UNIVERSIDAD DE GRANADA

Departamento de Electrónica
y Tecnología de Computadores

TESIS DOCTORAL

Study and Simulation of Advanced Si–based Nanodevices: Schottky–Barrier MOSFETs and Tunnel FETs

JOSÉ LUIS PADILLA DE LA TORRE

Granada, 2012

Editor: Universidad de Granada. Tesis Doctorales
Autor: José Luis Padilla de la Torre
ISBN: 978-84-9125-826-1
URI: http://hdl.handle.net/10481/43553

A todos los que la presente vieren y entendieren: sabed [...]

FRANCISCO JESÚS GÁMIZ PÉREZ y **ANDRÉS GODOY MEDINA**, Catedráticos de Universidad del Área de Electrónica de la Universidad de Granada,

AUTORIZAN

La presentación de la memoria de investigación titulada "*Study and Simulation of Advanced Si–based Nanodevices: Schottky–Barrier MOSFETs and Tunnel FETs*" realizada por **D. José Luis Padilla de la Torre** para optar al TÍTULO DE DOCTOR CON MENCIÓN INTERNACIONAL por la Universidad de Granada, y que ha sido desarrollada en su totalidad bajo su dirección en el Grupo de Investigación en Nanoelectrónica (TIC–216) del Departamento de Electrónica y Tecnología de Computadores de la Universidad de Granada.

Granada, trece de julio del dos mil doce

Francisco Jesús Gámiz Pérez
Catedrático de Universidad
Universidad de Granada

Andrés Godoy Medina
Catedrático de Universidad
Universidad de Granada

Tesis Doctoral: ***Study and Simulation of Advanced Si–based Nanodevices: Schottky–Barrier MOSFETs and Tunnel FETs***

Autor: **D. José Luis Padilla de la Torre**
Licenciado en Física

Directores: **D. Francisco Jesús Gámiz Pérez**
Doctor en Física
Catedrático de Universidad

D. Andrés Godoy Medina
Doctor en Física
Catedrático de Universidad

Departamento: Electrónica y Tecnología de Computadores
Universidad de Granada

El Tribunal de Calificación queda compuesto por:

PRESIDENTE:
VOCAL SECRETARIO:
VOCALES:

Celebrado el acto de defensa y lectura de la Tesis en Granada a día ..... de ..... del año ..... en la Facultad de Ciencias de la Universidad de Granada.
El tribunal acuerda otorgarle la calificación de:

A quienes se fueron,
a quienes lo son todo,
a quienes están por llegar.

Resistir sin esperanza es la suprema dignidad del ser humano.
José Saramago

Agradecimientos

A lo largo del complejo camino de maduración personal y profesional que supone la elaboración de una tesis, hay personas que resultan determinantes y sin las cuales el mosaico final que conforma la imagen global del conjunto sería irremediablemente diferente. El reconocimiento y la gratitud que estas líneas pretenden atestiguar sólo suponen una pequeña parte de todo lo que ha significado compartir esta etapa vital con quienes han formado parte de ella.

En primer lugar, a mis directores. Verdaderos guías que me han ido mostrando el camino e inestimables compañeros de viaje a lo largo de él. Francisco Gámiz, referente permanente y espejo en el que encontrar siempre el reflejo de una concepción de la ciencia basada en el trabajo duro y bien hecho. Por mostrarme que la perseverancia fruto de la convicción en lo que hacemos acaba siendo reconocida y valorada. Siempre me sentí respaldado por ti y apoyado especialmente en los momentos menos gratos. Gracias por confiar en mí. Andrés Godoy, un ejemplo constante del gusto por la perfección. Siempre enseñándome y planteándome la pregunta adecuada en el momento justo. Haciéndome dudar, y haciéndome sentir que siempre era posible mejorar. Contigo compartí —y espero seguir haciéndolo— el gusto por correr. Nunca conseguí ganarte, pero incluso en eso, me sirves permanentemente de modelo.

A Isabel Tienda, porque nunca olvidaré a través de quién tuve la oportunidad de conocer y vincularme al grupo de nanoelectrónica. Estuviste en el origen de esta travesía que ahora llega a puerto, y por ello mi gratitud va mucho más allá de lo que pueden expresar estas pala-

bras. Has conocido la historia desde sus orígenes y comprenderás el lugar destacado que para mi ocupas formando parte de todo esto.

