$) y θ_v^{mgs} . (*) indica los datos simulados específicamente para este estudio. Los σ_r mayores al 10% están marcados en negrita.

Arch.	θ_v^{mgs} [mV]	GS [nm]	σV_{Th}^P [mV]	σV_{Th}^S [mV]	σ_r [%]	$L_g \times P_g$ [nm ²]	Ref.		
FinFET	108	3	18.2	18.1	+0.6	10.7×30.0	[44]		
			9.2	10.1	-9.2	12.0×105.0	[20]		
			8.4	8.6	-2.3	25.0×60.0	[14]		
		5	30.3	28.9	+4.8	10.7×30.0	[44]		
			28.1	26.0	+8.1	10.0×37.3	[29]		
			21.0	20.1	+4.3	14.0×47.8	[29]		
			15.2	14.7	+3.3	20.0×63.3	[29]		
			15.3	16.0	-4.6	12.0×105.0	[20]		
		7	14.0	13.7	+2.2	25.0×60.0	[14]		
			42.4	40.2	+5.5	10.7×30.0	[44]		
		8	21.4	22.3	-4.0	12.0×105.0	[20]		
			22.4	21.8	+2.8	25.0×60.0	[14]		
		10	60.6	52.2	+16.1	10.7×30.0	[44]		
			56.1	42.5	+32.0	10.0×37.3	[29]		
			41.9	35.2	+16.0	14.0×47.8	[29]		
			29.1	28.3	+2.7	20.0×63.3	[29]		
			30.6	32.8	-6.7	12.0×105.0	[20]		
			28.0	25.4	+10.2	25.0×60.0	[14]		
		NWFET	107	3	21.4	22.7	-5.7	10.0×22.8	[46]
					10.3	11.4	-9.6	22.0×44.9	[46]
5	35.6			33.5	+6.3	10.0×22.8	[46]		
	28.7			26.5	+8.3	10.0×35.2	[47]		
	19.6			18.7	+4.8	20.0×37.7	[45]		
	17.1			18.4	-7.1	22.0×44.9	[46]		
7	49.8			48.7	+2.3	10.0×22.8	[46]		
	23.9			25.7	-7.0	22.0×44.9	[46]		
10	71.2			62.9	+13.2	10.0×22.8	[46]		
	34.2			36.5	-6.3	22.0×44.9	[46]		
NSFET	124	3	10.2	9.6	+6.3	12.0×110.0	[18]		
			6.8	6.9	-1.4	18.0×165.3	[*]		
		5	17.0	17.0	0.0	12.0×110.0	[18]		
			11.3	10.9	+3.7	18.0×165.3	[*]		
		7	23.9	24.0	-0.4	12.0×110.0	[18]		
			15.9	16.3	-2.5	18.0×165.3	[*]		
		10	34.1	33.0	+3.3	12.0×110.0	[18]		
			22.7	21.8	+4.1	18.0×165.3	[*]		

estudio de variabilidad (un dispositivo afectado por varios GS diferentes). Una vez que θ_v^{mgg} está determinado, el impacto de MGG para cualquier GS y cualquier dimensión de dispositivo (diferentes combinaciones de $L_g, P_g, w_{ch}/h_{ch}, t_{ox}$) puede estimarse utilizando el modelo PBP según la ecuación (3.3).

Las mayores imprecisiones en la predicción se deben a la saturación de las desviaciones estándar, algo que ocurre cuando el número de granos metálicos en la puerta es muy pequeño. El número de granos lo podemos estimar como el ratio entre A_G y el área promedio de los granos (GS^2 , asumiendo granos cuadrados). Por ejemplo, para un GS de 10 nm, solo habría en promedio 3 granos metálicos en la puerta del FinFET con $A_G = 10.7 \times 30.0 = 321 \text{ nm}^2$, transformando la distribución normal de V_{Th} [14] en una distribución bimodal con picos alrededor de los valores extremos de ϕ_m (4.4 y 4.6 eV) y rompiendo la linealidad asumida por el modelo PBP. Excepto los casos extremos mencionados anteriormente, que corresponden con las estimaciones marcadas en negrita en la Tabla 3.1, todas las predicciones tienen un error relativo del 10 % o inferior, destacando que el 91 % de los errores son del 8 % o menores.

Modelo PBP subumbral aplicado a LER

En esta sección, consideramos que la variabilidad LER será el único contribuyente a las variaciones recogidas en A_v . La σV_{Th} inducida por la variabilidad LER aumenta linealmente con RMS, como se muestra en la Figura 3.12(a) y en la literatura [14, 19, 20, 44, 46]. Asimismo, σV_{Th} es inversamente proporcional a la raíz de la anchura del canal (w_{ch} , véase la Figura 3.9), que corresponde con la dimensión crítica de la deformación, ya que cuanto más estrecho sea el dispositivo, mayor será el efecto esperado de LER. Además, σV_{Th} es proporcional a una función de la longitud de correlación ($f(CL)$). Esta $f(CL)$ se define empíricamente a partir de la función de autocorrelación que reproduce la rugosidad de la superficie a través de la transformada de Fourier del espectro gaussiano [10, 14]. Teniendo en cuenta todos estos factores, el coeficiente Pelgrom de proporcionalidad A_v será tal que:

$$A_v = \theta_v^{ler} \cdot \frac{\text{RMS}}{\sqrt{w_{ch}}} \cdot f(CL), \quad (3.4)$$

siendo,

$$f^2(CL) = CL \cdot \left(1 - e^{-\sqrt{L_g/CL}}\right), \quad (3.5)$$

donde θ_v^{ler} es el desfase del voltaje umbral para la variabilidad LER. La Figura 3.12(b) muestra cómo σV_{Th} aumenta con la relación CL/L_g para un NWFET de 10 nm con $\text{RMS} = 0.4 \text{ nm}$ y un

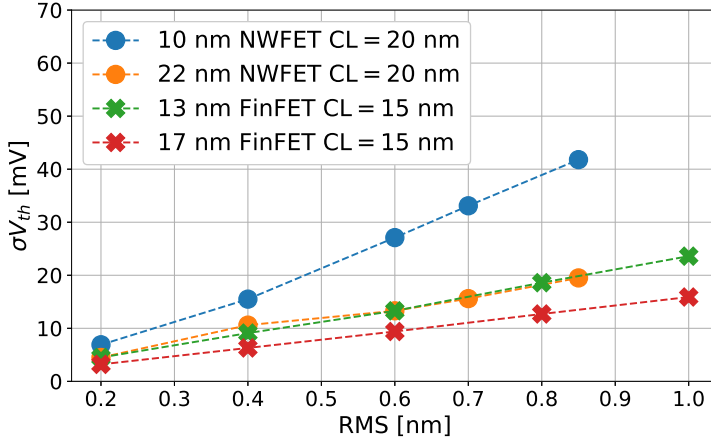
FinFET de 12 nm con RMS= 0.6 nm. Las líneas continuas de esta figura representan la $f(\text{CL})$ (véase la ecuación (3.5)) ajustada a los datos y las barras de error verticales corresponden al 10% de los valores de σV_{Th} . Dado que el LER afecta a la anchura del semiconductor estrechándola o ensanchándola, en la Ley de Pelgrom (ecuación (3.1)), P^* será el perímetro del semiconductor (P_{sc}). Con los conceptos anteriores, el modelo PBP para la variabilidad subumbral inducida por el LER es:

$$\sigma V_{Th} = \theta_v^{ler} \cdot \frac{\text{RMS}}{\sqrt{w_{ch}}} \cdot \frac{f(\text{CL})}{\sqrt{L_g P_{sc}}} \quad (3.6)$$

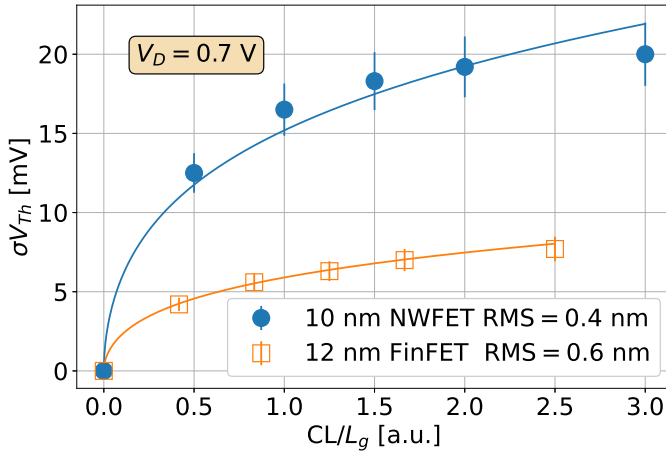
A diferencia de la variabilidad MGG, la V_D tiene un gran impacto en el aumento de σV_{Th} para LER [19, 44, 46]. En la Figura 3.13 se presenta σV_{Th} frente a $\text{RMS} \cdot w_{ch}^{-1/2} \cdot f(\text{CL}) \cdot (L_g P_{sc})^{-1/2}$ para diferentes dispositivos FinFET (figura superior) y NWFET (figura inferior). También se muestran los valores de θ_v^{ler} , que se determinan como la pendiente de las gráficas y como se observa crece con V_D para las dos arquitecturas mostradas.

La variabilidad LER no se estudia para el NSFET ya que, como se demostró en la sección 3.2 el impacto en esta arquitectura es despreciable ($\sigma V_{Th} \leq 1.0$ mV). Las simulaciones debidas a la variabilidad LER reportadas en Indalecio [14], Elmessary [19], Seoane21 [20], Wang [29] y Seoane16 [44] se realizaron con la metodología QC DD. Para cada arquitectura y V_D , todos los dispositivos independientemente de su forma (ver Figura 3.9(a)-(e)) comparten un θ_v^{ler} común. Una vez determinado θ_v^{ler} , el modelo PBP puede utilizarse para predecir el impacto del LER para cualquier conjunto de deformaciones (RMS, CL y dimensiones de los dispositivos (L_g, P_{sc}, w_{ch})).

Nótese que aunque θ_v^{ler} no tiene significado físico, la relación $\theta_v^{ler} / \sqrt{L_g P_{sc}}$ proporciona una estimación de la sensibilidad de un dispositivo a la variabilidad del LER. Las Tablas 3.2 y 3.3 muestran la comparación entre los valores predichos (σV_{Th}^P) y los datos simulados (σV_{Th}^S), con el fin de validar el modelo para FinFETs y NWFETs impactados por LER. Se presenta un amplio rango de valores con CL comprendidos entre 10 nm y 50 nm y RMS entre 0.2 nm y 1.0 nm. Visualizando los resultados, todos los errores relativos (σ_r) entre los valores predichos y simulados de σV_{Th} están por debajo del 9%, con el 94% de las predicciones siendo menores del 6%, demostrando la validez de la estimación de σV_{Th} realizada por el modelo PBP para estudiar el impacto del LER en la región subumbral.



(a)



(b)

Figura 3.12: (a) σV_{Th} frente a RMS debido a la variabilidad LER para NWFET y FinFET con diferentes longitudes de puerta, siendo CL es la longitud de correlación. Nótese la relación lineal entre σV_{Th} y RMS. (b) σV_{Th} frente a la relación entre la longitud de correlación y la longitud de puerta (CL/L_g) para dos arquitecturas: FinFET (14) y NWFET (19). También se muestra el valor RMS para cada arquitectura. Las líneas continuas representan el ajuste de cada dispositivo a su propia $f(CL)$ y las barras de error y corresponden al 10% de cada valor de σV_{Th} . Obsérvese el buen ajuste de $f(CL)$ para los dos dispositivos presentados. Las simulaciones se realizaron con un V_D de 0.7 V.

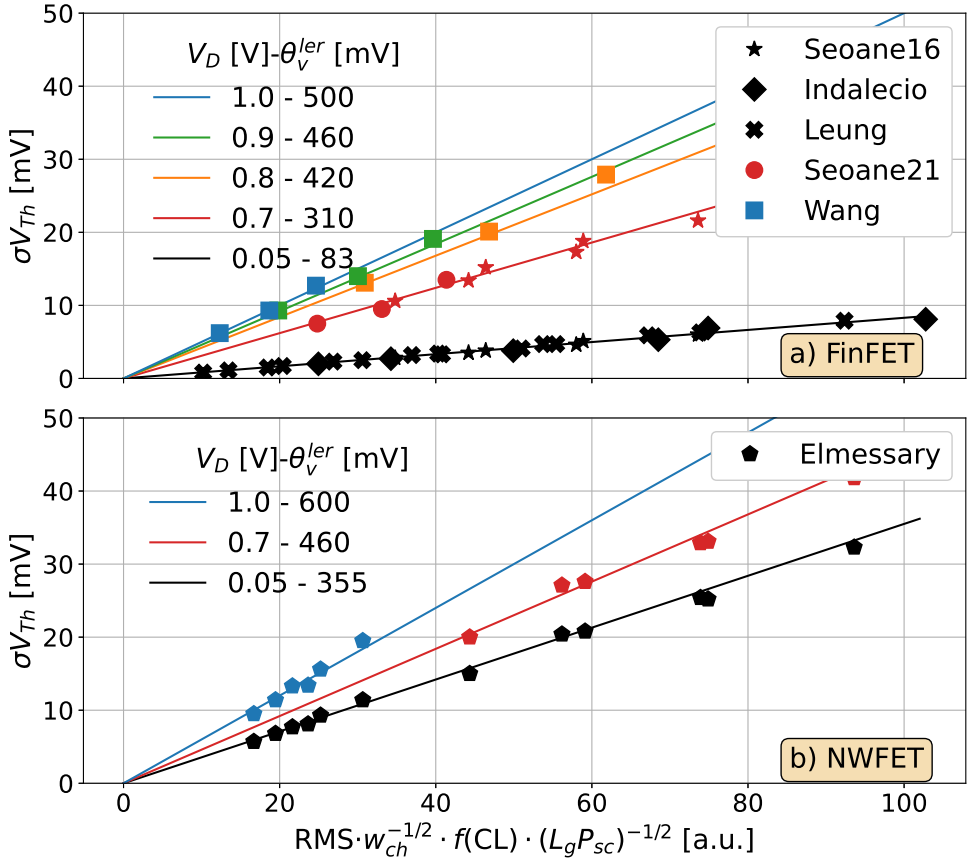


Figura 3.13: σV_{Th} debidos al impacto del LER frente a $RMS \cdot w_{ch}^{-1/2} \cdot f(CL) \cdot (L_g P_{sc})^{-1/2}$ a diferentes valores de V_D , para: (a) FinFETs, (b) NWFETs. Datos extraídos de Seoane16 (44), Seoane21 (20), Indalecio (14), Wang (29), Leung (48) y Elmessary (19). Obsérvese el excelente ajuste del modelo con los datos simulados y el aumento de θ_v^{ler} con V_D .

Tabla 3.2: Validación del modelo PBP para LER aplicado sobre el FinFET para la estimación de σV_{Th} . V_D es el potencial de drenaje, θ_v^{ler} el desfase del voltaje umbral para la variabilidad LER y $L_g \times P_{sc}$ el área del semiconductor. También se muestran CL, RMS, los valores de σV_{Th} predichos σV_{Th}^P y simulados σV_{Th}^S junto con su error relativo σ_r .

V_D [mV]	θ_v^{ler} [mV]	CL [nm]	$L_g \times P_{sc}$ [nm ²]	RMS [nm]	σV_{Th}^P [mV]	σV_{Th}^S [mV]	σ_r [%]	Ref.
0.05	83	50	25.0×72.0	1.00	2.8	2.7	+3.7	[14]
				2.00	5.6	5.3	+5.7	
				3.00	8.4	8.1	+3.7	
				0.60	3.7	3.5	+5.7	
				0.80	4.9	5.1	-3.9	
				1.00	6.1	6.0	+1.7	
		20	10.7×35.0	0.60	3.7	3.5	+5.7	[44]
				0.80	4.9	5.1	-3.9	
				1.00	6.1	6.0	+1.7	
				1.00	2.1	2	+5.0	
				2.00	4.1	3.8	+7.9	
				3.00	6.3	6.9	-8.7	
		15	13.0×12.8	0.20	1.5	1.5	0.0	[48]
				0.40	3.1	3.2	-3.1	
				0.60	4.6	4.7	-2.1	
				0.80	6.1	6.3	-3.2	
				1.00	7.7	7.9	-2.5	
				0.20	1.1	1.1	0.0	
		15	17.0×16.0	0.40	2.3	2.3	0.0	[48]
				0.60	3.4	3.4	0.0	
				0.80	4.6	4.7	-2.1	
				1.00	5.7	5.9	-3.4	
				0.20	0.8	0.8	0.0	
				0.40	1.7	1.7	0.0	
15	22.0×19.2	0.60	2.5	2.5	0.0	[48]		
		0.80	3.4	3.3	+3.0			
		1.00	4.2	4.1	+2.4			
		0.60	2.9	2.8	+3.6			
		0.80	3.9	3.8	+2.6			
		1.00	4.8	4.6	+4.2			
0.70	310	20	10.7×35.0	0.60	13.7	13.4	+2.2	[44]
				0.80	18.3	18.8	-2.7	
				1.00	22.8	21.6	+5.6	
				0.60	7.7	7.5	+2.7	
				0.80	10.3	9.5	+8.4	
				1.00	12.8	13.5	-5.2	
0.80	420	30	10.0×37.3	0.60	10.8	10.6	+1.9	[44]
				0.80	14.4	15.2	-5.3	
				1.00	18.0	17.3	+4.0	
				0.33	13.0	13.1	-0.8	
				0.50	19.7	20.1	-2.0	
				0.66	26.0	27.9	-6.8	
0.90	460	30	14.0×47.8	0.33	9.1	9.3	+2.1	[29]
				0.50	13.8	14.0	+1.4	
				0.66	18.2	19.1	-4.7	
				0.33	6.2	6.2	0.0	
				0.50	9.3	9.3	0.0	
				0.66	12.3	12.7	+3.1	
1.00	500	30	20.0×63.3	0.33	6.2	6.2	0.0	[29]
				0.50	9.3	9.3	0.0	
				0.66	12.3	12.7	+3.1	

Tabla 3.3: Validación del modelo PBP para LER aplicado sobre el NWFET para la estimación de σV_{Th} . V_D es el potencial de drenaje, θ_V^{ler} el desfase del voltaje umbral para la variabilidad LER, la longitud de correlación CL, $L_g P_{sc}$ el área del semiconductor y la raíz cuadrática media de la deformación RMS. También se han calculado los valores predichos σV_{Th}^P y la tensión umbral simulada σV_{Th}^S junto con su error relativo σ_r .

V_D [mV]	θ_V^{ler} [mV]	CL [nm]	$L_g \times P_{sc}$ [nm ²]	RMS [nm]	σV_{Th}^P [mV]	σV_{Th}^S [mV]	σ_r [%]	Ref.
0.05	355	20	10.0×20.3	0.60	19.9	20.4	-2.5	[19]
				0.70	26.6	25.2	+5.6	
				0.85	33.2	32.3	+2.8	
		10	22.0×40.2	0.60	7.7	7.7	0.0	[19]
				0.70	9.0	9.3	-3.2	
				0.85	10.9	11.4	-4.4	
	460	20	10.0×20.3	0.60	15.7	15.0	+4.6	[19]
				0.70	21.0	20.8	+1.0	
				0.85	26.2	25.4	+3.1	
		10	22.0×40.2	0.60	5.9	5.7	+3.5	[19]
				0.70	6.9	6.8	+1.5	
				0.85	8.4	8.1	+3.7	
0.70	460	20	10.0×20.3	0.60	25.8	27.1	-4.8	[19]
				0.70	34.4	33.1	+3.9	
				0.85	43.1	41.8	+3.1	
	10	10.0×20.3	0.60	20.4	20.0	+2.0	[19]	
			0.70	27.2	27.6	-1.5		
			0.85	34.0	32.9	+3.3		
1.00	600	20	22.0×40.2	0.60	13.0	13.3	-2.3	[19]
				0.70	15.1	15.6	-3.2	
				0.85	18.4	19.5	-5.6	
	10	22.0×40.2	0.60	10.0	9.5	+5.3	[19]	
			0.70	11.7	11.4	+2.6		
			0.85	14.2	13.3	+6.8		

3.3.2. Modelo PBP para la predicción de σI_{on}

Una vez hemos desarrollado el modelo PBP subumbral, procedemos a realizar un procedimiento análogo para conseguir un método similar para la región de encendido. En este caso, el coste computacional realizar estudios de variabilidad 3D en la región de encendido es más prohibitivo. Esto se debe a que, para capturar adecuadamente los fenómenos de transporte fuera del equilibrio, se necesita aplicar la metodología QC MC o otras metodologías completamente cuánticas (NEGF), ya que el modelo de transporte de DD no es capaz de capturar con exactitud estos efectos. Al ser estas metodologías mucho más costosas computacionalmente, la necesidad de desarrollar nuevas estrategias para el estudio de las figuras de mérito

correspondientes a la región de encendido, es aun más crítica que para la región subumbral.

En esta sección, desarrollaremos el modelo PBP para estimar la variabilidad inducida en la corriente de encendido I_{on} por las dos mismas fuentes de variabilidad estudiadas en la sección anterior MGG y LER. Por tanto se sustituirá la variable ψ en la ecuación (3.1) por I_{on} :

$$\sigma I_{on} = \frac{A_i}{\sqrt{L_g P^*}} \quad (3.7)$$

Para el estudio de la región de encendido a través de la I_{on} combinaremos datos extraídos de la literatura [18, 44, 46, 47] con simulaciones QC MC realizadas en VENDES (véase descripción en la sección 2.2).

Modelo PBP de encendido aplicado a MGG

Para desarrollar el modelo PBP aplicado a MGG en la región de encendido, realizamos de nuevo la asunción de que el MGG es la única fuente que contribuye a las variaciones recogidas en A_i . Al igual que para el caso de V_{Th} la dependencia respecto al GS es lineal como se ha demostrado en [18, 44] o se puede observar en la Figura 3.8(a), por ello, el modelo PBP para MGG en la región de encendido será:

$$A_i = \theta_i^{mgs} \cdot GS, \quad (3.8)$$

que sustituyendo en la ecuación (3.1):

$$\sigma I_{on} = \frac{\theta_i^{mgs}}{\sqrt{L_g P_g}} \cdot GS \quad (3.9)$$

donde θ_i^{mgs} es el desfase de corriente de encendido inducido por la variabilidad MGG, que es un parámetro tecnológico que establece cómo afecta la variabilidad a una cierta arquitectura. Una vez este parámetro está determinado para una arquitectura, puede predecirse la σI_{on} debida al impacto de MGG para dispositivos con diferentes dimensiones y valores de GS.