A Pablo Sánchez, compañero, amigo y partícipe de innumerables correrías académicas, personales... e incluso cinematográficas. Permanentemente con una perspectiva sincera, y con una palabra de ánimo lista en la recámara. Me siento afortunado de tenerte cerca y agradecido por dejarme compartir tu visión de la vida.

A Elena Ocaña. Por tantísimos momentos, por tantísimo cariño, por tu amistad inquebrantable. Porque estuvieras donde estuvieras, por lejos que fuera, siempre te he sentido a mi lado. Hemos reído juntos, llorado juntos, y compartido infinidad de cosas con las que atesoro en mí una pequeña parte de tu humanidad.

A Rosalía, por tu contrapunto de entrañable excentricidad con la que siempre supiste poner buena cara a la adversidad. Tu enorme capacidad y afán de superación te harán salir adelante frente a cualquier obstáculo. Fue un honor compartir contigo el inicio del camino. A Consuelo, por tus ideas e integridad. A Dani Guirado, por demostrar que el listón de la locura siempre se puede poner un poco más allá. A Dani Partal, Isa y Juanito, la entrañable familia que se encarga de poner el toque de cordura en los viajes.

A Antonio Palomares, por tu ejemplo de integridad y dedicación docente, por tu visión pragmática a la hora de enfrentar cualquier problema, por saber escuchar, por indignarte conmigo, por echar una mano siempre que hiciera falta, por las infinitas cenas viendo House... y por mil y un motivos que nunca terminaría de enumerar.

A Jorge de Blas, por la camaradería que compartimos y por tu enorme bondad. Llegarás muy lejos porque cuando la luz brilla en la oscuridad, la oscuridad desaparece. A Mari Carmen Carmona, por tu ejemplo de trabajo metódico y porque no soy capaz de recordar una sola vez en que no sonrieras, pasara lo que pasara.

A Pedro y Laura, por vuestra hospitalidad, espontaneidad, y por dejar siempre para otro día el visitar la catedral de Málaga. A Laura y Pablo, que desde la distancia también forman parte del grupo que —aunque en otra rama de la ciencia— compartió la aventura de embarcarse

en una tesis. A Ian, por ser el contrapunto matemático en medio de tanto biólogo y por hacerme sentir en casa cada vez que estoy en Málaga. A Vero, porque sabes todo lo que me has dado y porque haces que intente estar a tu altura siquiera para expresarte la enorme gratitud que supone encontrar a alguien tan especial como tú. Confío en que seamos compañeros atravesando la ignota senda que se abre ante nosotros tras la tesis.

A Enrique, compañero de despacho en espíritu desde tu destierro en el otro pasillo. A Abraham, defenderemos nuestras tesis con un pequeño intervalo entre ambas y, por eso mismo, me he sentido reflejado en ti y animado por tu ejemplo todo este tiempo compartido. A Pilar, por no rendirte y por tu enorme capacidad de trabajo. A Karam, por el contrapeso cultural que das al despacho. A Trinidad, porque los puntos de vista diferentes enriquecen y porque los nuestros no han impedido que fragüemos una sólida y sincera amistad. A Celso, prohombre bastetano, por esa especial conexión que da compartir un sentido del humor como el nuestro, que hace memorables las incontables horas juntos. Siempre decías que habías llegado al grupo de casualidad, ojalá las casualidades trajeran siempre a gente como tú. Y a Nika, sencillamente por aparecer... por haber llegado a formar parte tan importante de todo. Por nuestra especialísima complicidad, por tu franqueza, por tu confianza, por hacerme reír con la tontería más impensable, por sentirte al lado en cualquier situación y por compartir los mejores y los peores momentos.

A Fran y Luca, fuisteis mis primeros compañeros de despacho cuando empecé y difícilmente se puede tener más suerte por ello. A Carlos, por estar ahí siempre que necesité tu ayuda o pedirte cualquier cosa. A Noel, por ser y dar un gran ejemplo a quienes te rodean. Y a todos los miembros del Departamento de Electrónica con los que, de una manera u otra, he compartido parte de estos años.

Finalmente, a mis hermanos y mis padres. Es para vosotros la más emocionada expresión de mi gratitud, admiración y sentimientos, que siento no ser capaz de expresar con palabras. Siempre resultarían insuficientes y nunca contendrían todo lo que querría poner en ellas. Os reservo el espacio íntimo donde las emociones se comparten y donde todo

se sabe sin ser dicho. Lo sois todo y lo seguiréis siendo cuando ya nada sea.

Abstract

The aim of the work herein presented in this thesis is to deepen the simulation study of devices based of new injection mechanisms, which are currently regarded as potentially interesting to replace conventional MOSFETs and overcome their fundamental subthreshold swing limitation of 60mV/dec. The physical impossibility of breaking this limit fosters most of the ongoing research precisely in the direction of exploring novel devices such as those considered in this work: the Tunneling Field–Effect Transistors (TFETs) and the Schottky Barrier MOSFETs (SB–MOSFETs). For each one of the considered devices, and assuming the existing State–of–the–Art as starting point, we have structured and organized the performed work according to the following roadmap:

(i) Exposition of the physical phenomena and mechanisms to be studied for a precise understanding and description of the devices.

(ii) Identification of the existing limitations, incompatibilities and problems that arise during simulation processes. In our case, using Silvaco ATLAS.

(iii) Development of simple simulation algorithms that allow to overcome the encountered difficulties and keep using this widely employed commercial simulator.

(iv) Presentation of simulation results obtained from the application of these proposed simulating approaches.

The scope and orientation of the work in the case of SB–MOSFETs was set by the fact that before performing it, we had experimental results

that appropriately aimed our efforts at fitting them. The main problem that we found studying these devices was that barrier lowering processes were not completely implemented in ATLAS when applied to carrier injection mechanisms involving tunneling (field emission and thermionic field emission). Considering the relevant impact that small variations in barrier heights may have on the total current, and taking into account that depending on the bias conditions the relative importance of the different injection mechanisms changes, it becomes essential to suitably include those barrier lowering processes in our simulations. For that purpose, in this work we developed an iterative procedure inside ATLAS to account for barrier lowering (which also applies for tunneling processes), making it vertically dependent on the depth inside the channel. Very accurate fits between experimental results and simulations have been obtained especially for those regions where tunneling processes proved to be dominant. In addition, some short channel effects like the observed current reduction when decreasing the channel length due to the overlap of the potential profiles of the Schottky barriers are also satisfactorily reproduced.

In the case of TFETs, as most of the existing research on them still involves semiclassical approaches and considering their progressive reduction in size, we wanted to take a step forward by somehow performing a more complete treatment which needed to include the effect of quantum confinement. The necessity of such an approach was obvious regarding that whenever the presence of confinement is significant, the existence of a discrete sprectum of energy levels replacing the formerly continuous conduction and valence bands should be greatly affecting the so–called band–to–band tunneling injection of carriers. In that context, the work developed in this thesis lies between those semiclassical models not accounting for the effects of confinement, and the more recent approaches involving rigorous quantum mechanical treatments. The inclusion of confinement made us realize that the numerical solvers employed by ATLAS when using the non–local band–to–band tunneling model (to inject carriers), and the self consistent Schrödinger–Poisson model (to account for subband quantization) were not compatible. At

that stage, we decided to exploit the capabilities of the simulator by designing an iterative approach that reasonably allowed to account for confinement in a way that offers great possibilities for researchers that may be potentially interested in the study of these devices. Thanks to the development of this approach we have been able to analyze the impact that confinement indeed has over the underlying physics in TFETs, and how it modifies their total current levels or affects the global trends of electrical parameters of utmost importance for their characterization (threshold voltages and subthreshold swings).

Contents

List of Abbreviations and Symbols

Abbreviations

ATM	Airy–transfer–matrix
BL	Barrier lowering
BOX	Buried oxide (isolation layer of a SOI structure)
BTBT	Band–to–band tunneling
CMOS	Complementary MOS technology
CNT	Carbon nanotube
DG	Double gate
DIBL	Drain induced barrier lowering
DL	Dipole lowering
FET	Field–effect transistor
GIDL	Gate induced drain leakage
IFL	Image force lowering
MIGS	Metal induced gap states
MOS	Metal–oxide–semiconductor
MOSFET	Metal–oxide–semiconductor field–effect transistor
PMOS	p–channel MOSFET
RTA	Rapid Thermal Annealing
SB	Schottky barrier
SBH	Schottky barrier height
SB–MOSFET	Schottky barrier MOSFET
SB–NMOS	n–channel SB–MOSFET