En la Figura 3.14(a)-(c) se muestra la comparativa del impacto del MGG para las tres arquitecturas. σI_{on}^S son los datos provenientes de simulación y σI_{on}^P las predicciones realizadas con esta versión del modelo PBP. La línea negra sólida es la bisectriz del primer cuadrante que denotaría correlación perfecta ($y = x$) y las barras de error corresponden al 10 % del error relativo con respecto a σI_{on}^S , una tolerancia que consideramos aceptable. Como puede verse, la mayoría de las predicciones están dentro de este margen del 10 %.

La Tabla 3.4 presenta las σI_{on}^P y σI_{on}^S junto con el error relativo (σ_r) entre ellas, el GS, el θ_i^{mgs} y el área de la puerta ($L_g \times P_g$). Las estimaciones fuera del margen de error del 10 % se destacan en negrita en la tabla y se corresponden a la saturación de la desviación típica en dispositivos con pocos granos en la puerta. Este caso se da cuando el número de granos, calculado como el ratio entre A_G y el área promedio de los granos (GS^2 , asumiendo granos cuadrados), es muy pequeño. Por ejemplo, un NWFET con $A_G = 10 \times 22.8 = 228 \text{ nm}^2$ con $GS = 7 \text{ nm}$ o 10 nm tendría únicamente 5 o 2 granos metálicos, este fenómeno se describe ampliamente en [50]. En este caso, tal como se muestra en la tabla, el 93 % de los valores de σI_{on} predichos están dentro del margen del 10 % de σ_r .

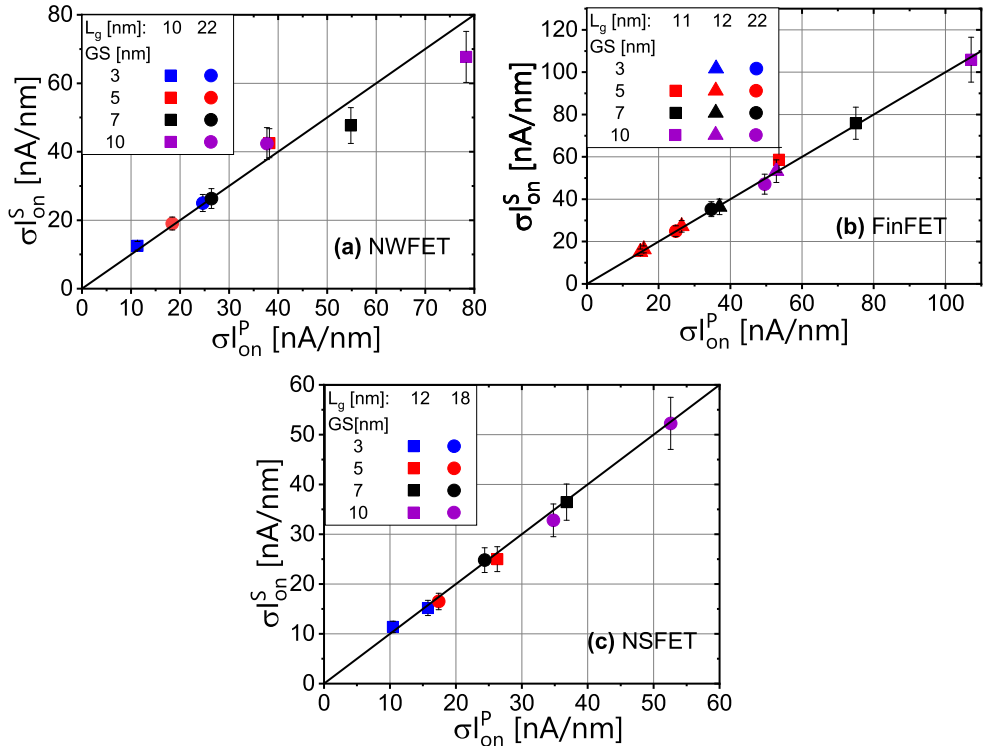


Figura 3.14: σI_{on} simulada en VENDES (σI_{on}^S) frente a la σI_{on} predicha por el modelo PBP (σI_{on}^P), para diferentes arquitecturas (a) NWFET, (b) FinFET y (c) NSFET, con diferentes longitudes de puerta (L_g) e impactados por diferentes GS. La línea sólida negra es la bisectriz del primer cuadrante e indica correlación perfecta entre los valores predichos con el modelo y simulados. Las barras de error corresponden al 10% de los valores simulados, margen que consideramos aceptable.

Tabla 3.4: σI_{on} predichas (σI_{on}^P) y simuladas (σI_{on}^S), a diferentes valores de GS para cada arquitectura. En la tabla también listamos el desfase de corriente para MGG θ_i^{mgs} , el área efectiva de la puerta ($L_g \times P_g$) y el error relativo σ_r . Las simulaciones realizadas específicamente para este trabajo se denotan con *. Los σ_r mayores al 10 % están marcados en negrita.

Arq.	θ_i^{mgs} [nA nm ⁻²]	GS [nm]	σI_{on}^P [A m ⁻¹]	σI_{on}^S [A m ⁻¹]	σ_r [%]	$L_g \times P_g$ [nm ²]	Ref.		
NWFET	120	3	23.8	25.0	-4.8	10.0×22.8	[46]		
			11.5	12.5	-8.0	22.0×44.9	[46]		
		5	39.7	42.5	-6.6	10.0×22.8	[46]		
			19.1	19.0	+0.5	10.0×35.2	[47]		
			32.0	33.0	-5.3	22.0×44.9	[46]		
			26.7	26.3	+1.5	22.0×44.9	[46]		
		10	79.5	67.7	+17.4	10.0×22.8	[46]		
			38.2	42.4	-9.9	22.0×44.9	[46]		
		FinFET	192	3	14.9	15.0	-0.7	25.0×60.0	*
					16.3	16.3	+0.0	12.0×105.0	*
5	52.7			58.6	-8.4	10.7×30.0	[44]		
	24.9			24.9	+0.0	25.0×60.0	*		
	27.1			27.1	+0.0	12.0×105.0	*		
	38.0			35.3	+7.6	12.0×105.0	*		
10	107.4			105.9	+1.4	10.7×30.0	[44]		
	49.7			47.1	+5.5	25.0×60.0	*		
	54.2			53.3	+1.7	12.0×105.0	*		
NSFET	191			3	15.8	15.2	+3.9	12.0×110.0	[18]
		10.5	11.4		-7.9	18.0×165.3	*		
		5	26.3	25.0	+5.2	12.0×110.0	[18]		
			17.5	16.5	+6.1	18.0×165.3	*		
		7	36.8	36.5	+1.0	12.0×110.0	[18]		
			24.5	24.8	-1.2	18.0×165.3	*		
		10	52.6	52.3	+0.7	12.0×110.0	[18]		
			35.0	32.8	+6.7	18.0×165.3	*		

Modelo PBP de encendido aplicado a LER

En este caso, el LER es el único factor que contribuye a la variabilidad recogida por A_i . Además, suponemos que σI_{on} aumenta linealmente con RMS (véase [20, 46] y Figura 3.8(b)). También depende de la raíz cuadrada de la dimensión de deformación del LER (anchura del canal, w_{ch}) y una de función de $f(\text{CL})$ que se define empíricamente igual que para la tensión

umbral con la relación (3.5). Por lo tanto, el modelo PBP para predecir el impacto del LER sobre la región de encendido será,

$$\sigma I_{on} = \theta_i^{ler} \cdot \text{RMS} \cdot f(\text{CL}) \cdot \sqrt{\frac{w_{ch}}{L_g P_{sc}}} \quad (3.10)$$

θ_i^{ler} es el desfase en corriente de encendido para LER, un parámetro que establece la dependencia de la variabilidad debida a una arquitectura. Una vez este parámetro es calibrado para una arquitectura y V_D , se puede estimar σI_{on} para dispositivos con cualquier combinación de L_g , P_{sc} , o w_{ch} , afectados por diferentes deformaciones recogidas en RMS y CL.

Las predicciones del modelo PBP de encendido aplicado a LER se muestran en la Figura 3.15(a)-(b) para NWFET y FinFET. Análogamente, al caso del LER para la región subumbral, el NSFET no se estudia para la variabilidad LER, ya que, el impacto en σI_{on} es despreciable para esta arquitectura.

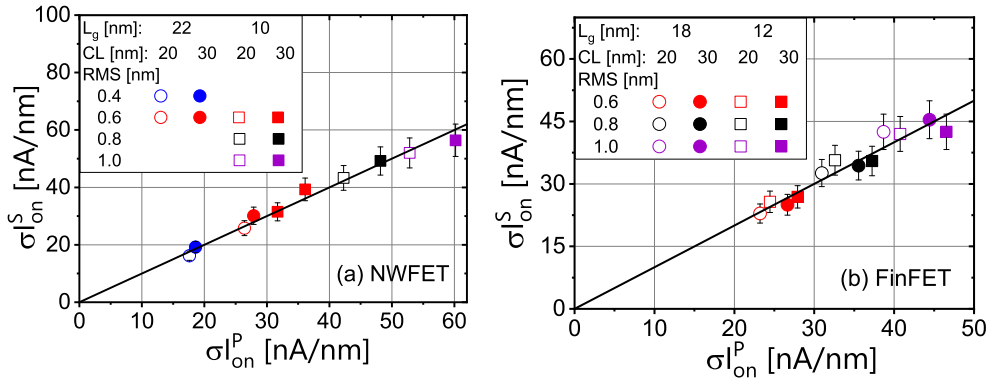


Figura 3.15: σI_{on} simulada en VENDES (σI_{on}^S) frente a la σI_{on} predicha por el modelo PBP para LER (σI_{on}^P), para dispositivos NWFET y FinFET con diferentes L_g y diferentes parámetros de deformación RMS y CL. La línea sólida negra es la bisectriz del primer cuadrante e indica correlación perfecta entre los valores predichos con el modelo y simulados. Las barras de error corresponden al 10 % de los valores simulados, margen que consideramos aceptable.

En la Tabla 3.5 mostramos una comparación entre el σI_{on} simulado y el predicho debido a la variabilidad LER, junto con los errores relativos (σ_r) entre ambos para cada conjunto de parámetros (L_g , P_{sc} , RMS, CL y θ_i^{ler}). En ella podemos observar que todas las predicciones tienen σ_r inferiores al 10 % del valor esperado para los 22 dispositivos estudiados, validando la precisión de las predicciones realizadas con el modelo PBP aplicado a LER en la región de encendido.

Tabla 3.5: σI_{on} predicha (σI_{on}^P) y simulada (σI_{on}^S), a diferentes valores de RMS y CL para las arquitecturas NWFET y FinFET. En la tabla también listamos el desfase de corriente para LER θ_i^{ler} , el área efectiva del semiconductor ($L_g \times P_{sc}$) y el error relativo σ_r . Las simulaciones realizadas específicamente para este trabajo se denotan con *.

Arq.	θ_i^{ler} [nA nm ⁻²]	CL [nm]	$L_g \times P_g$ [nm ²]	RMS [nm]	σI_{on}^P [A m ⁻¹]	σI_{on}^S [A m ⁻¹]	σ_r [%]	Ref.
NWFET	99	20	10.0×20.3	0.6	31.7	31.5	+0.6	[46]
				0.8	42.3	43.3	-2.3	*
				1.0	52.8	52.0	+1.5	*
		30	22.0×40.2	0.4	17.6	16.2	+8.6	*
				0.6	26.4	25.8	+2.3	*
				1.0	60.2	56.4	+6.7	*
	197	20	12.0×105.0	0.6	24.5	25.7	-5.1	[20]
				0.8	32.6	35.7	-8.7	[20]
				1.0	40.8	42.0	-2.9	[20]
		30	18.0×106.0	0.6	23.2	22.9	+1.3	*
				0.8	30.9	32.6	-5.2	*
				1.0	38.7	42.5	-8.9	*
FinFET	20	12.0×105.0	0.6	27.9	26.9	+3.7	*	
			0.8	37.2	35.5	+4.8	*	
			1.0	46.5	42.5	+9.4	*	
	30	18.0×106.0	0.6	26.7	25.0	+6.8	*	
			0.8	35.5	34.4	+3.2	*	
			1.0	44.4	45.4	-2.2	*	

Coste computacional

Una vez se ha demostrado la precisión del modelo PBP aplicado a diferentes arquitecturas, figuras de mérito y fuentes de variabilidad, presentaremos otra ventaja de este modelo, la reducción del coste computacional. Para ello se comparará el tiempo que se tarda en realizar un estudio de variabilidad con metodología TCAD frente a la calibración y aplicación del modelo PBP.

Con el fin de tener significancia estadística, necesitamos al menos la simulación de 300 dispositivos deformados para cada parámetro de variabilidad (GS o RMS y CL). Para estimar el tiempo de cálculo se ha tomado el caso de un NWFET de 10 nm de longitud de puerta, impactado por MGG simulado con la metodología más costosa, QC MC, necesaria para la obtención de resultados precisos en la región de encendido. En las condiciones previamente descritas, el tiempo que tarda una de estas simulaciones con VENDES en una CPU Intel^(R)

Core^(TM) i9-10850K a 3.60 GHz con una memoria DDR4 es de unas 25 h.

La Figura 3.16 muestra una comparación del tiempo computacional necesario para la realización de un estudio QC MC (denotado como TCAD) para dispositivos con una arquitectura común (NWFET) afectados por la variabilidad MGG. Todo esto definiendo como estudio de variabilidad la realización de las simulaciones necesarias para un dispositivo específico (L_g , P_g) con 4 valores GS diferentes. También se muestra el tiempo de calibración para el modelo PBP aplicado a la región de encendido (denotado en la figura como PBP). Obsérvese que este tiempo coincide con el coste computacional de un estudio de variabilidad realizado con metodología TCAD. Una vez calibrado el modelo PBP, el tiempo de cálculo para estimar el impacto de la variabilidad en los nuevos dispositivos no aumenta. Definimos la reducción del tiempo de cálculo ($t_{reduction}$) a partir del tiempo computacional en la realización de un estudio de variabilidad con TCAD (t_{TCAD}) y el tiempo en calibrar el modelo PBP (t_{PBP}):

$$t_{reduction} = \frac{t_{TCAD} - t_{PBP}}{t_{TCAD}} \quad (3.11)$$

Como puede verse en el eje Y izquierdo de la Figura 3.16, la reducción de tiempo computacional al aplicar el modelo PBP es realmente significativa a medida que aumenta el número de estudios de variabilidad. Si tenemos en cuenta que para este trabajo se han realizado 9 estudios TCAD de variabilidad MGG sobre diferentes NWFET, la aplicación completa del modelo PBP para este número de estudios en la región de encendido, produciría un ahorro computacional de $\sim 90\%$. Por tanto, este modelo puede ayudar a predecir el impacto de la variabilidad en futuros nodos tecnológicos con un coste computacional tremendamente reducido.

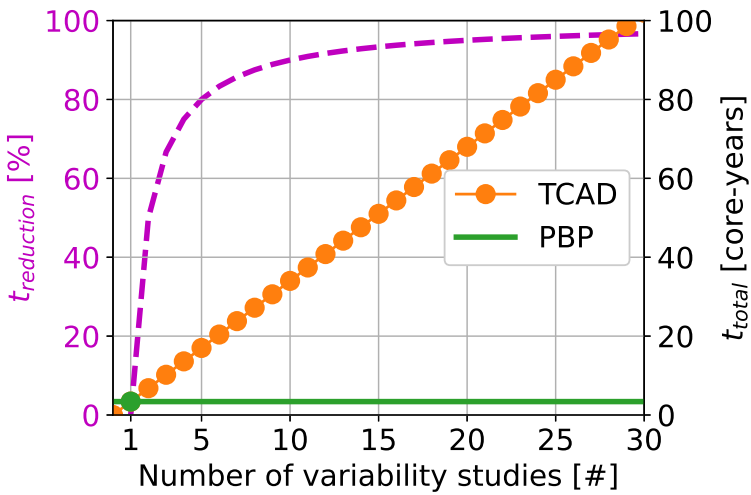


Figura 3.16: Tiempo computacional en la estimación σI_{on} debida a MGG frente al número de estudios de variabilidad. Definiendo como estudio de variabilidad la realización de las simulaciones para un dispositivo específico (L_g, P_g) a cuatro valores de GS diferentes. El tiempo de simulación de QC MC (TCAD) se compara con el necesario para calibrar el modelo PBP, para lo cual solo se necesita un estudio de variabilidad. Cada punto TCAD del gráfico corresponde a un estudio de variabilidad completo. En el eje Y derecho se muestra el tiempo total de simulación y el eje Y izquierdo muestra la reducción de tiempo computacional debido a la aplicación del modelo PBP en la región de encendido en comparación con el TCAD.

Conclusiones

En esta tesis se ha estudiado el impacto de diferentes fuentes de variabilidad sobre diferentes dispositivos basados en las arquitecturas FinFET, NWFET y NSFET que, o bien, se encuentran en el estado del arte, o son candidatos a ser los futuros nodos tecnológicos. También se han desarrollado modelos empíricos para intentar reducir el elevado coste computacional que suponen estos estudios estadísticos de variabilidad, estando basados estos modelos en la ley de Pelgrom.

En este capítulo, una vez desarrollada la metodología y la discusión de los resultados, procedemos a mostrar las principales conclusiones de este trabajo. Finalmente, también se mostrarán líneas de trabajo futuro fundamentadas en los resultados obtenidos en esta tesis.

Estudio de variabilidad para las tres arquitecturas

En la comparación del impacto de MGG, LER, GER y RDD sobre dispositivos similares (en términos de perfil de dopado, L_g y L) diseñados para las tres arquitecturas estudiadas (FinFET, NWFET y NSFET) podemos mostrar las siguientes conclusiones:

- MGG y LER son las dos fuentes de variabilidad que más impacto tienen de manera general sobre las características I_D-V_G de las tres arquitecturas. Aun así, el impacto de estas puede mitigarse en gran medida siempre que los procesos de fabricación sean capaces de reducir los valores de GS (deposiciones metálicas que reduzcan el tamaño de los granos) y RMS (mejora de procesos litográficos).
- Valorando únicamente la variabilidad inducida por MGG, el NWFET es la arquitectura menos resistente, tanto en la región subumbral como en la región de encendido, siendo las respuestas del FinFET y el NSFET prácticamente idénticas entre sí. La sensibilidad

al impacto de MGG del NWFET será en el caso de σV_{Th} y σI_{on} al menos un 50 % y 20 % mayor que en el resto de arquitecturas.

- Al estudiar el LER se observa que, aunque como comentamos anteriormente es una de las dos fuentes de variabilidad que más impacto tiene, su efecto sobre el NSFET es prácticamente imperceptible en ambas regiones de funcionamiento, ya que, $\sigma V_{Th} \leq 1$ mV y $\sigma I_{on} \leq 3$ A m⁻¹. Esto se debe al proceso de fabricación (crecimiento epitaxial) que provoca que la deformación en el NSFET se produzca en la dimensión no crítica del canal de conducción ($w_{ch} = 50$ nm), al contrario que en el caso del FinFET ($w_{ch} = 5$ nm) y NWFET ($w_{ch} = 7.5$ nm). Por ello, para un RMS= 0.6 nm, σV_{Th} es 12 y 19 veces menor para el NSFET que para FinFET y NWFET, respectivamente.
- En lo respectivo a la fuente de variabilidad GER, los resultados muestran una respuesta similar en los dispositivos que tienen puerta alrededor de todo el canal (NWFET y NSFET), con un impacto ligeramente mayor al del FinFET por ser este un dispositivo tri-puerta y no tener todo el canal rodeado por esta. Aun así, excepto en el caso del NSFET por lo explicado previamente con LER, será la fuente de variabilidad que menos impacto tiene en las características I_D - V_G subumbrales.
- El impacto de RDD es similar para las tres arquitecturas ($\sigma V_{Th} \sim 8$ mV), lo cual se debe a la aplicación del mismo perfil de dopado en la modelización. Al tener el mismo perfil de dopado gaussiano a lo largo del canal, el impacto de la implantación aleatoria de dopantes es similar. Aun así, podemos concluir que esta fuente de variabilidad es más significativa cuando los dopantes se encuentran en el canal de conducción o cerca de la frontera entre la puerta y las otras dos regiones, fuente y drenador.

Teniendo en cuenta únicamente la resistencia de cada arquitectura al impacto de las diferentes fuentes de variabilidad estudiadas, el NSFET sería la arquitectura más adecuada para sustituir a los FinFET por su elevada resistencia al LER.

Desarrollo de modelos empíricos para la reducción de coste computacional

Con el objetivo de reducir el elevado coste computacional de la realización de estudios estadísticos de variabilidad con metodología TCAD, se han desarrollado en esta tesis modelos

empíricos basados en la ley de Pelgrom (PBP) para la predicción del impacto de dos de las fuentes de variabilidad que mayor efecto tienen en el funcionamiento de las tecnologías estudiadas en este trabajo (MGG y LER). Una vez presentados los diferentes modelos, aplicados a la región subumbral y de encendido de los transistores, mostramos las conclusiones separadas por región de estudio.

Modelo PBP aplicado sobre la región subumbral

- Los errores de predicción del modelo PBP correspondientes a la región subumbral son menores o iguales a un 10 % en comparación con los resultados de los estudios TCAD, siempre que los dispositivos se encuentren dentro de los límites de validez del modelo. Esto demuestra la adaptabilidad del modelo PBP subumbral para una gran variedad de arquitecturas, dimensiones, secciones transversales y parámetros de variabilidad (GS para MGG, RMS y CL para LER). Los límites de validez del modelo vienen determinados por la saturación de σV_{Th} , producida cuando hay pocos granos en la puerta.
- Para el modelo PBP aplicado a la región subumbral y a la fuente de variabilidad MGG, obtenemos que un 91 % de los errores relativos de las predicciones son iguales o menores que el 8 %, mientras que para el caso del LER, el 94 % de los errores relativos son menores del 6 %.

Modelo PBP aplicado sobre la región de encendido

- El modelo PBP aplicado sobre la región de encendido ha sido testado para las tres arquitecturas, proporcionando errores menores del 8 % en el 92 % de los casos estudiados, combinando resultados debidos a MGG y LER. Al igual que en el caso de la región subumbral, las únicas limitaciones del modelo se producen en el caso de la variabilidad MGG cuando hay escasez de granos en la puerta metálica.
- Todos los errores relativos de las predicciones de σI_{on} debidos a LER están por debajo del 10 %.
- Al evaluar la reducción del coste computacional, se observa un beneficio claro en la aplicación del modelo a la región de encendido, ya que la realización de 9 estudios de

variabilidad con este método produciría un beneficio del $\sim 90\%$ respecto al coste de aplicar metodología TCAD.

Con lo previamente comentado se demuestra que el modelo PBP aplicado a las regiones subumbral y de encendido proporciona una estimación simple, rápida y precisa del impacto de dos de las fuentes que más contribuyen a la variabilidad de transistores con diferentes arquitecturas y dimensiones.

Trabajo futuro

Después de comentar el trabajo realizado y los principales retos que este aspira a resolver, mostraremos diferentes líneas de investigación derivadas de los resultados de esta tesis.

El silicio ha sido el material que fundamenta los dispositivos estudiados en este trabajo, debido a que es el material que domina la industria de semiconductores actual. Esto se debe principalmente a su abundancia en la naturaleza, que conjuntamente con sus propiedades eléctricas ha posibilitado su enorme desarrollo en la electrónica. Aun así, es interesante estudiar transistores basados en nuevos materiales semiconductores, como por ejemplo, aquellos basados en elementos III-V de la tabla periódica. El estudio las propiedades de los dispositivos FET basados en semiconductores III-V puede ser interesante para superar actuales limitaciones del silicio o para mejorar su funcionamiento en aplicaciones electrónicas específicas.

Otra futura línea de trabajo es la implementación de nuevas fuentes de variabilidad en nuestro simulador VENDES y su posterior estudio. Una de esas fuentes de variabilidad podría ser, las variaciones en la anchura del óxido (*Oxide Thickness Variations* (OTV)), que impacta sobre la interfaz entre óxido y semiconductor, produciendo fluctuaciones en las características electrostáticas del canal de conducción.

Una vez mostrado el poder predictivo del modelo PBP aplicado sobre MGG y LER, sería interesante realizar un análisis exhaustivo de otras fuentes de variabilidad para crear nuevas versiones del modelo PBP. También sería interesante seguir profundizando en las dependencias del modelo, por ejemplo, determinar que tipo de relación tienen las desviaciones estándar de las figuras de mérito respecto a V_D para el caso del LER o estudiar el modelo para diferentes metales en la puerta para MGG.

Para finalizar, ya que uno de los objetivos era la reducción del coste computacional, también se puede prolongar este trabajo con el desarrollo de nuevas técnicas para estudiar la variabilidad en estos dispositivos. En esta línea se han dado los primeros pasos, calibrando redes neuronales

para predecir el impacto de la variabilidad en dispositivos específicos afectados por MGG y LER, como se puede ver en las publicaciones complementarias de esta tesis. De todos modos, existe un gran abanico de posibilidades para explorar en relación con la aplicación de técnicas basadas en *machine learning*, abriendo un campo muy interesante de estudio, teniendo en cuenta la gran cantidad de datos accesibles de simulaciones TCAD.

Conclusions

The impact of different sources of variability on different devices based on the FinFET, NWFET and NSFET architectures, that are in the state-of-art or are candidates to the future technological nodes, has been studied in this thesis. Also, an empirical model based on the Pelgrom's law has been developed in order to reduce the high computational cost associated with the statistical variability studies performed with TCAD.

Variability study for the three architectures

In the comparison between the impact of MGG, LER, GER and RDD induced variabilities on similar devices (in terms of doping profile, L_g and L) designed for the three architectures studied in this work (FinFET, NWFET and NSFET) we obtained the following conclusions:

- MGG and LER are two of the variability sources that generally have the biggest impact on the I_D - V_G characteristics for the three architectures. Even so, the impact of these sources can be mitigated if the manufacturing processes were able to lower the values of GS (metallic depositions that reduce the grain size) and RMS (improvement of the lithographic processes).
- Focusing only in the variability induced by MGG, the NWFET is the most affected architecture in the subthreshold region and also in the on region, being the impact on the FinFET and the NSFET practically identical. The sensibility due to MGG for NWFET is for σV_{Th} and σI_{on} at least 50 % and 20 % higher than for the other architectures.
- When studying the LER, it is observed that, although as we mentioned earlier it is one of the two sources of variability with the greatest impact, its effect on the NSFET is practically imperceptible in both operating regions, given that $\sigma V_{Th} \leq 1$ mV and

$\sigma I_{on} \leq 3 \text{ A m}^{-1}$. This is due to the manufacturing process (epitaxial growth) that causes the deformation in the NSFET to occur in the non-critical dimension of the conduction channel ($w_{ch} = 50 \text{ nm}$), unlike the case of the FinFET ($w_{ch} = 5 \text{ nm}$) and NWFET ($w_{ch} = 7.5 \text{ nm}$). Therefore, for an RMS= 0.6 nm, σV_{Th} is 12 and 19 times smaller for the NSFET compared to the FinFET and NWFET, respectively.

- Regarding the GER induced variability, the results exhibit a similar response in devices with gates enclosing the entire channel (NWFET and NSFET). These devices show a slightly greater impact than the FinFET, which is a tri-gate device and does not enclose the entire channel with its gate. Nevertheless, except for the case of the NSFET due to the earlier explanation on LER, GER will be the least influential source of variability on the subthreshold I_D - V_G characteristics.
- The similar impact of RDD on the three architectures ($\sigma V_{Th} \sim 8 \text{ mV}$) results from the consistent application of the same doping profile in the modelling process. Since the gaussian doping profile is applied equally along the channel, the influence of the random dopant implantation remains similar for all the studied devices. However, it can still be deduced that this source of variability becomes more pronounced when the dopants are located within the conduction channel itself or in proximity to the boundaries between the gate and the other two regions, the source and drain.

Taking into consideration solely the resilience of each architecture against the influence of the studied variability sources, the NSFET emerges as the most promising candidate to replace FinFET largely due to its remarkable resistance to LER.

Development of empirical model for the reduction of the computational cost

To address the considerable computational costs linked to performing statistical variability studies using TCAD methodology, this thesis has formulated empirical models grounded in Pelgrom's Law (PBP). These models are developed to forecast the effects of two prominent sources of variability that significantly influence the operation of the devices investigated in this study (MGG and LER). After introducing the models, which have been applied to both the subthreshold and on regions of transistor's performance, we proceed to present the conclusions organized by their respective study regions.

PBP model applied to the subthreshold region

- The prediction errors of the PBP model in the subthreshold region are lower or equal to 10 % when compared to the results from TCAD studies, provided that the devices adhere to the model's validity constraints. This highlights the adaptability of the subthreshold PBP model across a wide range of architectures, dimensions, cross-sections, and variability parameters (GS for MGG, and RMS and CL for LER). The model's validity limits are determined by the saturation of σV_{Th} , which occurs when there are few grains on the gate.
- For the PBP model applied to the subthreshold region and the MGG variability source, we find that 91 % of the relative prediction errors are equal to or less than 8 %, while in the case of LER all relative errors are inside the 10 % margin, being 94 % of the relative errors below the 6 %.

PBP model applied to the on region

- The PBP model applied to the on region has been tested for all three architectures, yielding to errors lower than 8 % in 92 % of the studied cases by combining results from MGG and LER. Similar to the subthreshold region, the only limitations of the model arise in the case of MGG variability when there is a shortage of grains in the gate.
- All the relative prediction errors of σI_{on} due to LER are below the 10 %.
- Upon analysing the decrease in computational costs, a clear advantage emerges from employing the model in the on region. For example, applying the PBP model to predict 9 variability studies would yield a significant 90 % reduction in expenses compared to the costs associated with TCAD methodology.

As previously discussed, the application of the PBP model to both the subthreshold and on regions demonstrates its capability to provide a simple, fast, and accurate estimation of the impact induced from two of the most influential sources contributing to the variability of transistors with varying architectures and dimensions.

Future work

After reviewing the completed work and the central challenges it addresses, we will outline diverse areas of research that have been inspired by the conclusions of this thesis.

Silicon has been the foundation of the devices studied in this research due to its dominance in the semiconductor industry. This is largely because of its natural abundance and electrical properties, which have greatly contributed to its advancement in electronics. Even so, it's interesting to explore transistors using new materials like Group III-V elements. Exploring the properties of FET devices using these materials could help to overcome the limitations associated with silicon and enhance performance in targeted electronic applications.

Another future line of work involves incorporating new sources of variability into our VENDES simulator and subsequently studying their effects. One potential source of variability could be the oxide thickness variations (OTV), which influence the interface between the oxide and the semiconductor, leading to fluctuations in the electrostatic characteristics of the conduction channel.

Once the predictive capabilities of the PBP model applied to MGG and LER have been demonstrated, it would be valuable to perform a comprehensive analysis of additional variability sources to develop enhanced versions of the PBP model. It would also be interesting to further explore the model's dependencies. For example, understanding the relationship between standard deviations of different figures of merit and V_D for LER, or examining how the model behaves with different gate metals in the case of MGG, could provide valuable insights.

Finally, considering that one of the goals was to reduce computational costs, this work can also be extended by developing new techniques for studying variability in these devices. Taking steps in this direction, initial works have been done by calibrating neural networks to forecast the influence of variability on particular devices affected by MGG and LER, as evidenced in the supplementary publications of this thesis. However, there is a wide range of possibilities to explore regarding the application of machine learning techniques. This opens up an incredibly interesting field of study, especially considering the amount of data available from TCAD simulations.

Bibliografía

- [1] J. Bardeen and W. H. Brattain, “The transistor, a semi-conductor triode,” *Physical Review*, vol. 74, no. 2, p. 230, 1948.
- [2] G. Moore, “Moore’s law,” *Electronics Magazine*, vol. 38, no. 8, p. 114, 1965.
- [3] IRDS. (2021) International Roadmap for Devices and Systems: More Moore.
- [4] V. S. Basker, T. Standaert, H. Kawasaki, C.-C. Yeh, K. Maitra, T. Yamashita, J. Faltermeier, H. Adhikari, H. Jagannathan, J. Wang, H. Sunamura, S. Kanakasabapathy, S. Schmitz, J. Cummings, A. Inada, C. H. Lin, P. Kulkarni, Y. Zhu, J. Kuss, T. Yamamoto, A. Kumar, J. Wahl, A. Yagishita, L. F. Edge, R. H. Kim, E. McLellan, S. J. Holmes, R. C. Johnson, T. Levin, J. Demarest, M. Hane, M. Takayanagi, M. Colburn, V. K. Paruchuri, R. J. Miller, H. Bu, B. Doris, D. McHerron, E. Leobandung, and J. O’Neill, “A 0.063 μm^2 FinFET SRAM cell demonstration with conventional lithography using a novel integration scheme with aggressively scaled fin and gate pitch,” *2010 Symposium on VLSI Technology*, pp. 19–20, 2010.
- [5] S. Bangsaruntip, K. Balakrishnan, S.-L. Cheng, J. Chang, M. Brink, I. Lauer, R. L. Bruce, S. U. Engelmann, A. Pyzyna, G. M. Cohen, L. M. Gignac, C. M. Breslin, J. S. Newbury, D. P. Klaus, A. Majumdar, J. W. Sleight, and M. A. Guillorn, “Density scaling with gate-all-around silicon nanowire MOSFETs for the 10 nm node and beyond,” in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 20.2.1–20.2.4.
- [6] N. Loubet, T. Hook, P. Montanini, C.-W. Yeung, S. Kanakasabapathy, M. Guillom, T. Yamashita, J. Zhang, X. Miao, J. Wang, A. Young, R. Chao, M. Kang, Z. Liu, S. Fan, B. Hamieh, S. Sieg, Y. Mignot, W. Xu, S.-C. Seo, J. Yoo, S. Mochizuki, M. Sankaranandian, O. Kwon, A. Carr, A. Greene, Y. Park, J. Frougier, R. Galatage, R. Bao, J. Shearer,

- R. Conti, H. Song, D. Lee, D. Kong, Y. Xu, A. Arceo, Z. Bi, P. Xu, R. Muthinti, J. Li, R. Wong, D. Brown, P. Oldiges, R. Robison, J. Arnold, N. Felix, S. Skordas, J. Gaudiello, T. Standaert, H. Jagannathan, D. Corliss, M.-H. Na, A. Knorr, T. Wu, D. Gupta, S. Lian, R. Divakaruni, T. Gow, C. Labelle, S. Lee, V. Paruchuri, H. Bu, and M. Khare, “Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET,” in *2017 Symposium on VLSI Technology*, 2017, pp. T230–T231.
- [7] C. Geuzaine and J.-F. Remacle, “Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities,” *International Journal for Numerical Methods in Engineering*, vol. 79, no. 11, pp. 1309–1331, May 2009. [Online]. Available: <https://doi.org/10.1002/nme.2579>
- [8] N. Seoane, D. Nagy, G. Indalecio, G. Espiñeira, K. Kalna, and A. García-Loureiro, “A Multi-Method Simulation Toolbox to Study Performance and Variability of Nanowire FETs,” *Materials*, vol. 12, no. 15, 2019. [Online]. Available: <https://www.mdpi.com/1996-1944/12/15/2391>
- [9] H. Dadgour, K. Endo, V. De, and K. Banerjee, “Modeling and analysis of grain-orientation effects in emerging metal-gate devices and implications for SRAM reliability,” in *2008 IEEE International Electron Devices Meeting*, 2008, pp. 1–4.
- [10] A. Asenov, S. Kaya, and A. Brown, “Intrinsic parameter fluctuations in decananometer MOSFETs introduced by gate line edge roughness,” *IEEE Transactions on Electron Devices*, vol. 50, no. 5, pp. 1254–1260, 2003.
- [11] G. Indalecio, A. J. García-Loureiro, N. Seoane Iglesias, and K. Kalna, “Study of Metal-Gate Work-Function Variation Using Voronoi Cells: Comparison of Rayleigh and Gamma Distributions,” *IEEE Transactions on Electron Devices*, vol. 63, no. 6, pp. 2625–2628, 2016.
- [12] M. Lemme, J. Efavi, T. Mollenhauer, M. Schmidt, H. Gottlob, T. Wahlbrink, and H. Kurz, “Nanoscale TiN metal gate technology for CMOS integration,” *Microelectronic Engineering*, vol. 83, no. 4, pp. 1551–1554, 2006, micro- and Nano-Engineering MNE 2005. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167931706001286>

- [13] S. A. Vitale, J. Kedzierski, P. Healey, P. W. Wyatt, and C. L. Keast, “Work-Function-Tuned TiN Metal Gate FDSOI Transistors for Subthreshold Operation,” *IEEE Transactions on Electron Devices*, vol. 58, no. 2, pp. 419–426, 2011.
- [14] G. Indalecio, M. Aldegunde, N. Seoane, K. Kalna, and A. J. García-Loureiro, “Statistical study of the influence of LER and MGG in SOI MOSFET,” *Semiconductor Science and Technology*, vol. 29, no. 4, p. 045005, feb 2014. [Online]. Available: <https://dx.doi.org/10.1088/0268-1242/29/4/045005>
- [15] E. Giannatou, G. Papavieros, V. Constantoudis, H. Papageorgiou, and E. Gogolides, “Deep learning denoising of SEM images towards noise-reduced LER measurements,” *Microelectronic Engineering*, 2019.
- [16] G. F. Lorusso, O. Inoue, T. Ohashi, E. A. Sanchez, V. Constantoudis, and S. Koshihara, “Line width roughness accuracy analysis during pattern transfer in self-aligned quadruple patterning process,” in *Metrology, Inspection, and Process Control for Microlithography XXX*, M. I. Sanchez, Ed., vol. 9778, International Society for Optics and Photonics. SPIE, 2016, p. 97780V. [Online]. Available: <https://doi.org/10.1117/12.2218863>
- [17] G. Espiñeira, D. Nagy, G. Indalecio, A. J. García-Loureiro, K. Kalna, and N. Seoane, “Impact of Gate Edge Roughness Variability on FinFET and Gate-All-Around Nanowire FET,” *IEEE Electron Device Letters*, vol. 40, no. 4, pp. 510–513, 2019.
- [18] D. Nagy, G. Espiñeira, G. Indalecio, A. J. García-Loureiro, K. Kalna, and N. Seoane, “Benchmarking of FinFET, Nanosheet, and Nanowire FET Architectures for Future Technology Nodes,” *IEEE Access*, vol. 8, pp. 53 196–53 202, 2020.
- [19] M. A. Elmessary, D. Nagy, M. Aldegunde, N. Seoane, G. Indalecio, J. Lindberg, W. Dettmer, D. Perić, A. J. García-Loureiro, and K. Kalna, “Scaling/LER study of Si GAA nanowire FET using 3D finite element Monte Carlo simulations,” *Solid-State Electronics*, vol. 128, pp. 17–24, 2017, extended papers selected from EUROSOI-ULIS 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0038110116301836>
- [20] N. Seoane, J. G. Fernandez, K. Kalna, E. Comesaña, and A. García-Loureiro, “Simulations of Statistical Variability in n-Type FinFET, Nanowire, and Nanosheet FETs,” *IEEE Electron Device Letters*, vol. 42, no. 10, pp. 1416–1419, 2021.

- [21] C. Millar, D. Reid, G. Roy, S. Roy, and A. Asenov, “Accurate Statistical Description of Random Dopant-Induced Threshold Voltage Variability,” *IEEE Electron Device Letters*, vol. 29, no. 8, pp. 946–948, 2008.
- [22] A. Asenov, J. Watling, A. Brown, and D. Ferry, “The Use of Quantum Potentials for Confinement and Tunnelling in Semiconductor Devices,” *Journal of Computational Electronics*, vol. 1, no. 4, pp. 503–513, 2002. [Online]. Available: <https://doi.org/10.1023/a:1022905508032>
- [23] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. Springer Vienna, 1984. [Online]. Available: <https://doi.org/10.1007/978-3-7091-8752-4>
- [24] F. Assad, K. Banoo, and M. Lundstrom, “The drift-diffusion equation revisited,” *Solid-State Electronics*, vol. 42, no. 3, pp. 283–295, 1998. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0038110197002633>
- [25] C. Jacoboni and P. Lugli, *The Monte Carlo method for semiconductor device simulation*. Springer Science & Business Media, 2012.
- [26] S. Stanzione, D. Puntin, and G. Iannaccone, “CMOS Silicon Physical Unclonable Functions Based on Intrinsic Process Variability,” *IEEE Journal of Solid-State Circuits*, vol. 46, no. 6, pp. 1456–1463, 2011.
- [27] A. Ortiz-Conde, F. J. García-Sánchez, J. Muci, A. Terán Barrios, J. J. Liou, and C.-S. Ho, “Revisiting MOSFET threshold voltage extraction methods,” *Microelectronics Reliability*, vol. 53, no. 1, pp. 90–104, 2013, reliability of Micro-Interconnects in 3D IC Packages. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0026271412004684>
- [28] K. Patel, T.-J. K. Liu, and C. J. Spanos, “Gate Line Edge Roughness Model for Estimation of FinFET Performance Variability,” *IEEE Transactions on Electron Devices*, vol. 56, no. 12, pp. 3055–3063, 2009.
- [29] X. Wang, A. R. Brown, B. Cheng, and A. Asenov, “Statistical variability and reliability in nanoscale FinFETs,” in *2011 International Electron Devices Meeting*, 2011, pp. 5.4.1–5.4.4.

- [30] D. Nagy, G. Indalecio, A. J. García-Loureiro, G. Espiñeira, M. A. Elmessary, K. Kalna, and N. Seoane, “Drift-Diffusion Versus Monte Carlo Simulated ON-Current Variability in Nanowire FETs,” *IEEE Access*, vol. 7, pp. 12 790–12 797, 2019.
- [31] S. Yu, S. M. Won, H. W. Baac, D. Son, and C. Shin, “Quantitative Evaluation of Line-Edge Roughness in Various FinFET Structures: Bayesian Neural Network With Automatic Model Selection,” *IEEE Access*, vol. 10, pp. 26 340–26 346, 2022.
- [32] C. Akbar, Y. Li, and W. L. Sung, “Machine Learning Aided Device Simulation of Work Function Fluctuation for Multichannel Gate-All-Around Silicon Nanosheet MOSFETs,” *IEEE Transactions on Electron Devices*, vol. 68, no. 11, pp. 5490–5497, 2021.
- [33] C. Akbar, Y. Li, and W.-L. Sung, “Deep Learning Algorithms for the Work Function Fluctuation of Random Nanosized Metal Grains on Gate-All-Around Silicon Nanowire MOSFETs,” *IEEE Access*, vol. 9, pp. 73 467–73 481, 2021.
- [34] J. Lee, T. Park, H. Ahn, J. Kwak, T. Moon, and C. Shin, “Prediction Model for Random Variation in FinFET Induced by Line-Edge-Roughness (LER),” *Electronics*, vol. 10, no. 4, 2021. [Online]. Available: <https://www.mdpi.com/2079-9292/10/4/455>
- [35] H. Carrillo-Nuñez, N. Dimitrova, A. Asenov, and V. Georgiev, “Machine Learning Approach for Predicting the Effect of Statistical Variability in Si Junctionless Nanowire Transistors,” *IEEE Electron Device Letters*, vol. 40, no. 9, pp. 1366–1369, 2019.
- [36] G. Indalecio, N. Seoane, K. Kalna, and A. J. García-Loureiro, “Fluctuation Sensitivity Map: A Novel Technique to Characterise and Predict Device Behaviour Under Metal Grain Work-Function Variability Effects,” *IEEE Transactions on Electron Devices*, vol. 64, no. 4, pp. 1695–1701, 2017.
- [37] A. Wettstein, O. Penzin, E. Lyumkis, and W. Fichtner, “Random dopant fluctuation modelling with the impedance field method,” in *International Conference on Simulation of Semiconductor Processes and Devices, 2003. SISPAD 2003.*, 2003, pp. 91–94.
- [38] W.-L. Sung and Y. Li, “Statistical Prediction of Nanosized-Metal-Grain-Induced Threshold-Voltage Variability for 3D Vertically Stacked Silicon Gate-All-Around Nanowire n-MOSFETs,” *Journal of Electronic Materials*, vol. 49, no. 11, pp. 6865–6871, Aug. 2020.

- [39] M. Pelgrom, A. Duinmaijer, and A. Welbers, “Matching properties of MOS transistors,” *IEEE Journal of Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, 1989.
- [40] A. Sudarsanan, S. Venkateswarlu, and K. Nayak, “Impact of Fin Line Edge Roughness and Metal Gate Granularity on Variability of 10-nm Node SOI n-FinFET,” *IEEE Transactions on Electron Devices*, vol. 66, no. 11, pp. 4646–4652, 2019.
- [41] K. Nayak, S. Agarwal, M. Bajaj, P. J. Oldiges, K. V. R. M. Murali, and V. R. Rao, “Metal-Gate Granularity-Induced Threshold Voltage Variability and Mismatch in Si Gate-All-Around Nanowire n-MOSFETs,” *IEEE Transactions on Electron Devices*, vol. 61, no. 11, pp. 3892–3895, 2014.
- [42] K. Takeuchi, T. Fukai, T. Tsunomura, A. T. Putra, A. Nishida, S. Kamohara, and T. Hiramoto, “Understanding Random Threshold Voltage Fluctuation by Comparing Multiple Fabs and Technologies,” in *2007 IEEE International Electron Devices Meeting, 2007*, pp. 467–470.
- [43] P. Stolk, F. Widdershoven, and D. Klaassen, “Modeling statistical dopant fluctuations in MOS transistors,” *IEEE Transactions on Electron Devices*, vol. 45, no. 9, pp. 1960–1971, 1998.
- [44] N. Seoane, G. Indalecio, M. Aldegunde, D. Nagy, M. A. Elmessary, A. J. García-Loureiro, and K. Kalna, “Comparison of Fin-Edge Roughness and Metal Grain Work Function Variability in InGaAs and Si FinFETs,” *IEEE Transactions on Electron Devices*, vol. 63, no. 3, pp. 1209–1216, 2016.
- [45] W.-F. Lü and L. Dai, “Impact of work-function variation on analog figures-of-merits for high-k/metal-gate junctionless FinFET and gate-all-around nanowire MOSFET,” *Microelectronics Journal*, vol. 84, pp. 54–58, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S002626921830497X>
- [46] D. Nagy, G. Indalecio, A. J. García-Loureiro, M. A. Elmessary, K. Kalna, and N. Seoane, “FinFET Versus Gate-All-Around Nanowire FET: Performance, Scaling, and Variability,” *IEEE Journal of the Electron Devices Society*, vol. 6, pp. 332–340, 2018.
- [47] W.-L. Sung, Y.-S. Yang, and Y. Li, “Work-Function Fluctuation of Gate-All-Around Silicon Nanowire n-MOSFETs: A Unified Comparison Between Cuboid and Voronoi Methods,” *IEEE Journal of the Electron Devices Society*, vol. 9, pp. 151–159, 2021.

- [48] G. Leung and C. O. Chui, “Variability of Inversion-Mode and Junctionless FinFETs due to Line Edge Roughness,” *IEEE Electron Device Letters*, vol. 32, no. 11, pp. 1489–1491, 2011.
- [49] D. Nagy, G. Indalecio, A. J. García-Loureiro, M. A. Elmessary, K. Kalna, and N. Seoane, “Metal Grain Granularity Study on a Gate-All-Around Nanowire FET,” *IEEE Transactions on Electron Devices*, vol. 64, no. 12, pp. 5263–5269, 2017.
- [50] J. G. Fernandez, N. Seoane, E. Comesaña, K. Kalna, and A. García-Loureiro, “Impact of metal grain granularity on three gate-all-around advanced architectures,” in *2021 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 2021, pp. 201–205.

APÉNDICE A

TRABAJOS PUBLICADOS

A.1. Artículos en esta tesis

A.1.1. Simulations of statistical variability in n-type FinFET, nanowire, and nanosheet FETs



Simulations of Statistical Variability in n-Type FinFET, Nanowire, and Nanosheet FETs

Author: Natalia Seoane
Publication: IEEE Electron Device Letters
Publisher: IEEE
Date: October 2021

Copyright © 2021, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

- 1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
- 2) In the case of illustrations or tabular material, we require that the copyright line © [year of original publication] IEEE appear prominently with each reprinted figure and/or table.
- 3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

- 1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
- 2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
- 3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK

CLOSE WINDOW

- Publication quality indicators:
 - Impact factor (JCR 2021): 4.816
 - Quartile: Q1
 - Category: Engineering, electrical & electronic
 - Rank: 66/344
- PhD candidate contribution:
 - Simulation of variability affected devices.
 - Data processing.
 - Co-writing of the manuscript.

Simulations of Statistical Variability in n -Type FinFET, Nanowire, and Nanosheet FETs

Natalia Seoane¹, Julian G. Fernandez¹, Karol Kalna¹, *Senior Member, IEEE*,
Enrique Comesaña, and Antonio García-Loureiro¹, *Member, IEEE*

Abstract—Four sources of variability, metal grain granularity (MGG), line-edge roughness (LER), gate-edge roughness (GER), and random discrete dopants (RDD), affecting the performance of state-of-the-art FinFET, nanosheet (NS), and nanowire (NW) FETs, are analysed via our in-house 3D finite-element drift-diffusion/Monte Carlo simulator that includes 2D Schrödinger equation quantum corrections. The MGG and LER are the sources of variability that influence device performance of the three multi-gate architectures the most. The FinFET and the NS FET are similarly affected by the MGG variations with threshold voltage and on-current standard deviations significantly lower (at least 20%) than those of the NW FET. The LER variability has a negligible influence in the NS FET performance with σV_T values around 12 and 42 times lower than those of the FinFET and the NW FET. The three architectures are equally affected by the RDD ($\sigma V_T = 8$ mV) and minimally influenced by the GER ($\sigma V_T \approx 4$ mV). The variability of NS FETs makes them strong candidates to replace FinFETs.

Index Terms—Drift-diffusion, FinFET, nanosheet, nanowire, monte carlo, Schrödinger quantum correction, variability.

I. INTRODUCTION

FINFET architecture has been dominant for digital applications during the last years [1]. However, a changeover will be needed for future CMOS technology generations in order to maintain a good electrostatic control [2]. Gate-all-around (GAA) nanowire (NW) FETs are currently considered as one of the strongest contenders to replace the FinFETs because of their superior gate control [3] but, their adoption implies a substantial change in the fabrication processes [4]. The GAA nanosheet (NS) FETs have been proposed as an

intermediate step between both architectures due to a slightly better performance than the FinFETs [5] while reusing, with minimal changes [6], its fabrication process.

The continuous scaling of device dimensions brought an increase in process induced variability. The main variability sources affecting the reliability of multi-gate transistors are: metal-gate work-function granularity (MGG), line edge roughness (LER), gate edge roughness (GER), random discrete dopants (RDD), oxide thickness variation (OTV), and interface trap charges (ITC) [7]. The minimisation of this device variability is essential to reduce further a supply voltage and thus a power dissipation [8]. The impact of variability in multi-gate transistor architectures [9]–[12] has been intensively analysed but, to our best knowledge, no thorough studies exist that would compare a variability of nanoscale FinFET, NS and NW FET architectures. Therefore, in this work, we investigate the effect that four different sources of variability (MGG, LER, GER and RDD) have on the performance of realistic, Si-based FinFET, NW FET, and NS FET. VENDES [13], an in-house-built 3D finite-element physically-based simulation toolbox, is used to analyse the variability effect on the devices performance. VENDES combines quantum-corrected drift-diffusion and ensemble Monte Carlo transport models via the 2D Schrödinger equation to precisely model both sub-threshold and on-region device characteristics.

II. 3D MODELING OF VARIABILITY

Figs. 1(a)–(c) present 3D schemes of the three device architectures and show their physical dimensions. The 12 nm NS FET structure is designed after the experimental device reported in [6] and it was previously validated in [5]. Because doping profiles in the source/drain and channel regions are typically unknown, the profiles are determined from the experimental I-V characteristics in the sub-threshold by a reverse engineering process [5], [14] (see resulting values in Fig. 1). The 12 nm FinFET could be seen as a tri-gate version of the GAA NS FET that has a similar doping profile. The 12 nm gate length NW FET was scaled-down from a 22 nm gate length experimental transistor [15] previously studied in [14].

VENDES employs the finite-element (FE) method and combines 2D solutions of the Schrödinger (SCH) equation across the channel with the 3D drift-diffusion (DD) transport model and the 3D ensemble Monte Carlo (MC) technique in order to simulate the FETs in the sub-threshold and the on-regions, respectively [13]. The FE method allows to generate tetrahedral meshes that perfectly depict the real geometries as seen in Fig. 1. The 2D SCH equation quantum corrections incorpo-

Manuscript received August 2, 2021; revised August 24, 2021; accepted August 30, 2021. Date of publication September 1, 2021; date of current version September 27, 2021. This work was supported by the Spanish Ministerio de Ciencia e Innovación, Xunta de Galicia, and the Fondo Europeo de Desarrollo Regional (FEDER) funds under Grant RYC-2017-23312, Grant PID2019-104834GB-I00, and Grant ED431F 2020/008. The review of this letter was arranged by Editor U. Ganguly. (Corresponding author: Natalia Seoane.)

Natalia Seoane, Julian G. Fernandez, Enrique Comesaña, and Antonio García-Loureiro are with the CITIUS, Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain (e-mail: natalia.seoane@usc.es).

Karol Kalna is with the Nanoelectronic Devices Computational Group, Faculty of Science and Engineering, Swansea University, Swansea, Wales SA1 8EN, U.K.

Color versions of one or more figures in this letter are available at <https://doi.org/10.1109/LED.2021.3109586>.

Digital Object Identifier 10.1109/LED.2021.3109586

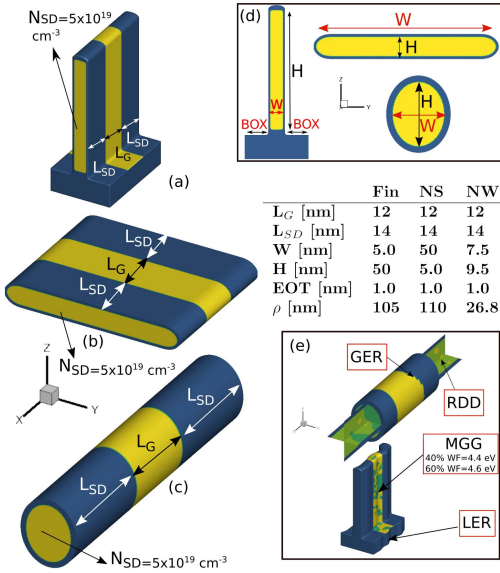


Fig. 1. 3D schemes of the FinFET (a), NS FET (b), and NW FET (c). 2D schemes of their respective cross-sections are shown in (d). The device dimensions: physical gate (L_G), source/drain lengths (L_{SD}), channel width (W) and height (H), effective oxide thickness (EOT), and semiconductor perimeter (ρ), are included in a table. The devices have a uniform p -type doping ($1.0 \times 10^{19} \text{ cm}^{-3}$) in the channel and a n -type Gaussian doping in the S/D regions (NSD). (e) Examples of devices affected by GER, RDD, MGG and LER variability.

rated into the 3D DD model permit a precise characterisation of the sub-threshold region. The 3D MC simulations consider all Si scattering mechanisms with Fermi-Dirac statistics in the screening of electron-ionised impurity scattering (self-consistently calculating Fermi level and electron temperature from electron density and average energy at the device mesh during each MC time step [16]).

This accurate modelling of electron transport in the highly doped source/drain regions assures a correct injection of carriers into the channel and reproduces correctly the experimental resistance in the source/drain [14], [17]. More details of the simulation procedures can be found in [13] whereas in-depth descriptions of the DD and MC physical models can be found in [18], [19].

To model the MGG, the gate work-function (WF) is modified using the Voronoi approach to represent the experimental shapes and values of the different grain orientations [20] (Fig. 1(e) lists the WF values and probabilities used for a TiN metal gate [21]). To introduce the LER, the edges of the device are deformed in the y -direction, which is perpendicular to the transport direction (see the reference axis in Fig. 1(d)). To generate the roughness profiles, Fourier synthesis with Gaussian auto-correlation [22] is used. The deformations are described by a root mean square (RMS) height that determines the amplitude of the roughness, and by a correlation length (CL) that accounts for the correlation in the transport direction between deformations (see the example in Fig. 1(e)). The GER follows a similar methodology to the LER, but the deformation profile is applied to the device gate (see the

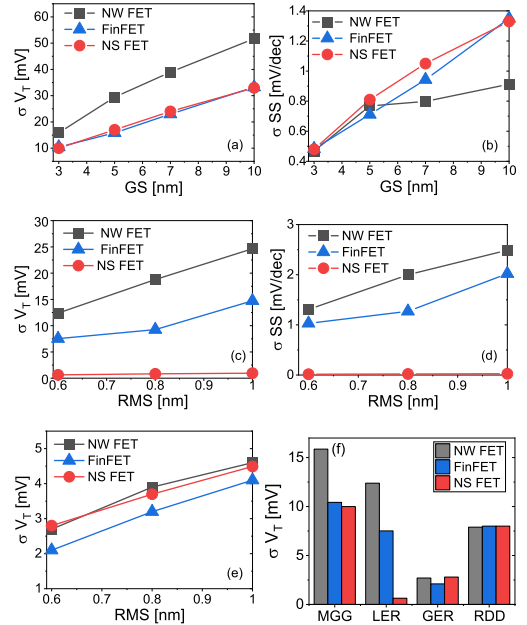


Fig. 2. (a) Threshold voltage (σV_T) and (b) sub-threshold slope (σSS) standard deviations due to the MGG vs. the grain size, (c) σV_T and (d) σSS due to the LER vs. the root mean square (RMS) for a correlation length (CL) of 20 nm, (e) σV_T due to the GER vs. the RMS value (CL = 20 nm), and (f) a comparison of σV_T vs. the source of variability for the three architectures (GS = 3 nm and RMS = 0.6 nm).

top example in Fig. 1(e)) [23]. To model the RDD in the n -type source/drain, a rejection technique is employed from the continuous doping of the ideal device (see the values in Fig. 1) [24], mapping the dopants to the tetrahedral mesh via the cloud-in-cell approach (see an example of the resulting electron concentration in Fig. 1(e)).

III. COMPARISON OF VARIABILITY

All the study is done at a high drain bias, $V_{DD} = 0.7 \text{ V}$. The threshold voltage (V_T) is calculated using the constant current method [25] set to $I_D = 2.0 \mu\text{A}/\mu\text{m}$. The on-current (I_{ON}) is the drain current at $V_G = V_{DD} + V_T$. To obtain statistical significance, ensembles of 300 devices are generated for each source of variability and for each analysed parameter.

Figs. 2(a)-(b) show the V_T and SS standard deviations (σ) due to MGG as a function of the metal-gate grain size (GS), respectively, for the three analysed architectures. The NS FET and the FinFET exhibit less V_T variability than the NW architecture. We can observe that, the smaller the gate area (G_A), the larger the expected variations. For instance, the σV_T for the NW FET ($G_A \approx 395 \text{ nm}^2$) is 59% larger for a GS = 3 nm than those of the FinFET and the NS FET ($G_A > 1000 \text{ nm}^2$). The FinFET and NS FET exhibit an almost identical response to the V_T induced MGG variability with their σV_T s linearly increasing with the GS. However, for the NW FET, σV_T begins to saturate at large GSs because of the very low number of grains present in the gate (≈ 4 grains when GS = 10 nm) and the statistics follows a bimodal distribution.

This saturation is even more noticeable in the σSS for the NW FET at GS equal or larger than 7 nm. However, the FinFET and the NS FET V_T and SS distributions are normal for all the analysed GS values.

Figs. 2(c)-(d) show, respectively, σV_T and σSS due to LER as a function of the RMS height for a CL of 20 nm. Previous SEM (Scanning Electron Microscope) studies [22], [26] showed CL values in the range of 20 – 80 nm. The influence of LER on the NS FET is negligible, with σV_T values around 12 and 19 times lower than those of the FinFET and NW FET, respectively (at RMS = 0.6 nm). LER is induced by resist patterning [27], and as previously mentioned, the LER deformation is applied in the y -direction, which corresponds with the critical dimensions for the FinFET (5.0 nm) and NW FET (7.5 nm). However, the critical dimension of the NS FET is its height (z -direction) that will not be affected by LER since is defined by epitaxial growth [6]. The width of the NS FET (50 nm) is large enough to remain practically unaffected by any LER deformation. Although the FinFET has a smaller width than that of NW FET, its larger height (50 nm) implies a larger conduction area, which reduces the potential impact of a LER deformation and explains the lower σV_T and σSS when compared to those of the NW FET. In addition, for the FinFET and NW FET, the LER induced σSS is generally larger than that observed for the MGG variability. Bear in mind that MGG, unlike LER, mainly implies a threshold voltage shift in the I-V characteristics.

Fig. 2(e) shows σV_T due to GER as a function of the RMS height for a CL of 20 nm. The influence of GER is significantly smaller than that of the other sources of variability analysed. Results show that the FinFET is less disturbed by the sub-threshold GER variability than the NS and the NW FET architectures. For example, at a RMS = 0.6 nm, the FinFET has a $\approx 20\%$ lower σV_T than those of the NS and the NW FETs (which are very similar). To understand the physical reason of this behaviour, for the FinFET and NS FETs we simulated configurations in which, keeping always the total device length fixed (40 nm), we increased/reduced L_G and therefore reduced/increased L_{SD} by 1 nm to account for the extreme deformation scenarios. We observed that a reduction/increase in L_G lead to a reduction/increase in the V_T value with respect to that of the 12 nm gate length device (δV_T). The FinFET and the NS FET presented the exact same δV_T values, which indicated a similar sensitivity of the two architectures to gate deformations. Therefore, the lower variability yielded by the FinFET can be explained due to the part of the gate located over the BOX regions (see schemes in Figs. 1(a) and (d)) which span over 10 nm. Any GER deformation falling in these zones will not affect the device behaviour, but will lower the statistical impact of the GER variability in the FinFET.

Fig. 2(f) shows the V_T variation for the four different sources of variability. For each source, the configurations that yield the lowest variability (i.e., a GS = 3 nm for MGG and a RMS = 0.6 nm for LER and GER) are selected. The three architectures have the same resilience ($\sigma V_T = 8$ mV) to the RDD variations. Note that, as indicated in Fig. 1, they all have the same S/D length and N_{SD} doping. The RDD variations in the FinFET are larger than those previously reported for a similar gate length device [11], with a σV_T around 5 mV. Whereas in [11] they report rare extreme cases with a single

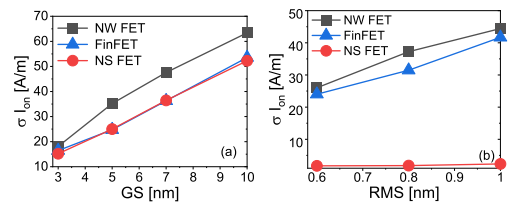


Fig. 3. On-current standard deviation (σI_{ON}) due to (a) the MGG and (b) the LER variability as a function of the GS and the RMS height, respectively.

dopant in the channel, in our devices the decay of the gaussian doping happens relatively close to the gate region, increasing the probability of finding one (or even more) dopants in this zone. If the exponential decay of the S/D doping were to happen 1 nm farther away from the gate, there is a 50% decrease in the observed σV_T due to RDD, highlighting the need for a nanoscale precision of ion implantation. In general, the MGG and the LER are the two sources of variability that influence the device performance in the sub-threshold region the most. Therefore, the on-region study of variability is limited to these two sources since MC-based variability studies are very computationally demanding.

Figs. 3(a)-(b) show the I_{ON} standard deviation due to MGG and LER, respectively. The I_{ON} has been normalised by the semiconductor perimeter. Only the sidewalls and the top of the FinFET (see table included in Fig. 1) are considered to calculate this perimeter since the thick BOX at the bottom of the structure prevents any real effect in the device conduction. Similarly to what we observed for the sub-threshold region figures of merit, the NS FET and the FinFET are equally affected by the MGG variability in the on-region being more resilient architectures than the NW FET. At a GS = 3 nm, the NW FET σI_{ON} is around a 20% larger than those of the NS FET and the FinFET. The LER induced variability also exhibits the same behaviour in the sub-threshold and on-region with a minimal influence of LER on σI_{ON} in the NS FET (of around 2 A/m). The LER variability in the FinFET makes this architecture a slightly more resilient than the NW FET one. For instance, at a RMS = 0.6 nm, σI_{ON} due to the LER for the FinFET is 24 A/m, a value 10% lower than that of the NW FET.

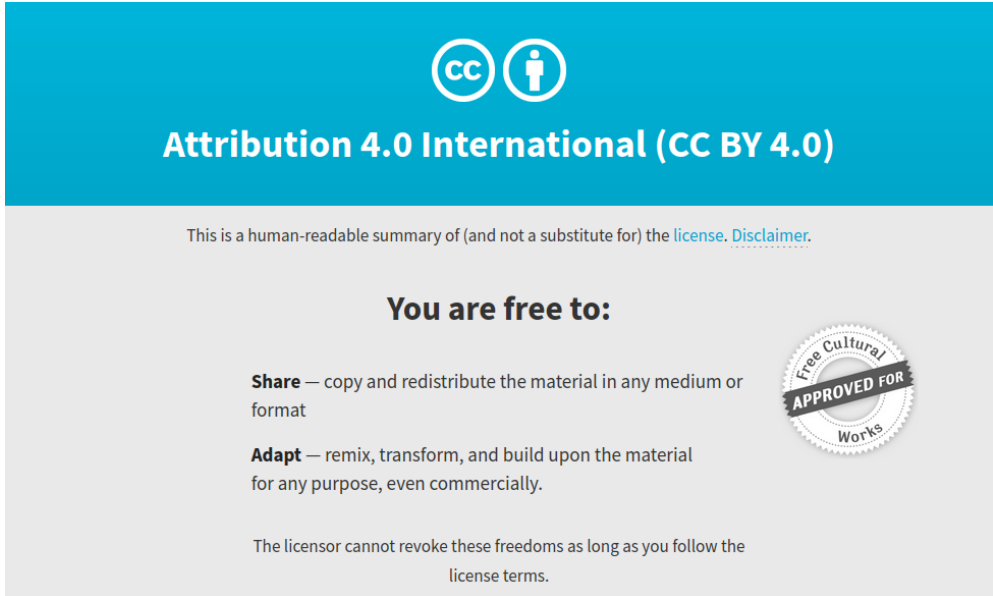
IV. CONCLUSION

The simulation study demonstrated that MGG and LER are the two sources of variability that influence the performance of the three analysed architectures (FinFET, NS FET, and NW FET) the most. However, the impact of MGG and LER can be greatly mitigated if the fabrication process substantially reduces the RMS and the GS values. The NW FET is the least resilient architecture against these two sources of variability, especially in the sub-threshold region, with σV_T values at least 50% larger than those of the FinFET and the NS FET. The three architectures are equally affected by the RDD variability while the effect of the GER is very small. When compared to the FinFET, the NS FET is similarly affected by the MGG variations but the effect of the LER on the NS FET is negligible indicating that this architecture is a suitable candidate to replace FinFETs from a variability point of view.

REFERENCES

- [1] J.-P. Colinge, *FinFETs and Other Multi-Gate Transistors*. Boston, MA, USA: Springer, 2008.
- [2] X. He, J. Fronheiser, P. Zhao, Z. Hu, S. Uppal, X. Wu, Y. Hu, R. Sporer, L. Qin, R. Krishnan, E. M. Bazizi, R. Carter, K. Tabakman, A. K. Jha, H. Yu, O. Hu, D. Choi, J. G. Lee, S. B. Samavedam, and D. K. Sohn, "Impact of aggressive fin width scaling on FinFET device characteristics," in *IEDM Tech. Dig.*, Dec. 2017, pp. 493–496, doi: 10.1109/IEDM.2017.8268427.
- [3] P. Feng, S. Song, G. Nallapati, J. Zhu, J. Bao, V. Moroz, M. Choi, X. Lin, Q. Lu, B. Colombeau, N. Breil, M. Chudzik, and C. Chidambaram, "Comparative analysis of semiconductor device architectures for 5-nm node and beyond," *IEEE Electron Device Lett.*, vol. 38, no. 12, pp. 1657–1660, Dec. 2017, doi: 10.1109/LED.2017.2769058.
- [4] Y. Tian, R. Huang, Y. Wang, J. Zhuge, R. Wang, J. Liu, X. Zhang, and Y. Wang, "New self-aligned silicon nanowire transistors on bulk substrate fabricated by epi-free compatible CMOS technology: Process integration, experimental characterization of carrier transport and low frequency noise," in *IEDM Tech. Dig.*, Dec. 2007, pp. 895–898, doi: 10.1109/IEDM.2007.4419094.
- [5] D. Nagy, G. Espiñeira, G. Indalecio, A. J. García-Loureiro, K. Kalna, and N. Seoane, "Benchmarking of FinFET, nanosheet, and nanowire FET architectures for future technology nodes," *IEEE Access*, vol. 8, pp. 53196–53202, 2020, doi: 10.1109/ACCESS.2020.2980925.
- [6] N. Loubet, T. Hook, P. Montanini, C.-W. Yeung, S. Kanakasabapathy, M. Guillom, T. Yamashita, J. Zhang, X. Miao, J. Wang, A. Young, R. Chao, M. Kang, Z. Liu, S. Fan, B. Hamieh, S. Sieg, Y. Mignot, W. Xu, S.-C. Seo, J. Yoo, S. Mochizuki, M. Sankarapandian, O. Kwon, A. Carr, A. Greene, Y. Park, J. Frougier, R. Galatage, R. Bao, J. Shearer, R. Conti, H. Song, D. Lee, D. Kong, Y. Xu, A. Arceo, Z. Bi, P. Xu, R. Muthinti, J. Li, R. Wong, D. Brown, P. Oldiges, R. Robison, J. Arnold, N. Felix, S. Skordas, J. Gaudiello, T. Standaert, H. Jagannathan, D. Corliss, M.-H. Na, A. Knorr, T. Wu, D. Gupta, S. Lian, R. Divakaruni, T. Gow, C. Labelle, S. Lee, V. Paruchuri, H. Bu, and M. Khare, "Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET," in *Proc. Symp. VLSI Technol.*, Jun. 2017, pp. T230–T231, doi: 10.23919/VLSIT.2017.7998183.
- [7] IRDS. (2020). *International Roadmap for Devices and Systems: More Moore*. [Online]. Available: <https://irds.ieee.org/editions/2020/more-moore>
- [8] K. J. Kuhn, M. D. Giles, D. Becher, P. Kolar, A. Kornfeld, R. Kotlyar, S. T. Ma, A. Maheshwari, and S. Mudanai, "Process technology variation," *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2197–2208, Aug. 2011, doi: 10.1109/TED.2011.2121913.
- [9] Y. Li, H.-T. Chang, C.-N. Lai, P.-J. Chao, and C.-Y. Chen, "Process variation effect, metal-gate work-function fluctuation and random dopant fluctuation of 10-nm gate-all-around silicon nanowire MOSFET devices," in *IEDM Tech. Dig.*, Dec. 2015, pp. 887–890, doi: 10.1109/IEDM.2015.7409827.
- [10] S. H. Rasouli, K. Endo, J. F. Chen, N. Singh, and K. Banerjee, "Grain-orientation induced quantum confinement variation in FinFETs and multi-gate ultra-thin body CMOS devices and implications for digital design," *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2282–2292, Aug. 2011, doi: 10.1109/TED.2011.2151196.
- [11] X. Wang, A. R. Brown, B. Cheng, and A. Asenov, "Statistical variability and reliability in nanoscale FinFETs," in *IEDM Tech. Dig.*, Dec. 2011, pp. 4–5, doi: 10.1109/IEDM.2011.6131494.
- [12] T. Yu, R. Wang, R. Huang, J. Chen, J. Zhuge, and Y. Wang, "Investigation of nanowire line-edge roughness in gate-all-around silicon nanowire MOSFETs," *IEEE Trans. Electron Devices*, vol. 57, no. 11, pp. 2864–2871, Nov. 2010, doi: 10.1109/TED.2010.2065808.
- [13] N. Seoane, D. Nagy, G. Indalecio, G. Espiñeira, K. Kalna, and A. García-Loureiro, "A multi-method simulation toolbox to study performance and variability of nanowire FETs," *Materials*, vol. 12, no. 15, p. 2391, Jul. 2019, doi: 10.3390/ma12152391.
- [14] M. A. Elmessary, D. Nagy, M. Aldegunde, N. Seoane, G. Indalecio, J. Lindberg, W. Dettmer, D. Perić, A. J. García-Loureiro, and K. Kalna, "Scaling/LER study of Si GAA nanowire FET using 3D finite element Monte Carlo simulations," *Solid-State Electron.*, vol. 128, pp. 7–24, 2017, doi: 10.1016/j.sse.2016.10.018.
- [15] S. Bangsaruntip, K. Balakrishnan, S.-L. Cheng, M. Brink, I. Lauer, R. L. Bruce, S. U. Engelmann, A. Pyzyna, G. M. Cohen, L. M. Gignac, C. M. Breslin, J. S. Newbury, D. P. Klaus, A. Majumdar, J. W. Sleight, and M. A. Guillorn, "Density scaling with gate-all-around silicon nanowire MOSFETs for the 10 nm node and beyond," in *IEDM Tech. Dig.*, Dec. 2013, pp. 526–529, doi: 10.1109/IEDM.2013.6724667.
- [16] A. Islam and K. Kalna, "Monte Carlo simulations of mobility in doped GaAs using self-consistent Fermi–Dirac statistics," *Semiconductor Sci. Technol.*, vol. 26, no. 5, Mar. 2011, Art. no. 055007, doi: 10.1088/0268-1242/26/5/055007.
- [17] M. Aldegunde, A. J. García-Loureiro, and K. Kalna, "3D finite element Monte Carlo simulations of multigate nanoscale transistors," *IEEE Trans. Electron Devices*, vol. 60, no. 5, pp. 1561–1567, May 2013, doi: 10.1109/TED.2013.2253465.
- [18] A. J. García-Loureiro, N. Seoane, M. Aldegunde, R. Valin, A. Asenov, A. Martinez, and K. Kalna, "Implementation of the density gradient quantum corrections for 3-D simulations of multigate nanoscaled transistors," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 30, no. 6, pp. 841–851, Jun. 2011, doi: 10.1109/TCAD.2011.2107990.
- [19] J. Lindberg, M. Aldegunde, D. Nagy, W. G. Dettmer, K. Kalna, A. J. García-Loureiro, and D. Perić, "Quantum corrections based on the 2-D Schrödinger equation for 3-D finite element Monte Carlo simulations of nanoscaled FinFETs," *IEEE Trans. Electron Devices*, vol. 61, no. 2, pp. 423–429, Feb. 2014, doi: 10.1109/TED.2013.2296209.
- [20] G. Indalecio, A. J. García-Loureiro, N. S. Iglesias, and K. Kalna, "Study of metal-gate work-function variation using Voronoi cells: Comparison of Rayleigh and gamma distributions," *IEEE Trans. Electron Devices*, vol. 63, no. 6, pp. 2625–2628, Jun. 2016, doi: 10.1109/TED.2016.2556749.
- [21] H. F. Dadvour, K. Endo, V. K. De, and K. Banerjee, "Grain-orientation induced work function variation in nanoscale metal-gate transistors—Part I: Modeling, analysis, and experimental validation," *IEEE Trans. Electron Devices*, vol. 57, no. 10, pp. 2504–2514, Oct. 2010, doi: 10.1109/TED.2010.2063191.
- [22] A. Asenov, S. Kaya, and A. R. Brown, "Intrinsic parameter fluctuations in decanometer MOSFETs introduced by gate line edge roughness," *IEEE Trans. Electron Devices*, vol. 50, no. 5, pp. 1254–1260, May 2003, doi: 10.1109/TED.2003.813457.
- [23] G. Espiñeira, D. Nagy, G. Indalecio, A. J. García-Loureiro, K. Kalna, and N. Seoane, "Impact of gate edge roughness variability on FinFET and gate-all-around nanowire FET," *IEEE Electron Device Lett.*, vol. 40, no. 4, pp. 510–513, Apr. 2019, doi: 10.1109/LED.2019.2900494.
- [24] C. Millar, D. Reid, G. Roy, S. Roy, and A. Asenov, "Accurate statistical description of random dopant-induced threshold voltage variability," *IEEE Electron Device Lett.*, vol. 29, no. 8, pp. 946–948, Aug. 2008, doi: 10.1109/LED.2008.2001030.
- [25] D. Nagy, G. Indalecio, A. J. García-Loureiro, M. A. Elmessary, K. Kalna, and N. Seoane, "FinFET versus gate-all-around nanowire FET: Performance, scaling, and variability," *IEEE J. Electron Devices Soc.*, vol. 6, pp. 332–340, Feb. 2018, doi: 10.1109/JEDS.2018.2804383.
- [26] G. F. Lorusso, O. Inoue, T. Ohashi, E. A. Sanchez, V. Constantoudis, and S. Koshihara, "Line width roughness accuracy analysis during pattern transfer in self-aligned quadruple patterning process," in *Proc. SPIE, 30th Metrol., Inspection, Process Control Microlithography*, M. I. Sanchez, Ed., vol. 9778, 2016, pp. 282–289, doi: 10.1117/12.2218863.
- [27] K. Patel, T.-J. K. Liu, and C. J. Spanos, "Gate line edge roughness model for estimation of FinFET performance variability," *IEEE Trans. Electron Devices*, vol. 56, no. 12, pp. 3055–3063, Dec. 2009, doi: 10.1109/TED.2009.2032605.

A.1.2. Pelgrom-based predictive model to estimate metal grain granularity and line edge roughness in advanced multigate MOS-FETs



The image shows a Creative Commons Attribution 4.0 International License banner. At the top, there are two icons: 'CC' and a person icon. Below them, the text reads 'Attribution 4.0 International (CC BY 4.0)'. A smaller line of text states: 'This is a human-readable summary of (and not a substitute for) the [license](#). [Disclaimer](#).' The main heading is 'You are free to:'. Underneath, two freedoms are listed: 'Share' (copy and redistribute) and 'Adapt' (remix, transform, and build upon). A circular seal on the right says 'Free Cultural APPROVED FOR Works'. At the bottom, it states: 'The licensor cannot revoke these freedoms as long as you follow the license terms.'

- Publication quality indicators:
 - Impact factor (JCR 2022): 2.3
 - Quartile: Q2
 - Category: Engineering, electrical & electronic
 - Rank: 157/349
- PhD candidate contribution:
 - Development and validation of mathematical models.
 - Data processing.
 - Writing of the manuscript.

Received 8 August 2022; revised 23 September 2022 and 11 October 2022; accepted 12 October 2022. Date of publication 17 October 2022; date of current version 26 October 2022. The review of this article was arranged by Editor P.-W. Li.

Digital Object Identifier 10.1109/JEDS.2022.3214928

Pelgrom-Based Predictive Model to Estimate Metal Grain Granularity and Line Edge Roughness in Advanced Multigate MOSFETs

JULIAN G. FERNANDEZ¹, NATALIA SEOANE¹ (Member, IEEE), ENRIQUE COMESAÑA (Member, IEEE), AND ANTONIO GARCÍA-LOUREIRO¹ (Member, IEEE)

Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain

CORRESPONDING AUTHOR: J. G. FERNANDEZ (e-mail: julian.garcia.fernandez2@usc.es)

This work was supported by the Spanish MICINN, Xunta de Galicia, and FEDER Funds under Grant RYC-2017-23312, Grant PID2019-104834GB-I00, Grant ED431F 2020/008, and Grant ED431C 2022/16.

ABSTRACT The impact of different variability sources on the transistor performance increases as devices are scaled-down, being the metal grain granularity (MGG) and the line edge roughness (LER) some of the major contributors to this increase. Variability studies require the simulation of large samples of different device configurations to have statistical significance, increasing the computational cost. A novel Pelgrom-based predictive (PBP) model that estimates the impact of MGG and LER through the study of the threshold voltage standard deviation (σV_{Th}), is proposed. This technique is computationally efficient since once the threshold voltage mismatch is calculated, σV_{Th} can be predicted for different gate lengths (L_g), cross-sections, and intrinsic variability parameters, without further simulations. The validity of the PBP model is demonstrated for three state-of-the-art architectures (FinFETs, nanowire FETs, and nanosheet FETs) with different L_g , cross-sections, and drain biases (V_D). The relative errors between the predicted and simulated data are lower than 10%, in the 92% of the cases.

INDEX TERMS FinFET, line edge roughness, metal grain granularity, nanosheet, nanowire, Pelgrom, prediction model, variability.

I. INTRODUCTION

Next generations of advanced transistors require an improvement in the gate control, together with the reduction of variability effects [1]. Some of the most relevant sources of variability are the line edge roughness (LER) [2], [3], [4], produced by lithographic processes; the metal grain granularity (MGG) [4], [5], [6], caused by the metal gate deposition; or the random dopant fluctuations (RDF) [7], that are due to the presence of dopants in the channel.

To accurately analyze the impact of different variability sources on these state-of-the-art architectures it is necessary: i) to use three-dimensional simulation approaches that incorporate the effects of quantum-mechanical confinement, and ii) to simulate large ensembles of different device configurations to obtain statistical significance. These two factors

notably increase the computational cost of these variability studies. Therefore, complementary techniques to predict variability were developed to lower the computational time, such as those based on machine learning (ML) [2], [8], [9], [10], [11], [12], the fluctuation sensitivity map (FSM) [13], the impedance field method [14], or the statistical model reported in [15].

In classical planar MOSFETs, the impact of the variability on the devices was estimated by the Pelgrom's Law [16], which states that the threshold voltage standard deviation (σV_{Th}) is proportional to the inverse square root of the effective metal gate area (A_g) and has been applied to different sources of variability [4], [5], [17]. This law can be extended to advanced tri-gate or gate-all-around (GAA) architectures by defining the A_g as the product between the gate length

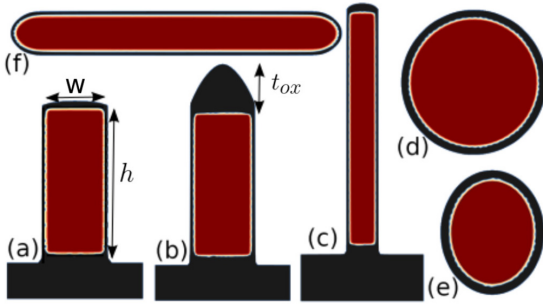


FIGURE 1. Cross-section of the three architectures used in this work. FinFETs: (a) from Wang et al. [19], and Leung and Chui [20], (b) from Seoane et al. [21], Indalecio et al. [22], and (c) from Seoane et al. [23]. NW FETs: (d) from Lü and Dai [24], Sung et al. [6], and (e) from Nagy et al. [25] and Elmessary et al. [3]. NS FETs: (f) from Nagy et al. [26] and this study.

(L_g) and the effective gate perimeter (W^*):

$$\sigma V_{Th} = \frac{A_v}{\sqrt{A_g}} = \frac{A_v}{\sqrt{L_g W^*}}, \quad (1)$$

where A_v is the threshold voltage matching coefficient, which is determined by the contributions of all possible sources of transistor variations [18].

In this work, we present a novel Pelgrom-based predictive (PBP) model as a simple, fast, and reliable tool to estimate the impact of variability on advanced semiconductor devices. This model has been developed by analyzing the dependence of A_v with the intrinsic parameters for each variability source. We have successfully applied the PBP model to predict the σV_{Th} due to MGG and LER on three state-of-the-art architectures: FinFET, GAA nanowire (NW) FET and GAA nanosheet (NS) FET.

The structure of this paper is organized as follows. Section II includes the description of the benchmark devices and the simulation methods used in this work. In Section III and Section IV, we develop the PBP model for MGG and LER, respectively, comparing the predicted results with those extracted from literature. Finally, Section V summarizes the main conclusions of this paper.

II. BENCHMARK DEVICES AND SIMULATION METHODOLOGY

In this section, we initially present the benchmark devices used in this work. We study three semiconductor advanced architectures: FinFETs, NW FETs, and NS FETs, see 2D cross-sectional schemes in Fig. 1. Most of the FinFETs [21], [22], NW FETs [3], [6], [25], and NS FETs [26] are either directly based on experimental devices (FinFET [27], NW FET [28], NS FET [29]) or are scaled down versions of them. The FinFET used in [23] (Fig. 1(c)) is a tri-gate version of the NS FET from [29], based on current IRDS predictions [1]. Note that, the FinFETs from Figs. 1(a)-1(b) have similar semiconductor width/height ratio, $w/h \sim 0.4$, and differ on the top oxide height (t_{ox}), whereas the FinFET

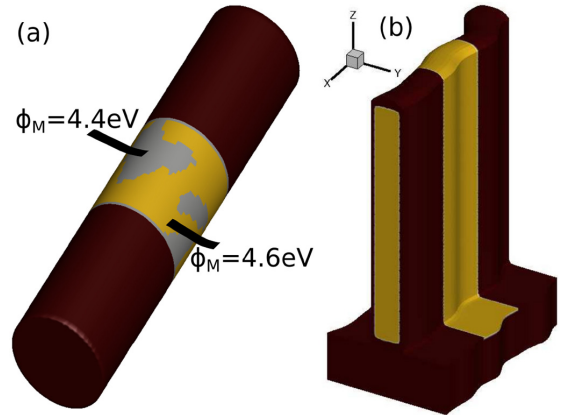


FIGURE 2. Scheme of (a) a gate-all-around (GAA) nanowire (NW) FET with a titanium nitride (TiN) metal gate affected by metal grain granularity (MGG) and (b) a FinFET affected by line edge roughness (LER) induced variability.

from Fig. 1(c) has a completely different ratio, $w/h \sim 0.1$. Also, the NW FETs have either circular ($w = h$, see Fig. 1(d)) or elliptical ($w \neq h$, see Fig. 1(e)) cross-sections. The NS FETs from this work (Fig. 1(f)) have a ratio of $w/h \sim 10.0$ based on [29].

All devices of this work have a titanium nitride (TiN) metal gate. TiN is a commonly used metal in the gate deposition due to several factors: its excellent thermal stability for the annealing process [30], its compatibility with silicon dioxide (SiO_2), and its low capacitance for ultralow-power circuit [31]. TiN has two possible metal grain orientations with work-functions (ϕ_M) of 4.4/4.6 eV, and occurrence probabilities of 40%/60%, respectively [32]. The MGG profiles are generated using Poisson-Voronoi diagrams depending on the average grain size (GS) in most of the analyzed works [3], [6], [21], [22], [23], [25], [26]. Fig. 2(a) shows a scheme of a NW FET affected by MGG variability.

The LER is applied in TCAD studies using the Fourier transform of the Gaussian spectra [22], [33]. The spectra depends on two parameters, the correlation length (Λ) and the root mean square height (Δ). The Λ is a parameter that indicates how the deformation propagates to neighboring points of the surface. The Δ is the height of the deformation in the direction of the normal vector of a surface. The deformation is applied in the perpendicular direction (Y-axis) to the transport direction (X-axis), as can be seen in Fig. 2(b) for a FinFET affected by LER. The profiles studied in this work correspond to random uncorrelated LER. The Λ is reported to vary between 5 and 40 nm [34], which is consistent with SEM (scanning electron microscope) studies from EUV (extreme ultraviolet) lithography reported by [33], [35].

For completeness, when data are not available in literature, we carried out simulations using VENDES [36], an in-house-built simulation toolbox that reproduces the physics via 3D quantum corrected (QC) drift-diffusion (DD) transport

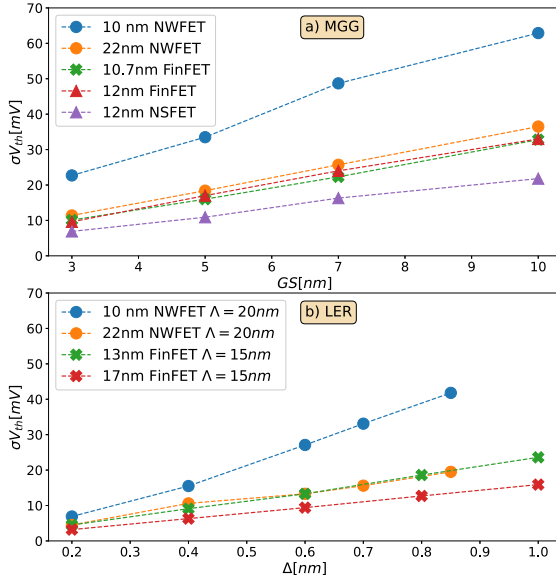


FIGURE 3. Threshold voltage standard deviation σV_{Th} versus: (a) the grain size (GS) for MGG variability, and (b) the root mean square (Δ) for LER variability. With these figures we validate the linearity assumption for the GS and Δ .

method. For each simulation, the threshold voltage (V_{Th}) has been extracted using the MLFoMpy post-processing tool [37], applying the constant current criteria [38]. This value is set as the drain current that corresponds to the V_{Th} extracted by the lineal extrapolation method from the ideal undeformed device I-V characteristics.

III. PBP MODEL FOR METAL GRAIN GRANULARITY

MGG appears because of the random orientations of the metal grains produced in the gate deposition process [39]. The average size of these grains (GS) depends on the annealing temperature [32]. In this section, we consider that the MGG is the only source of variability that contributes to A_v . σV_{Th} increases linearly with the GS, as shown in Fig. 3(a), and as also previously demonstrated in [40], therefore:

$$A_v = \theta_{m_{gg}} \cdot GS, \quad (2)$$

being $\theta_{m_{gg}}$ the threshold voltage mismatch for the MGG variability. Also, as the MGG is a gate induced variability, in the Pelgrom's Law (1), W_g^* is the effective gate perimeter (W_g). The MGG induced variability for a particular device architecture can be predicted with the PBP model as follows:

$$\sigma V_{Th} = \frac{\theta_{m_{gg}}}{\sqrt{L_g W_g}} \cdot GS. \quad (3)$$

MGG does not depend on the drain bias (V_D) as can be seen on [21], [25], and $\theta_{m_{gg}}$ will depend on the device architecture and the metal used in the deposition process. Next, we are

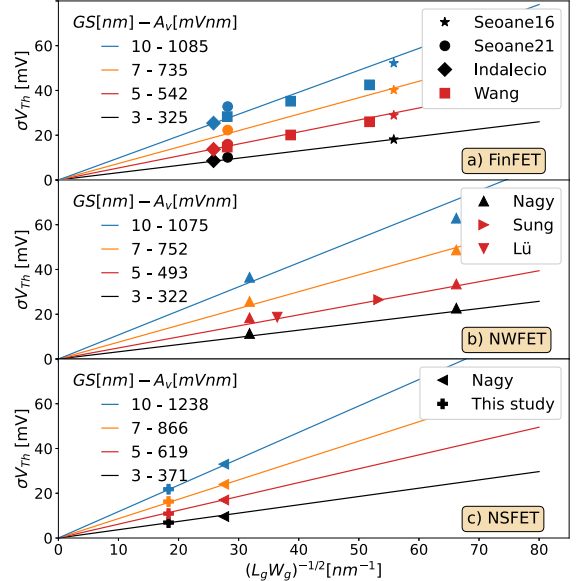


FIGURE 4. Pelgrom plots for the threshold voltage standard deviation (σV_{Th}) due to MGG variability at different grain sizes (GS), for: (a) FinFETs, (b) NW FETs, and (c) NS FETs. The analyzed data has been extracted from Seoane16 [21], Seoane21 [23], Indalecio et al. [22], Wang et al. [19], Nagy et al. [25], Lü and Dai [24], Sung et al. [6], Nagy et al. [26], or from simulations specifically done for this study. Note how the devices accurately follow the Pelgrom's Law and the lineal increase of the matching factor (A_v) with the GS.

going to validate the model, by analyzing literature data for the benchmark devices.

Fig. 4 shows the Pelgrom plots for σV_{Th} due to MGG variability for each architecture: Fig. 4(a) FinFETs, Fig. 4(b) NW FETs, and Fig. 4(c) NS FETs. The results for different GS are presented, as well as their corresponding A_v and $\theta_{m_{gg}}$ values. The simulations due to MGG reported in [6], [19], [21], [22], [23], [25], [26] together with this study simulations, were performed through QC DD methodology, but Lü and Dai [24] carried out classical DD simulations.

There is a linear relationship between σV_{Th} and $(L_g W_g)^{-1/2}$, following the PBP model for MGG (3). It is expected that the larger the gate area, the smaller the influence of the metal grains. Therefore, from a macroscopic point of view, when $(L_g W_g)^{-1/2} \rightarrow 0$ the gate will behave as a gate with a uniform ϕ_M value, and $\sigma V_{Th} \rightarrow 0$.

The slope of the Pelgrom plots (A_v) increases linearly with the GS for the three architectures (see Fig. 4) validating (2). Note that, all devices with the same architecture and metal on the gate, even having different shapes (see Fig. 1) and dimensions ($L_g W_g$), share a common $\theta_{m_{gg}}$. The W_g corresponds to the perimeter of gate surrounding the semiconductor for all devices presented in Fig. 1, except for the Fig. 1(b)-type FinFETs since, as demonstrated in [41] the grains located on the top of the gate do not impact on σV_{Th} .

TABLE 1. Predicted (σV_{Th}^P) vs. simulated (σV_{Th}^S) threshold voltage standard deviations at different grain sizes (GS) for each architecture due to metal grain granularity (MGG) induced variability. The relative error between the predicted and simulated data (σ_r), the gate length (L_g), the effective gate perimeter (W_g), and the V_{Th} mismatch for MGG variability θ_{mgg} are also shown. [*] corresponds to simulations done specifically for this work.

	GS [nm]	σV_{Th}^P [mV]	σV_{Th}^S [mV]	$\theta_{mgg} = 108 \text{ mV}$		Ref.
				σ_r	$L_g \times W_g$ [nm ²]	
FinFET	3	18.2	18.1	+0.6%	10.7×30.0	[21]
		9.2	10.1	-9.2%	12.0×105.0	[23]
		8.4	8.6	-2.3%	25.0×60.0	[22]
	5	30.3	28.9	+4.8%	10.7×30.0	[21]
		28.1	26.0	+8.1%	10.0×37.3	[19]
		21.0	20.1	+4.3%	14.0×47.8	[19]
		15.2	14.7	+3.3%	20.0×63.3	[19]
		15.3	16.0	-4.6%	12.0×105.0	[23]
		14.0	13.7	+2.2%	25.0×60.0	[22]
	7	42.4	40.2	+5.5%	10.7×30.0	[21]
		21.4	22.3	-4.0%	12.0×105.0	[23]
	8	22.4	21.8	+2.8%	25.0×60.0	[22]
10	60.6	52.2	+16.1%	10.7×30.0	[21]	
	56.1	42.5	+32.0%	10.0×37.3	[19]	
	41.9	35.2	+16.0%	14.0×47.8	[19]	
	29.1	28.3	+2.7%	20.0×63.3	[19]	
	30.6	32.8	-6.7%	12.0×105.0	[23]	
	28.0	25.4	+10.2%	25.0×60.0	[22]	
	GS [nm]	σV_{Th}^P [mV]	σV_{Th}^S [mV]	$\theta_{mgg} = 107 \text{ mV}$		Ref.
				σ_r	$L_g \times W_g$ [nm ²]	
NW FET	3	21.4	22.7	-5.7%	10.0×22.8	[25]
		10.3	11.4	-9.6%	22.0×44.9	[25]
		35.6	33.5	+6.3%	10.0×22.8	[25]
	5	28.7	26.5	+8.3%	10.0×35.2	[6]
		19.6	18.7	+4.8%	20.0×37.7	[24]
	7	17.1	18.4	-7.1%	22.0×44.9	[25]
		49.8	48.7	+2.3%	10.0×22.8	[25]
	10	23.9	25.7	-7.0%	22.0×44.9	[25]
		71.2	62.9	+13.2%	10.0×22.8	[25]
	34.2	36.5	-6.3%	22.0×44.9	[25]	
	GS [nm]	σV_{Th}^P [mV]	σV_{Th}^S [mV]	$\theta_{mgg} = 124 \text{ mV}$		Ref.
				σ_r	$L_g \times W_g$ [nm ²]	
NS FET	3	10.2	9.6	+6.3%	12.0×110.0	[26]
		6.8	6.9	-1.4%	18.0×165.3	[*]
		17.0	17.0	0.0%	12.0×110.0	[26]
	5	11.3	10.9	+3.7%	18.0×165.3	[*]
		23.9	24.0	-0.4%	12.0×110.0	[26]
	7	15.9	16.3	-2.5%	18.0×165.3	[*]
		34.1	33.0	+3.3%	12.0×110.0	[26]
	22.7	21.8	+4.1%	18.0×165.3	[*]	

Therefore, to take into account only the grains located on the sides, for the Fig. 1(b)-type FinFET we define the gate effective perimeter as $W_g = 2h$. Note that, the $\theta_{mgg}/\sqrt{L_g W_g}$ ratio provides the sensitivity of a certain device to MGG variability.

In Table 1, we show a comparison of the threshold voltage standard deviation predicted by the PBP model (σV_{Th}^P), against the σV_{Th} coming from simulation studies (σV_{Th}^S) available in literature. Also, the GS, the relative error between predicted and simulated results (σ_r), the effective gate area ($L_g \times W_g$), and θ_{mgg} are shown. In Table 1, [*] indicates data simulated specifically for this study. Note that, as mentioned above, for a particular device architecture and metal on the gate, θ_{mgg} does not change, and once is determined (via the simulation of the MGG variability for two ensembles of devices with different $L_g W_g$ values) the impact of the MGG for any GS and device dimensions ($L_g, W_g, w/h, t_{ox}$) can be estimated using the PBP model according to (3). The larger prediction inaccuracies are due to the saturation

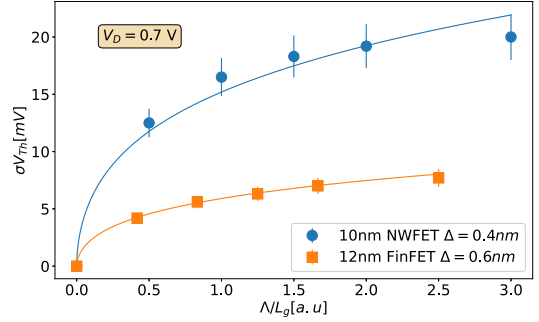


FIGURE 5. Threshold voltage standard deviation σV_{Th} versus the ratio between the correlation length and the gate length (λ/L_g) for two architectures: (a) FinFETs [21], [22], and (b) NW FETs [3]. The root mean square (Δ) of each architecture is also shown. The solid lines represent the fit of each device to its own $f(\lambda)$, and the y-error bars correspond to the 10% of each σV_{Th} value. The simulations were carried out at a drain bias (V_D) of 0.7 V. Note the good fit of $f(\lambda)$ for two devices presented in the plot.

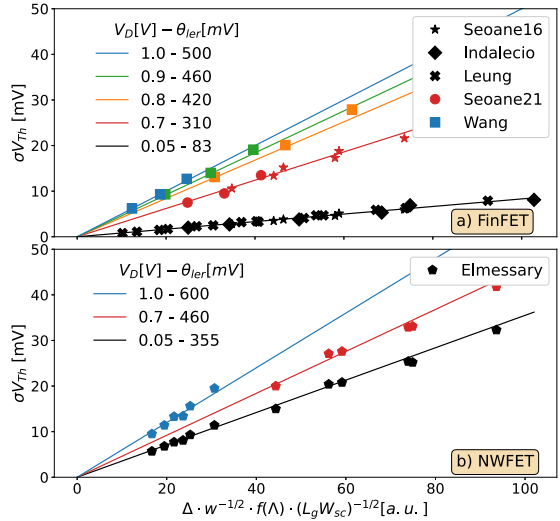


FIGURE 6. Threshold voltage standard deviation (σV_{Th}) versus $\Delta \cdot w^{-1/2} \cdot f(\lambda) \cdot (L_g W_{sc})^{-1/2}$ plots due to LER variability at different drain biases V_D , for: (a) FinFETs, (b) NW FETs. Data extracted from Seoane16 [21], Seoane21 [23], Indalecio et al. [22], Wang et al. [19], Leung and Chui [20], and Elmessary et al. [3]. Note the excellent fit of the model with simulated data and the increase of θ_{ler} with the V_D .

of the standard deviations in devices with small A_g and large GS. This effect occurs because there are a few metal grains on the gate (e.g., for GS 10 nm, there are only 4 grains on the 10.7 nm×30.0 nm FinFET gate), transforming the normal V_{Th} distribution [22] into a bimodal distribution with peaks around the extreme ϕ_M values (4.4/4.6 eV), and breaking the linearity assumed by the PBP model. Avoiding the extreme cases previously mentioned (marked in bold in Table 1), all the predictions have a relative error of 10% or lower, highlighting that 91% of errors are 8% or smaller.

TABLE 2. PBP model validation for LER. V_D is the drain bias, θ_{ler} the threshold voltage mismatch for LER variability, Λ the correlation length, $L_g \times W_{sc}$ the semiconductor area, and Δ the root mean square height. Also, the predicted (σV_{Th}^P), and the simulated (σV_{Th}^S) threshold voltage standard deviations are shown, together with their relative error (σ_r).

	V_D [V]	θ_{ler} [mV]	Λ [nm]	$L_g \times W_{sc}$ [nm]×[nm]	Δ [nm]	σV_{Th}^P [mV]	σV_{Th}^S [mV]	σ_r [%]	Ref.			
FinFET	0.05	83	50	25.0×72.0	1.00	2.8	2.7	+3.7	[22]			
					2.00	5.6	5.3	+5.7				
					3.00	8.4	8.1	+3.7				
				20	10.7×35.0	0.60	3.7	3.5		+5.7	[21]	
						0.80	4.9	5.1		-3.9		
						1.00	6.1	6.0		+1.7		
			25.0×72.0	1.00	2.1	2	+5.0	[22]				
				2.00	4.1	3.8	+7.9					
				3.00	6.3	6.9	-8.7					
			15	13.0×12.8	0.20	1.5	1.5	0.0	[20]			
					0.40	3.1	3.2	-3.1				
					0.60	4.6	4.7	-2.1				
	0.80	6.1			6.3	-3.2						
	1.00	7.7			7.9	-2.5						
	1.00	1.1			1.1	0.0	[20]					
	0.40	2.3	2.3	0.0								
	0.60	3.4	3.4	0.0								
	0.80	4.6	4.7	-2.1								
	1.00	5.7	5.9	-3.4								
	1.00	0.8	0.8	0.0	[20]							
	0.40	1.7	1.7	0.0								
	0.60	2.5	2.5	0.0								
	0.80	3.4	3.3	+3.0								
	1.00	4.2	4.1	+2.4								
1.00	2.9	2.8	+3.6	[21]								
0.80	3.9	3.8	+2.6									
1.00	4.8	4.6	+4.2									
0.70	310	20	10.7×35.0		0.60	13.7	13.4	+2.2	[21]			
					0.80	18.3	18.8	-2.7				
					1.00	22.8	21.6	+5.6				
		12.0×105.0	0.60	7.7	7.5	+2.7	[23]					
			0.80	10.3	9.5	+8.4						
			1.00	12.8	13.5	-5.2						
10	10.7×35.0	0.60	10.8	10.6	+1.9	[21]						
		0.80	14.4	15.2	-5.3							
		1.00	18.0	17.3	+4.0							
0.80	420	30	10.0×37.3	0.33	13.0	13.1	-0.8	[19]				
				0.50	19.7	20.1	-2.0					
				0.66	26.0	27.9	-6.8					
0.90	460	30	14.0×47.8	0.33	9.1	9.3	+2.1	[19]				
				0.50	13.8	14.0	+1.4					
				0.66	18.2	19.1	-4.7					
1.00	500	30	20.0×63.3	0.33	6.2	6.2	0.0	[19]				
				0.50	9.3	9.3	0.0					
				0.66	12.3	12.7	+3.1					
NW FET	0.05	355	20	10.0×20.3	0.60	19.9	20.4	-2.5	[3]			
					0.70	26.6	25.2	+5.6				
					0.85	33.2	32.3	+2.8				
				22.0×40.2	0.60	7.7	7.7	0.0		[3]		
					0.70	9.0	9.3	-3.2				
					0.85	10.9	11.4	-4.4				
			10	10.0×20.3	0.60	15.7	15.0	+4.6	[3]			
					0.70	21.0	20.8	+1.0				
					0.85	26.2	25.4	+3.1				
					22.0×40.2	0.60	5.9	5.7		+3.5	[3]	
						0.70	6.9	6.8		+1.5		
						0.85	8.4	8.1		+3.7		
	0.70	460	20	10.0×20.3	0.60	25.8	27.1	-4.8	[3]			
					0.70	34.4	33.1	+3.9				
					0.85	43.1	41.8	+3.1				
					10	10.0×20.3	0.60	20.4		20.0	+2.0	[3]
							0.70	27.2		27.6	-1.5	
							0.85	34.0		32.9	+3.3	
	1.00	600	20	22.0×40.2	0.60	13.0	13.3	-2.3	[3]			
					0.70	15.1	15.6	-3.2				
					0.85	18.4	19.5	-5.6				
					10	22.0×40.2	0.60	10.0		9.5	+5.3	[3]
							0.70	11.7		11.4	+2.6	
							0.85	14.2		13.3	+6.8	

IV. PBP MODEL FOR LINE EDGE ROUGHNESS

The lithographic processes used in the production of nano-electronic devices result in the appearance of surface roughness, affecting the device performance [33].

In this section, we consider that LER variability will be the sole contributor to A_v . LER induced σV_{Th} variability increases linearly with Δ , as shown in Fig. 3(b) and in the

literature [3], [21], [22], [23], [25]. Also, σV_{Th} is inversely proportional to the semiconductor width (w , see Fig. 1), the critical dimension of deformation, since the narrower the device, the larger the expected effect of LER. In addition, σV_{Th} is proportional to a function of the correlation length ($f(\Lambda)$). This $f(\Lambda)$ is empirically defined based on the auto-correlation function that reproduces the surface roughness through the Gaussian spectra. Considering all these factors, the matching factor A_v is as follows:

$$A_v = \theta_{ler} \cdot \frac{\Delta}{\sqrt{w}} \cdot f(\Lambda), \tag{4}$$

with,

$$f(\Lambda) = \sqrt{\Lambda \cdot \left(1 - e^{-\sqrt{L_g/\Lambda}}\right)}, \tag{5}$$

being θ_{ler} the threshold voltage mismatch for the LER variability. Fig. 5 shows how σV_{Th} increases with the Λ/L_g ratio for a 10 nm NWFET with $\Delta = 0.4$ nm and a 12 nm FinFET with $\Delta = 0.6$ nm. The solid lines in Fig. 5 represent the $f(\Lambda)$ (5) that fits the data, and the vertical error bars correspond to the 10% of the σV_{Th} values. Since LER impacts the semiconductor width by narrowing or widening it, in the Pelgrom’s Law (1), W^* is the semiconductor perimeter (W_{sc}). With the previous concepts, the PBP model for the LER induced variability is described as follows:

$$\sigma V_{Th} = \theta_{ler} \cdot \frac{\Delta}{\sqrt{w}} \cdot \frac{f(\Lambda)}{\sqrt{L_g W_{sc}}} \tag{6}$$

Unlike for the MGG variability, the V_D has a high impact on the increase of σV_{Th} for LER [3], [21], [25]. Therefore, θ_{ler} will depend on the device architecture and the V_D .

Fig. 6 shows σV_{Th} versus $\Delta f(\Lambda)(wL_gW_{sc})^{-1/2}$ for the FinFETs (top figure), and the NW FETs (bottom figure). Also, the θ_{ler} values are presented at different drain biases (V_D).

Note that, LER variability is not studied for the NS FET since, as seen in [23], the impact on this architecture is negligible ($\sigma V_{Th} \leq 1.0$ mV) because the roughness due to the etching processes is on the non-critical dimension. The simulations due to LER variability reported in [3], [19], [21], [22], [23] were performed with QC DD methodology.

θ_{ler} is determined as the slope of the Fig. 6 plots, increasing with the V_D for the two architectures. For each architecture and V_D , all devices independently on the shape (see Fig. 1(a)-(e)), Λ and Δ share a common θ_{ler} . Once θ_{ler} is determined, the PBP model can be used to predict the impact of the LER for any set of Λ , Δ parameters and device dimensions (L_g , W_{sc} , w). Note that although θ_{ler} has no physical meaning, the $\theta_{ler}/\sqrt{L_g W_{sc}}$ ratio provides an estimation of the sensitivity of a device to LER variability. Table 2 shows the comparison between the predicted (σV_{Th}^P) and simulated (σV_{Th}^S) data, to estimate the accuracy of the PBP model for LER. A wide range of Λ (10 nm to 50 nm), and Δ (0.2 nm to 1.0 nm) values are presented. We observe that for all data the relative errors (σ_r) between the predicted

and simulated σV_{Th} values are below 9%, with 94% of them being below 6%, demonstrating the validity of the estimation provided by the Pelgrom-based predictive model.

V. CONCLUSION

We have presented a novel Pelgrom-based predictive (PBP) model that characterizes the impact of metal grain granularity (MGG) and line edge roughness (LER) induced variabilities on semiconductor device performance. The PBP model was applied to state-of-the-art FinFETs, NW FETs, and NS FETs, with prediction errors of 10% or lower for all data inside the model limits, demonstrating its adaptability to a variety of architectures, dimensions, cross-sectional shapes, and intrinsic variability parameters. Also for MGG, the 91% of the relative errors are 8% or lower and for LER, the 94% of the relative errors are below the 6%. The PBP model only fails to predict the MGG when the gate area ($L_g W_g$) and the grain size (GS) are comparable.


We have demonstrated that the PBP model provides a simple, fast, and reliable estimation of the impact of the major contributors to nanoelectronic transistor variability, with several dimensions, and architectures. This predictive tool could guide the industry and academia in the development of new generation variability-resistant device architectures.

REFERENCES

- [1] *International Roadmap for Devices and Systems: More Moore*, IRDS, Piscataway, NJ, USA, 2021.
- [2] S. Yu, S. M. Won, H. W. Baac, D. Son, and C. Shin, "Quantitative evaluation of line-edge roughness in various FinFET structures: Bayesian neural network with automatic model selection," *IEEE Access*, vol. 10, pp. 26340–26346, 2022.
- [3] M. A. Elmessary et al., "Scaling/LER study of Si GAA nanowire FET using 3D finite element Monte Carlo simulations," *Solid-State Electron.*, vol. 128, pp. 17–24, Feb. 2017.
- [4] A. Sudarsanan, S. Venkateswarlu, and K. Nayak, "Impact of fin line edge roughness and metal gate granularity on variability of 10-nm node SOI n-FinFET," *IEEE Trans. Electron Devices*, vol. 66, no. 11, pp. 4646–4652, Nov. 2019.
- [5] K. Nayak, S. Agarwal, M. Bajaj, P. J. Oldiges, K. V. R. M. Murali, and V. R. Rao, "Metal-gate granularity-induced threshold voltage variability and mismatch in Si gate-all-around nanowire n-MOSFETs," *IEEE Trans. Electron Devices*, vol. 61, no. 11, pp. 3892–3895, Nov. 2014.
- [6] W.-L. Sung, Y.-S. Yang, and Y. Li, "Work-function fluctuation of gate-all-around silicon nanowire n-MOSFETs: A unified comparison between cuboid and Voronoi methods," *IEEE J. Electron Devices Soc.*, vol. 9, pp. 151–159, 2021.
- [7] N. Seoane, G. Indalecio, D. Nagy, K. Kalna, and A. J. Garcia-Loureiro, "Impact of cross-sectional shape on 10-nm gate length InGaAs FinFET performance and variability," *IEEE Trans. Electron Devices*, vol. 65, no. 2, pp. 456–462, Feb. 2018.
- [8] C. Akbar, Y. Li, and W. L. Sung, "Machine learning aided device simulation of work function fluctuation for multichannel gate-all-around silicon nanosheet MOSFETs," *IEEE Trans. Electron Devices*, vol. 68, no. 11, pp. 5490–5497, Nov. 2021.
- [9] C. Akbar, Y. Li, and W.-L. Sung, "Deep learning algorithms for the work function fluctuation of random nanosized metal grains on gate-all-around silicon nanowire MOSFETs," *IEEE Access*, vol. 9, pp. 73467–73481, 2021.
- [10] J. Lee, T. Park, H. Ahn, J. Kwak, T. Moon, and C. Shin, "Prediction model for random variation in FinFET induced by line-edge-roughness (LER)," *Electronics*, vol. 10, no. 4, p. 455, 2021.
- [11] J. Lim and C. Shin, "Machine learning (ML)-based model to characterize the line edge roughness (LER)-induced random variation in FinFET," *IEEE Access*, vol. 8, pp. 158237–158242, 2020.
- [12] H. Carrillo-Nuñez, N. Dimitrova, A. Asenov, and V. Georgiev, "Machine learning approach for predicting the effect of statistical variability in Si junctionless nanowire transistors," *IEEE Electron Device Lett.*, vol. 40, no. 9, pp. 1366–1369, Sep. 2019.
- [13] G. Indalecio, N. Seoane, K. Kalna, and A. J. Garcia-Loureiro, "Fluctuation sensitivity map: A novel technique to characterise and predict device behaviour under metal grain work-function variability effects," *IEEE Trans. Electron Devices*, vol. 64, no. 4, pp. 1695–1701, Apr. 2017.
- [14] A. Wettstein, O. Penzin, E. Lyumkis, and W. Fichtner, "Random dopant fluctuation modelling with the impedance field method," in *Proc. Int. Conf. Simul. Semicond. Process. Devices (SISPAD)*, 2003, pp. 91–94.
- [15] W.-L. Sung and Y. Li, "Statistical prediction of nanosized-metal-grain-induced threshold-voltage variability for 3D vertically stacked silicon gate-all-around nanowire n-MOSFETs," *J. Electron. Mater.*, vol. 49, no. 11, pp. 6865–6871, Aug. 2020.
- [16] M. J. M. Pelgrom, A. C. J. Duijnmaier, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, Oct. 1989.
- [17] K. Takeuchi et al., "Understanding random threshold voltage fluctuation by comparing multiple fabs and technologies," in *Proc. IEEE Int. Electron Devices Meeting*, 2007, pp. 467–470.
- [18] P. A. Stolk, F. P. Widdershoven, and D. B. M. Klaassen, "Modeling statistical dopant fluctuations in MOS transistors," *IEEE Trans. Electron Devices*, vol. 45, no. 9, pp. 1960–1971, Sep. 1998.
- [19] X. Wang, A. R. Brown, B. Cheng, and A. Asenov, "Statistical variability and reliability in nanoscale FinFETs," in *Proc. Int. Electron Devices Meeting*, 2011, pp. 1–4.
- [20] G. Leung and C. O. Chui, "Variability of inversion-mode and junctionless FinFETs due to line edge roughness," *IEEE Electron Device Lett.*, vol. 32, no. 11, pp. 1489–1491, Nov. 2011.
- [21] N. Seoane et al., "Comparison of fin-edge roughness and metal grain work function variability in InGaAs and Si FinFETs," *IEEE Trans. Electron Devices*, vol. 63, no. 3, pp. 1209–1216, Mar. 2016.
- [22] G. Indalecio, M. Aldegunde, N. Seoane, K. Kalna, and A. J. Garcia-Loureiro, "Statistical study of the influence of LER and MGG in SOI MOSFET," *Semicond. Sci. Technol.*, vol. 29, no. 4, Feb. 2014, Art. no. 45005.
- [23] N. Seoane, J. G. Fernandez, K. Kalna, E. Comesaña, and A. Garcia-Loureiro, "Simulations of statistical variability in n-type FinFET, nanowire, and nanosheet FETs," *IEEE Electron Device Lett.*, vol. 42, no. 10, pp. 1416–1419, Oct. 2021.
- [24] W.-F. Lü and L. Dai, "Impact of work-function variation on analog figures-of-merits for high- k /metal-gate junctionless FinFET and gate-all-around nanowire MOSFET," *Microelectron. J.*, vol. 84, pp. 54–58, Feb. 2019.
- [25] D. Nagy, G. Indalecio, A. J. Garcia-Loureiro, M. A. Elmessary, K. Kalna, and N. Seoane, "FinFET versus gate-all-around nanowire FET: Performance, scaling, and variability," *IEEE J. Electron Devices Soc.*, vol. 6, pp. 332–340, 2018.
- [26] D. Nagy, G. Espiñeira, G. Indalecio, A. J. Garcia-Loureiro, K. Kalna, and N. Seoane, "Benchmarking of FinFET, nanosheet, and nanowire FET architectures for future technology nodes," *IEEE Access*, vol. 8, pp. 53196–53202, 2020.
- [27] V. S. Basker et al., "A 0.063 μm^2 FinFET SRAM cell demonstration with conventional lithography using a novel integration scheme with aggressively scaled fin and gate pitch," in *Proc. Symp. VLSI Technol.*, 2010, pp. 19–20.
- [28] S. Bangsaruntip et al., "Density scaling with gate-all-around silicon nanowire MOSFETs for the 10 nm node and beyond," in *Proc. IEEE Int. Electron Devices Meeting*, 2013, pp. 1–4.
- [29] N. Loubet et al., "Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET," in *Proc. Symp. VLSI Technol.*, 2017, pp. T230–T231.
- [30] M. C. Lemme et al., "Nanoscale TiN metal gate technology for CMOS integration," *Microelectron. Eng.*, vol. 83, nos. 4–9, pp. 1551–1554, Apr.–Sep. 2006.
- [31] S. A. Vitale, J. Kedzierski, P. Healey, P. W. Wyatt, and C. L. Keast, "Work-function-tuned TiN metal gate FDSOI transistors for sub-threshold operation," *IEEE Trans. Electron Devices*, vol. 58, no. 2, pp. 419–426, Feb. 2011.

- [32] H. Dadgour, K. Endo, V. De, and K. Banerjee, "Modeling and analysis of grain-orientation effects in emerging metal-gate devices and implications for SRAM reliability," in *Proc. IEEE Int. Electron Devices Meeting*, 2008, pp. 1–4.
- [33] A. Asenov, S. Kaya, and A. R. Brown, "Intrinsic parameter fluctuations in decananometer MOSFETs introduced by gate line edge roughness," *IEEE Trans. Electron Devices*, vol. 50, no. 5, pp. 1254–1260, May 2003.
- [34] E. Giannatou, G. Papavarios, V. Constantoudis, H. Papageorgiou, and E. Gogolides, "Deep learning denoising of SEM images towards noise-reduced LER measurements," *Microelectron. Eng.*, vol. 216, Aug. 2019, Art. no. 111051. [Online]. Available: <https://www.science-direct.com/science/article/pii/S0167931719302084>
- [35] G. F. Lorusso, O. Inoue, T. Ohashi, E. A. Sanchez, V. Constantoudis, and S. Koshihara, "Line width roughness accuracy analysis during pattern transfer in self-aligned quadruple patterning process," in *Proc. SPIE Metrol. Inspect. Process Control Microlithogr. XXX*, Mar. 2016, Art. no. 97780V.
- [36] N. Seoane, D. Nagy, G. Indalecio, G. Espiñeira, K. Kalna, and A. García-Loureiro, "A multi-method simulation toolbox to study performance and variability of nanowire FETs," *Materials*, vol. 12, no. 15, p. 2391, 2019.
- [37] G. Espiñeira, N. Seoane, D. Nagy, G. Indalecio, and A. J. García-Loureiro, "FoMPy: A figure of merit extraction tool for semiconductor device simulations," in *Proc. Joint Int. EUROSOL Workshop Int. Conf. Ultimate Integr. Silicon (EUROSOL-ULIS)*, 2018, pp. 1–4.
- [38] A. Ortiz-Conde, F. J. García-Sánchez, J. Muci, A. T. Barrios, J. J. Liou, and C.-S. Ho, "Revisiting MOSFET threshold voltage extraction methods," *Microelectron. Rel.*, vol. 53, no. 1, pp. 90–104, 2013. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0026271412004684>
- [39] H.-W. Cheng et al., "3D device simulation of work function and interface trap fluctuations on high- κ /metal gate devices," in *Proc. Int. Electron Devices Meeting*, 2010, pp. 1–4.
- [40] D. Nagy, G. Indalecio, A. J. García-Loureiro, M. A. Elmessary, K. Kalna, and N. Seoane, "Metal grain granularity study on a gate-all-around nanowire FET," *IEEE Trans. Electron Devices*, vol. 64, no. 12, pp. 5263–5269, Dec. 2017.
- [41] J. G. Fernandez, N. Seoane, E. Comesaña, K. Kalna, and A. García-Loureiro, "Impact of metal grain granularity on three gate-all-around advanced architectures," in *Proc. Int. Conf. Simul. Semicond. Process. Devices (SISPAD)*, 2021, pp. 201–205.

A.1.3. A comprehensive Pelgrom-based on-current variability model for FinFET, NWFET and NSFET



A comprehensive Pelgrom-based on-current variability model for FinFET, NWFET and NSFET
Author: Julian G. Fernandez, Natalia Seoane, Enrique Comesaña, Antonio Garcia-Loureiro
Publication: Solid-State Electronics
Publisher: Elsevier
Date: January 2023
© 2022 The Author(s). Published by Elsevier Ltd.

Journal Author Rights

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

[BACK](#) [CLOSE WINDOW](#)

- Publication quality indicators:
 - Impact factor (JCR 2022): 1.7
 - Quartile: Q3
 - Category: Physics, Condensed matter
 - Rank: Rank 45/75
- PhD candidate contribution:
 - Development and validation of mathematical models.
 - Data processing.
 - Writing of the manuscript.



A comprehensive Pelgrom-based on-current variability model for FinFET, NWFET and NSFET[☆]

Julian G. Fernandez^{a,*}, Natalia Seoane^a, Enrique Comesaña^b, Antonio Garcia-Loureiro^a

^a Centro Singular de Investigación en Tecnoloxías Intelixentes, Universidade de Santiago de Compostela, Spain

^b Escola Politécnica Superior de Enxeñaría, Campus Terra, Universidade de Santiago de Compostela, Spain

ARTICLE INFO

Keywords:

TCAD
FinFET
Nanowire FET
Nanosheet FET
Pelgrom
Prediction model
Monte carlo

ABSTRACT

We present a novel Pelgrom-based predictive (PBP) model to estimate the impact of variability on the on-current of different state-of-the-art semiconductor devices. In this work, we focus on two of the most problematic sources of variability, the metal grain granularity (MGG) and the line edge roughness (LER). This model allows us to make an accurate prediction of the on-current standard deviation σI_{on} , being the relative error of the predicted data lower than 8% in 92% of the studied cases. The PBP model entails an immense reduction in the computational cost since once it is calibrated for an architecture, the prediction of the impact of a variability on devices with any given dimension can be made without any further simulations. This model could be useful for predicting the effect of variability on future technology nodes.

1. Introduction

The scaling of semiconductor devices is essential for the progress of the electronic industry, but their small size requirements make them more vulnerable to different sources of variability. The metal grain granularity (MGG) and the line edge roughness (LER) are two of the variability sources that have the greatest impact on state-of-the-art transistor architectures [1]: FinFETs, nanowire (NW) FETs, and nanosheet (NS) FETs. Technology computer-aided design (TCAD) is commonly used in variability studies since, to obtain statistical significance, we require the analysis of a large number of samples. The computational cost of on-region variability studies can be prohibitive because either quantum-corrected (QC) Monte Carlo (MC), or full quantum simulations are necessary to properly capture non-equilibrium transport effects. For this reason, new strategies are needed to reduce the computational times of statistical studies [2,3].

In this work, we propose a novel Pelgrom-based predictive (PBP) model to estimate the on-current (I_{on}) variability due to MGG and LER, for three state-of-the-art architectures (FinFET, NWFET, and NSFET). The structure of this paper is as follows. Section 2 presents the assumptions, the calibration, and the accuracy of the model for MGG (Section 2.1), and LER (Section 2.2), together with the computational cost of the model (Section 2.3). Finally, the main conclusions of this work are presented in Section 3.

2. Pelgrom-based σI_{on} predictive model

Pelgrom's Law [4] states that the standard deviation of a figure of merit, I_{on} in this particular case, is proportional to the inverse square root of the effective gate area, defined as the product between the gate length (L) and the effective gate perimeter (W) [5,6]:

$$\sigma I_{on} = \frac{A_i}{\sqrt{LW}}, \quad (1)$$

where A_i is the on-current matching factor, determined by the contributions of all possible sources of transistor variations [7].

There are previous studies for multigate transistors that show that MGG and LER variabilities follow the Pelgrom's law [8,9]. Therefore, we have developed two models based on Pelgrom's Law, for these two sources of variability, applying them to the three architectures shown in Fig. 1: (a) NWFET, (b) FinFET, and (c) NSFET. To validate the PBP model we have used data published by different authors [10–13] combined with simulations that were specifically done for this study using QC MC methodology in VENDES [14] tool.

2.1. PBP model for MGG

The MGG variability is implemented in TCAD studies by generating random grain distribution in the gate using the Poisson–Voronoi

[☆] The review of this paper was arranged by Francisco Gamiz.

* Corresponding author.

E-mail address: julian.garcia.fernandez2@usc.es (J.G. Fernandez).

¹ Work supported by the Spanish MICINN, Xunta de Galicia and FEDER funds (RYC-2017-23312, PID2019-104834GB-I00, ED431F 2020/008, ED431C 2022/16).

<https://doi.org/10.1016/j.sse.2022.108492>

Available online 11 November 2022

0038-1101/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

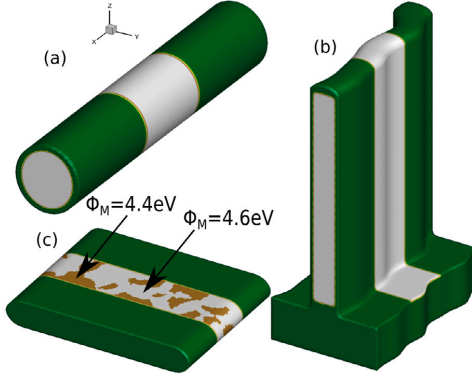


Fig. 1. Scheme of a (a) NWFET not affected by any variability source, (b) FinFET affected by LER variability, and (c) NSFET affected by TiN MGG with two gate work-functions (ϕ_M).

diagram methodology depending on the average grain size (GS) (see Fig. 1(c)), this methodology has been validated against experimental data in [15]. As all devices of this work have a TiN metal gate, the generated profiles have two metal grain orientations with work-functions of 4.4 eV, and 4.6 eV, and occurrence probabilities of 40%, and 60%, respectively. The GS is determined by the annealing temperature and deposition time of the metal gate deposition process [16]. To develop the PBP model for MGG variability, we assume that MGG is the sole contributor to transistor variations, and therefore $A_i = A_{mgs}$. Also, we assume that σI_{on} depends linearly on the GS as reported in [10,12], leading to the PBP model for MGG:

$$\sigma I_{on} = A_{mgs} \frac{1}{\sqrt{LW}} = \theta_{mgs} \frac{GS}{\sqrt{LW}} \quad (2)$$

We define the θ_{mgs} as the on-current mismatch for MGG, which is a technological parameter that states how the variability impacts a certain architecture, and once is determined, σI_{on} could be predicted for any device dimensions at any GS .

The PBP model plots for MGG are shown for the three architectures in Fig. 2(a)–(c). The minimum sample size used for each set of simulations is 300 in order to obtain statistical significance. The black line is the reference line that denotes the fitting of the model, where the error bars correspond to 10% of the relative error with respect to the simulated data, a tolerance that we consider acceptable. As can be seen, the majority of the predictions are inside the 10% margin. Table 1 presents the predicted (σI_{on}^P) and simulated (σI_{on}^S) on-current standard deviations together with the relative error (σ_r) between them, the GS , the θ_{mgs} , and dimensional information of the devices (LW). The estimations outside the 10% margin of error are highlighted in bold in the table, and are due to the saturation of the standard deviation in devices with small LW and large GS (i.e. NWFET $L = 10$ nm with $GS = 7$ nm and 10 nm), this phenomenon is described in [17].

2.2. PBP model for LER

LER variability is reproduced in TCAD studies using the Fourier transform of the Gaussian spectra [18] (see Fig. 1(b)), depending on two parameters: the root mean square height (Δ , depth of the roughness) and the correlation length (A , propagation of the roughness). In this case, the LER is the sole contributor to variability $A_i = A_{ler}$. Also, we assume that σI_{on} increases linearly with Δ (see [1,12]), depending on the square root of the product between the device width (w) and a function of A , which is empirically defined as:

$$f(A) = A \cdot \left[1 - e^{-\sqrt{\frac{A}{\lambda}}} \right] \quad (3)$$

Table 1

Predicted (σI_{on}^P) vs. simulated (σI_{on}^S) on-current standard deviations, at different GS for each architecture. In the table, we also list the on-current mismatch θ_{mgs} , the gate effective gate area (LW), and the relative error σ_r .

	GS		$\theta_{mgs} = 120$ nA/nm		Ref.	
	[nm]	[A/m]	σ_r	$L \times W$ [nm ²]		
NWFET	3	23.8	25.0	-4.8%	10.0 × 22.8	[12]
		11.5	12.5	-8.0%	22.0 × 44.9	[12]
	5	39.7	42.5	-6.6%	10.0 × 22.8	[12]
		19.1	19.0	+0.5%	10.0 × 35.2	[11]
		32.0	33.0	-5.3%	22.0 × 44.9	[12]
	7	55.6	47.7	+16.6%	10.0 × 22.8	[12]
26.7		26.3	+1.5%	22.0 × 44.9	[12]	
10	79.5	67.7	+17.4%	10.0 × 22.8	[12]	
	38.2	42.4	-9.9%	22.0 × 44.9	[12]	
	GS		$\theta_{mgs} = 192$ nA/nm		Ref.	
	[nm]	[A/m]	σ_r	$L \times W$ [nm ²]		
FinFET	3	14.9	15.0	-0.7%	25.0 × 60.0	^a
		16.3	16.3	+0.0%	12.0 × 105.0	^a
	5	52.7	58.6	-8.4%	10.7 × 30.0	[10]
		24.9	24.9	+0.0%	25.0 × 60.0	^a
		27.1	27.1	+0.0%	12.0 × 105.0	^a
	7	75.2	75.9	-0.9%	10.7 × 30.0	[10]
34.8		36.4	-4.5%	25.0 × 60.0	^a	
10	107.4	105.9	+1.4%	10.7 × 30.0	[10]	
	49.7	47.1	+5.5%	25.0 × 60.0	^a	
	54.2	53.3	+1.7%	12.0 × 105.0	^a	
	GS		$\theta_{mgs} = 191$ nA/nm		Ref.	
	[nm]	[A/m]	σ_r	$L \times W$ [nm ²]		
NSFET	3	15.8	15.2	+3.9%	12.0 × 110.0	[13]
		10.5	11.4	-7.9%	18.0 × 165.3	^a
	5	26.3	25.0	+5.2%	12.0 × 110.0	[13]
		17.5	16.5	+6.1%	18.0 × 165.3	^a
	7	36.8	36.5	+1.0%	12.0 × 110.0	[13]
		24.5	24.8	-1.2%	18.0 × 165.3	^a
10	52.6	52.3	+0.7%	12.0 × 110.0	[13]	
	35.0	32.8	+6.7%	18.0 × 165.3	^a	

^aThe simulations done for this work are referenced.

Thus, the PBP model for LER will be as follows:

$$\sigma I_{on} = A_{ler} \frac{1}{\sqrt{LW}} = \theta_{ler} \cdot \Delta \sqrt{\frac{f(A) \cdot w}{LW}}, \quad (4)$$

where θ_{ler} is the on-current mismatch for LER, a parameter that states the dependence of variability due to an architecture, and once is fixed, the estimation of σI_{on} could be done at any dimensions (L, W, w, Δ , or A).

The PBP model plots are shown in Fig. 3(a)–(b) for NWFET and FinFET due to LER-induced variability, respectively. Also, in Fig. 3(c) we show a comparison between the simulated and predicted σI_{on} due to LER variability, together with the shaded region where the deviations are lower than 10% from the expected value for the 22 devices studied. The NSFET is not studied for LER variability since, the impact on I_{on} is negligible for this architecture because the roughness due to the etching processes is on the non-critical dimension, this could be seen in [1]. We can see a good match between the predicted and the simulated data for different values of Δ , A , different dimensions, and architectures.

2.3. Computational cost

Once the model's accuracy has been tested, we present another advantage of the PBP model, the reduction of computational cost. In this section, we compare the computational cost of performing a

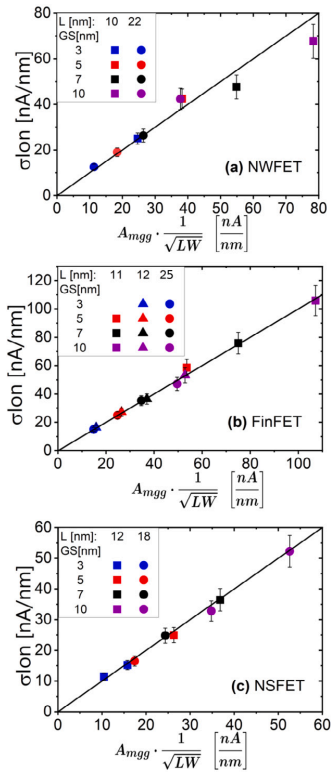


Fig. 2. Pelgrim-based predictive (PBP) plots for MGG variability for three state-of-the-art different architectures: (a) NWFET, (b) FinFET, and (c) NSFET. The gate length L of the devices, together with the GS are also shown.

variability study through TCAD simulations versus the calibration and application of the PBP model.

To calibrate the model for MGG (LER), we only need the simulation of the sets (300 simulations per set) for an architecture at four different GS (Δ or Λ) values. The computational time for a single simulation of a 10 nm NWFET affected by MGG or LER variability with QC MC methodology using VENDES software on Intel^(R) Core^(TM) i9-10850 K CPU at 3.60 GHz with a memory DDR4 with 3200 MT/s is around 25 h. Fig. 4 shows a comparison of the computational time required for a QC MC TCAD study for several dimensional devices with a common architecture (NWFET) affected by MGG variability. Also, the calibration time for the PBP model is shown. Note that this time matches the computational cost of one TCAD simulation variability study. Once the PBP model is calibrated, the computational time on estimating the impact of variability on new devices does not increase. As can be seen, the reduction in computational time by using the PBP model is significant as the number of devices increases. Therefore, this model can help to predict the impact of future technology nodes with a reduced computational cost.

3. Conclusions

In conclusion, we have presented a new comprehensive on-current variability prediction model for MGG and LER based on Pelgrim's Law.

PBP model has been tested for three state-of-the-art architectures, with prediction errors lower than 8% in the 92% of the cases. The application of the PBP model implies a huge reduction of the computational

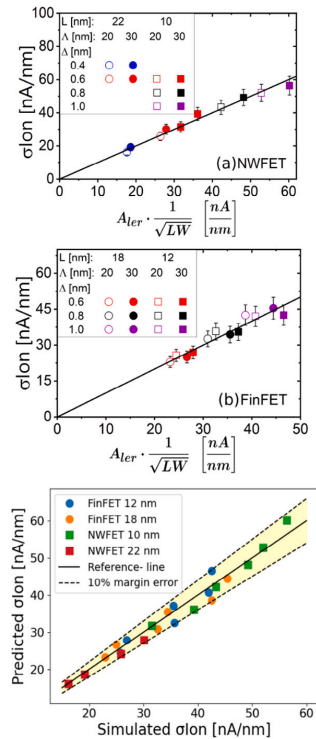


Fig. 3. PBP plots for LER variability impact on $\sigma_{I_{on}}$ for two state-of-the-art different architectures: (a) NWFET, (b) FinFET. Also, in (c) is shown a comparison between data predicted with the PBP model and the simulated data due to LER. The gate length L of the devices, together with the variability parameters for LER (Δ and Λ), are also shown.

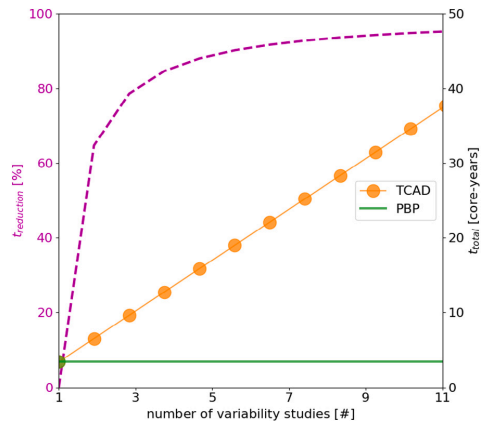


Fig. 4. Computational time in the estimation of the on-current standard deviation ($\sigma_{I_{on}}$) due to MGG versus the number of variability studies. The simulation time of QC MC TCAD is compared to that of the PBP model. Each TCAD point of the graph corresponds to a complete variability study (four GS values) for several devices. The reduction in computational cost derived from the application of the PBP model compared to TCAD is also shown.

time in comparison with QC MC TCAD variability studies. Also, the power of the model allows us to predict the impact of MGG and LER on any device for a calibrated architecture with no further simulations.

Hence, the PBP model is a simple, fast, and reliable strategy to estimate the impact on the transistor performance of a certain source of variability. This model could be useful to predict the impact of variability on future technology nodes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Spanish MICINN, Xunta de Galicia, and FEDER Funds under Grant RYC-2017-23312, Grant PID2019-104834GB-I00, Grant ED431F 2020/008, and Grant ED431C 2022/16.

References

- [1] Seoane N, Fernandez JG, Kalna K, Comesaña E, García-Loureiro A. IEEE Electron Dev Lett 2021;42(10):1416–9. <http://dx.doi.org/10.1109/LED.2021.3109586>.
- [2] Carrillo-Núñez H, Dimitrova N, Asenov A, Georgiev V. IEEE Electron Dev Lett 2019;40(9):1366–9. <http://dx.doi.org/10.1109/LED.2019.2931839>.
- [3] Indalecio G, Seoane N, Kalna K, García Loureiro A. IEEE Trans Electron Dev 2017;64:1–7. <http://dx.doi.org/10.1109/TED.2017.2670060>.
- [4] Pelgrom M, Duinmaijer A, Welbers A. IEEE J Solid-State Circuits 1989;24(5):1433–9. <http://dx.doi.org/10.1109/JSSC.1989.572629>.
- [5] Sheikholeslami A. IEEE Solid-State Circuits Mag 2015;7(1):8–9. <http://dx.doi.org/10.1109/MSSC.2014.2369331>.
- [6] Lu N, Brown JS, Thoma R, Kotecha PM, Wachnik RA. IEEE IEDM 2014;35.3.1–4. <http://dx.doi.org/10.1109/IEDM.2014.7047174>.
- [7] Stolk P, Widdershoven F, Klaassen D. IEEE Trans Electron Dev 1998;45(9):1960–71. <http://dx.doi.org/10.1109/16.711362>.
- [8] Nayak K, Agarwal S, Bajaj M, Oldiges PJ, Murali KVRM, Rao VR. IEEE Trans Electron Dev 2014;61(11):3892–5. <http://dx.doi.org/10.1109/TED.2014.2351401>.
- [9] Sudarsanan A, Venkateswarlu S, Nayak K. IEEE Trans Electron Dev 2019;66(11):4646–52. <http://dx.doi.org/10.1109/TED.2019.2941896>.
- [10] Seoane N, Indalecio G, Aldegunde M, Nagy D, Elmessary MA, García-Loureiro AJ, Kalna K. IEEE Trans Electron Dev 2016;63(3):1209–16. <http://dx.doi.org/10.1109/TED.2016.2516921>.
- [11] Sung W-L, Yang Y-S, Li Y. IEEE J Electron Dev Soc 2021;9:151–9. <http://dx.doi.org/10.1109/JEDS.2020.3046608>.
- [12] Nagy D, Indalecio G, García-Loureiro AJ, Elmessary MA, Kalna K, Seoane N. IEEE J Electron Dev Soc 2018;6:332–40. <http://dx.doi.org/10.1109/JEDS.2018.2804383>.
- [13] Nagy D, Espiñeira G, Indalecio G, García-Loureiro AJ, Kalna K, Seoane N. IEEE Access 2020;8:53196–202. <http://dx.doi.org/10.1109/ACCESS.2020.2980925>.
- [14] Seoane N, Nagy D, Indalecio G, Espiñeira G, Kalna K, García-Loureiro A. Materials 2019;12(15). <http://dx.doi.org/10.3390/ma12152391>.
- [15] Indalecio G, García-Loureiro AJ, Seoane Iglesias N, Kalna K. IEEE Trans Electron Devices 2016;63(6):2625–8. <http://dx.doi.org/10.1109/TED.2016.2556749>.
- [16] Dadgour H, Endo K, De V, Banerjee K. Ieee iedm. 2008, p. 1–4. <http://dx.doi.org/10.1109/IEDM.2008.4796792>.
- [17] Fernandez JG, Seoane N, Comesaña E, Kalna K, García-Loureiro A. SISPAD 2021;201–5. <http://dx.doi.org/10.1109/SISPAD54002.2021.9592600>.
- [18] Asenov A, Kaya S, Brown A. IEEE Trans Electron Dev 2003;50(5):1254–60. <http://dx.doi.org/10.1109/TED.2003.813457>.

A.2. Trabajos complementarios relacionados con esta tesis

A.2.1. Congresos internacionales

A comprehensive Pelgrom-based predictive model to estimate metal grain granularity and line edge roughness in advanced multigate MOSFETs

- Authors: J. G. Fernandez, N. Seoane, E. Comesaña, A. García-Loureiro
- Conference: 2022 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)

Impact of metal grain granularity on three gate-all-around advanced architectures

- Authors: J. G. Fernandez, N. Seoane, E. Comesaña, K. Kalna and A. García-Loureiro
- Conference: 2021 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)
- <https://doi.org/10.1109/SISPAD54002.2021.9592600>

A.2.2. Congresos nacionales

Threshold voltage variability study in a 12 nm gate length Nanosheet FET

- Authors: J. G. Fernandez, N. Seoane, K. Kalna and A. Garcia-Loureiro
- Conference: 13th Spanish Conference on Electron Devices (CDE) 2021

Analysis of the statistical impact of variability in a 12 nm nanosheet FET

- Authors: J. G. Fernandez, N. Seoane, E. Comesaña, A. García-Loureiro
- Conference: 14th Spanish Conference on Electron Devices (CDE) 2023

Índice de figuras

2.1.	Esquema de la metodología aplicada.	24
2.2.	Ejemplo de modelado y mallado por elementos finitos de un NSFET.	25
2.3.	Esquema del proceso y funcionalidades del simulador VENDES.	27
2.4.	Esquema de las diferentes fuentes de variabilidad estudiadas en este trabajo, aplicadas sobre un NWFET. (a) Muestra los granos en la puerta metálica que se corresponden a las diferentes funciones de trabajo. En (b) se observa la rugosidad superficial del dispositivo debida a los procesos litográficos. (c) Muestra la rugosidad producida sobre la puerta, haciendo que el contorno de esta sea no-uniforme. (d) Corte transversal del NWFET dónde se puede observar el efecto de la aparición de dopantes aleatorios sobre el potencial interno del dispositivo.	30
2.5.	Láminas transversales de un NWFET dónde se resuelve la ecuación 2D de Schrödinger bajo una polarización en el drenador (V_D) de 0.1 V.	32
2.6.	Esquema y funcionalidades del MLFoMpy.	34
3.1.	Esquema de las diferentes arquitecturas 3D modeladas siendo: (a) NSFET, (b) FinFET y (c) NWFET. También se muestra una tabla con los valores de los dispositivos simulados en (d) siendo: L_g la longitud de puerta, $L_{s/d}$ la longitud de fuente-drenador, L longitud total del dispositivo, w_{ch} ancho del canal, h_{ch} alto del canal, t_{ox} anchura del óxido efectiva, P_s el perímetro del semiconductor, $N_{s/d}$ el dopado tipo-n en fuente-drenador, X_m la posición de decaimiento del dopado gaussiano tipo-n y δ la forma en la que decae el dopado gaussiano.	39

3.2.	Perfil de dopado gaussiano de un NSFET con longitud de puerta de 12 nm. N es la concentración de dopado tipo-n en el dispositivo. $N_{s/d}$ es el máximo dopado en la fuente y drenador, X_m la posición de decaimiento del dopado gaussiano cercana a las fronteras fuente/canal y canal/drenado, y δ la forma con la que decae el perfil de dopado.	40
3.3.	Características I_D - V_G de los diferentes dispositivos calibrados para este trabajo a una tensión de drenador V_D de 0.7 V: (a) 12 nm NSFET, (b) 12 nm FinFET y (c) 12 nm NWFET. En las leyendas de las figuras se muestra la combinación método de transporte/corrección cuántica, siendo DD arrastre difusión, MC Monte Carlo, DG gradiente de densidad y SCH la ecuación de Schrödinger. Los datos experimentales se denotan como EXP.	41
3.4.	Métodos de extracción de figuras de mérito empleado en el desarrollo de este trabajo aplicados sobre la curva del NSFET con longitud de puerta 12 nm. . .	43
3.5.	Desviaciones estándar de (a) la tensión umbral (σV_{Th}) y de (b) la pendiente subumbral (σSS) debidas a la variabilidad MGG frente al GS para NWFET, FinFET y NSFET.	44
3.6.	Desviaciones estándar de (a) la tensión umbral (σV_{Th}) y (b) la pendiente subumbral (σSS) debidas al impacto del LER frente a la raíz cuadrática media de la deformación (RMS) para una longitud de correlación (CL) de 20 nm para NWFET, FinFET y NSFET.	45
3.7.	(a) σV_{Th} debida al impacto de GER frente al RMS (CL=20 nm) y (b) comparativa del impacto de las diferentes fuentes de variabilidad sobre σV_{Th} para las tres arquitecturas (GS=3 nm y RMS=0.6 nm).	46
3.8.	Desviación estándar de la corriente de encendido (σI_{on}) debido a las fuentes de variabilidad (a) MGG y (b) LER en función del GS y del RMS, respectivamente, para NWFET, FinFET y NSFET.	48
3.9.	Sección transversal de las tres arquitecturas usadas en este trabajo. Los FinFETs: (a) de Wang [29] y Leung [48], (b) de Seoane [44], Indalecio [14] y (c) de Seoane [20]. Los NWFETs: (d) de Lü [45], Sung [47] y (e) de Nagy [46] y Elmessary [19]. Los NSFETs: (f) de Nagy [18] y simulados para este trabajo. w_{ch} y h_{ch} son el ancho y alto del canal respectivamente. Denotamos como to_h la altura del óxido en la cúpula del FinFET tipo-(b).	50

3.10. Desviación estándar del voltaje umbral σV_{Th} frente al GS para la variabilidad MGG comparando NWFETs, FinFETs y NSFET con diferentes longitudes de puerta. 51

3.11. Gráficas de Pelgrom evaluando la σV_{Th} debida al impacto de MGG con diferentes GS para: (a) FinFETs extraídos de Seoane16 [44], Seoane21 [20], Indalecio [14], Wang [29], (b) NWFETs reportados por Nagy [46], Sung [47], Lü [45], y (c) NSFETs de Nagy [18] conjuntamente con simulaciones propias realizadas para este estudio. Obsérvese cómo los dispositivos siguen con exactitud la Ley de Pelgrom, además el coeficiente Pelgrom de proporcionalidad para V_{Th} (A_v) aumenta linealmente con el GS. 53

3.12. (a) σV_{Th} frente a RMS debido a la variabilidad LER para NWFET y FinFET con diferentes longitudes de puerta, siendo CL es la longitud de correlación. Nótese la relación lineal entre σV_{Th} y RMS. (b) σV_{Th} frente a la relación entre la longitud de correlación y la longitud de puerta (CL/L_g) para dos arquitecturas: FinFET [14] y NWFET [19]. También se muestra el valor RMS para cada arquitectura. Las líneas continuas representan el ajuste de cada dispositivo a su propia $f(CL)$ y las barras de error y corresponden al 10 % de cada valor de σV_{Th} . Obsérvese el buen ajuste de $f(CL)$ para los dos dispositivos presentados. Las simulaciones se realizaron con un V_D de 0.7 V. 57

3.13. σV_{Th} debidos al impacto del LER frente a $RMS \cdot w_{ch}^{-1/2} \cdot f(CL) \cdot (L_g P_{sc})^{-1/2}$ a diferentes valores de V_D , para: (a) FinFETs, (b) NWFETs. Datos extraídos de Seoane16 [44], Seoane21 [20], Indalecio [14], Wang [29], Leung [48] y Elmessary [19]. Obsérvese el excelente ajuste del modelo con los datos simulados y el aumento de θ_v^{ler} con V_D 58

3.14. σI_{on} simulada en VENDES (σI_{on}^S) frente a la σI_{on} predicha por el modelo PBP (σI_{on}^P), para diferentes arquitecturas (a) NWFET, (b) FinFET y (c) NSFET, con diferentes longitudes de puerta (L_g) e impactados por diferentes GS. La línea sólida negra es la bisectriz del primer cuadrante e indica correlación perfecta entre los valores predichos con el modelo y simulados. Las barras de error corresponden al 10 % de los valores simulados, margen que consideramos aceptable. 62

- 3.15. σI_{on} simulada en VENDES (σI_{on}^S) frente a la σI_{on} predicha por el modelo PBP para LER (σI_{on}^P), para dispositivos NWFET y FinFET con diferentes L_g y diferentes parámetros de deformación RMS y CL. La línea sólida negra es la bisectriz del primer cuadrante e indica correlación perfecta entre los valores predichos con el modelo y simulados. Las barras de error corresponden al 10 % de los valores simulados, margen que consideramos aceptable. 64
- 3.16. Tiempo computacional en la estimación σI_{on} debida a MGG frente al número de estudios de variabilidad. Definiendo como estudio de variabilidad la realización de las simulaciones para un dispositivo específico (L_g, P_g) a cuatro valores de GS diferentes. El tiempo de simulación de QC MC (TCAD) se compara con el necesario para calibrar el modelo PBP, para lo cual solo se necesita un estudio de variabilidad. Cada punto TCAD del gráfico corresponde a un estudio de variabilidad completo. En el eje Y derecho se muestra el tiempo total de simulación y el eje Y izquierdo muestra la reducción de tiempo computacional debido a la aplicación del modelo PBP en la región de encendido en comparación con el TCAD. 67

Índice de tablas

3.1.	Comparación de la σV_{Th} predicha por el modelo PBP (σV_{Th}^P), frente a la σV_{Th} procedente de estudios de simulación (σV_{Th}^S) disponibles en la literatura. También se muestran el GS, el error relativo entre los resultados predichos y simulados (σ_r), el área efectiva de puerta ($L_g \times P_g$) y θ_v^{mgs} . [*] indica los datos simulados específicamente para este estudio. Los σ_r mayores al 10 % están marcados en negrita.	54
3.2.	Validación del modelo PBP para LER aplicado sobre el FinFET para la estimación de σV_{Th} . V_D es el potencial de drenaje, θ_v^{ler} el desfase del voltaje umbral para la variabilidad LER y $L_g \times P_{sc}$ el área del semiconductor. También se muestran CL, RMS, los valores de σV_{Th} predichos σV_{Th}^P y simulados σV_{Th}^S junto con su error relativo σ_r	59
3.3.	Validación del modelo PBP para LER aplicado sobre el NWFET para la estimación de σV_{Th} . V_D es el potencial de drenaje, θ_v^{ler} el desfase del voltaje umbral para la variabilidad LER, la longitud de correlación CL, $L_g P_{sc}$ el área del semiconductor y la raíz cuadrática media de la deformación RMS. También se han calculado los valores predichos σV_{Th}^P y la tensión umbral simulada σV_{Th}^S junto con su error relativo σ_r	60
3.4.	σI_{on} predichas (σI_{on}^P) y simuladas (σI_{on}^S), a diferentes valores de GS para cada arquitectura. En la tabla también listamos el desfase de corriente para MGG θ_i^{mgs} , el área efectiva de la puerta ($L_g \times P_g$) y el error relativo σ_r . Las simulaciones realizadas específicamente para este trabajo se denotan con *. Los σ_r mayores al 10 % están marcados en negrita.	63

- 3.5. σI_{on} predicha (σI_{on}^P) y simulada (σI_{on}^S), a diferentes valores de RMS y CL para las arquitecturas NWFET y FinFET. En la tabla también listamos el desfase de corriente para LER θ_i^{ler} , el área efectiva del semiconductor ($L_g \times P_{sc}$) y el error relativo σ_r . Las simulaciones realizadas específicamente para este trabajo se denotan con * 65

Siglas

2D-SCH ecuación 2D de Schrödinger 3, 4, 31, 40, 42

CL *Correlation Length* 4, 6, 8, 12, 13, 15, 28, 29, 44–47, 55–57, 59, 60, 63–65, 71, 77, 108–112

DD *Drift-diffusion* 3, 5, 7, 11, 12, 15, 32, 34, 40–42, 47, 49, 50, 52, 56, 60, 108

DG *Density Gradient* 3, 31, 32, 40, 41, 108

FET *Field Effect Transistor* 1, 9, 18, 49, 72, 78

FoM *Figures of Merit* 5, 13, 34, 35, 37, 42, 47

GER *Gate Edge Roughness* 4, 6, 11, 12, 14, 26, 28, 30, 42, 45–47, 69, 70, 75, 76, 108

GS *Grain Size* 4–8, 12, 13, 15, 16, 26, 43, 44, 46–48, 51–55, 61–63, 65–67, 69, 71, 75, 77, 108–111

IRDS *IEEE International Roadmap for Devices and Systems* 1, 4, 9, 12, 18, 24, 37, 38, 50

LER *Line Edge Roughness* 4, 6–8, 11–15, 26, 28, 30, 42, 44–49, 55–61, 63–65, 69–73, 75–78, 108–112

MC Monte Carlo 3, 5, 7, 11, 12, 15, 32, 33, 40–42, 47, 49, 60, 61, 65–67, 108, 110

MGG *Metal Grain Granularity* 4–8, 11–16, 26, 30, 42–49, 51–53, 55, 56, 61, 63, 65–67, 69–73, 75–78, 108–111

ML *Machine Learning* 20

MLFoMpy *Machine Learning Figure of Merit Python extraction tool* 2, 10, 21, 23, 34, 42, 107

MOS *Metal Oxide Semiconductor* 18, 49

NEGF *Non-Equilibrium Green's Functions* 4, 11, 26, 31, 33

NSFET *nanosheet FET* 1, 2, 4–7, 9, 10, 12–14, 18, 20, 24, 25, 37–48, 50–54, 56, 62–64, 69, 70, 75, 76, 107–109

NWFET *nanowire FET* 1, 2, 4–10, 12–15, 18, 20, 24, 28, 30, 32, 37–41, 44–48, 50–58, 60, 62–66, 69, 70, 75, 76, 107–112

OTV *Oxide Thickness Variations* 72, 78

PBP *Pelgrom-based predictive model* 7, 8, 15, 16, 49–52, 54–56, 59–67, 71, 72, 76–78, 109–111

PUFs *Physical Unclonable Functions* 37

QC *Quantum Corrections* 3, 7, 31, 47, 49, 50, 52, 56, 60, 61, 65–67, 110

RDD *Random Discret Dopants* 4, 6, 11, 12, 14, 26, 29, 30, 42, 47, 69, 70, 75, 76

RMS *Root Mean Square height* 4, 6, 8, 12, 13, 15, 28, 29, 44–48, 55–57, 59, 60, 63–65, 69–71, 75–77, 108–112

SEM *Scanning Electron Microscope* 28, 44

TCAD *Technology Computer Aided Design* 2, 7–10, 15, 16, 19–21, 23, 26, 37, 65–67, 70–73, 75–78, 110

VENDES *Variability Enabled Nanometric DEvice Simulator* 3, 7, 11, 15, 23, 25–28, 31–33, 49, 50, 61, 62, 64, 65, 72, 78, 107, 109, 110



En esta tesis se abordan dos desafíos clave en el diseño de transistores para futuros nodos tecnológicos. El primero identificar la arquitectura basada en silicio más prometedora para reemplazar a los FinFETs, analizando el impacto de diferentes fuentes de variabilidad. Los resultados demuestran que el nanosheet FET exhibe una notable resistencia al impacto de las variabilidades estudiadas, posicionándolo como una alternativa adecuada para sustituir al FinFET. El segundo es reducir al máximo la demanda de recursos computacionales, manteniendo una elevada precisión en los resultados, para ello se desarrollaron modelos empíricos. Estos modelos presentan errores relativos inferiores al 10% reduciendo de manera sustancial el tiempo de computación necesario para la realización de estos estudios y su impacto ambiental.