SCE	Short Channel Effects
SDE	Source/drain extension
SEM	Scanning electron microscopy
SG	Single gate
S/D	Source/drain
SOI	Silicon on insulator
SS	Subthreshold swing
TCAD	Technology computer aided design
TFET	Tunneling field–effect transistor
UST	Universal Schottky tunneling
UTB	Ultra–thin body
WKB	Wenzel–Kramers–Brillouin approximation
XTEM	Cross–sectional transmission electron microscopy

Symbols

A^*	Richardson's constant
C_{ox}	Oxide capacitance
C_q	Quantum capacitance (semiconductor capacitance)
dn	Electron density in an incremental energy range
δ	Interfacial layer length
Δ	Potential drop across the interfacial layer
$\Delta\phi$	Barrier lowering potential
$\Delta\phi_{dl}$	DL Barrier lowering potential
$\Delta\phi_{ifl}$	IFL Barrier lowering potential
$\Delta\Phi$	Available energy range for BTBT
D_{it}	Interface trap density
E_m	Energy level above E_F at which thermionic emission is maximum
D_n	Diffusion coefficient
E_C	Conduction band edge energy
$E_{Cn(p)}$	n(p)–side conduction band energy
E_i	Intrinsic Fermi level energy
E_F	Fermi level energy
E_{Fm}	Fermi level energy in the metal
E_i	Conduction band bound energy
E_ν	Valence band bound energy
E_g	Bandgap energy
E_p	Phonon energy
E_{00}	Reduced energy
$E_\parallel$	Longitudinal energy
E_T	Transverse energy
E_V	Valence band edge energy
$E_{Vn(p)}$	n(p)–side valence band energy
ε_0	Vacuum permittivity
ε_{ox}	Oxide permittivity
ε_s	Semiconductor permittivity
ε_i	Interfacial layer permittivity
$f_{s(m)}$	Fermi distribution function in the semiconductor(metal)

F	Electric field
F_m	Electric field maximum
G	Generation rate function per unit volume
ϕ_m	Metal workfunction
ϕ_n	Potential difference between E_C and E_F
ϕ_p	Potential difference between E_V and E_F
ϕ_{bn}	Schottky barrier for electrons
ϕ_{bp}	Schottky barrier for holes
ϕ_0	Neutral energy level above E_V
ϕ_s	Surface channel potential in TFETs
$\tilde{\phi}_s$	Quantum surface channel potential in TFETs
$\psi_{Bn(p)}$	Potential difference for n(p)–type semiconductors inside the bulk between the Fermi level and the midgap
$\psi_{n(p)}(x)$	Position dependent potential accounting for band bending inside the semiconductor
ψ_s	Surface potential in SB–MOSFETs ($\psi_{n(p)}(x=0)$)
h	Planck constant ($\hbar$: Reduced Planck constant)
J_{btbt}	BTBT current density
$J_{s(m)\rightarrow m(s)}$	Current density from the semiconductor(metal) into the metal(semiconductor)
J_{TE}	Thermionic emission saturation current
J_D	Diffusion saturation current
$J_{n,te}$	Thermionic emission current density for electrons
$J_{n,tfe}$	Thermionic field–emission current density for electrons
$J_{n,fe}$	Field–emission current density for electrons
$J_{n,d}$	Diffusion current density for electrons
κ	Dielectric constant
k	Boltzmann's constant
$k(x,y)$	Carrier wave vector
λ	Screening length
λ_H	Heine tail length
Λ	Correction potential in the Density Gradient model

L_g	Gate length
m^*	Carrier effective mass
μ_n	Electron mobility
n	Electron concentration
n_{ie}	Effective intrinsic concentration
N_A	Acceptor–type doping concentration
$N_{C(V)}$	Density of states in the conduction(valence) band
N_D	Donor–type doping concentration
p	Hole concentration
q	Carrier charge
Q_M	Surface charge on the metal
Q_{sc}	Space charge
Q_{ss}	Charge at the semiconductor surface
ρ	Charge density
$\rho(E_T)$	2D density of states
S_{av}	Average swing
S_{pt}	Point swing
T	Temperature
$T(E)$	Tunneling probability at a given energy E
t_{ox}	Oxide thickness
t_{Si}	Silicon thickness
U	Potential energy
v_x	Carrier velocity in the x–direction
V	Applied bias
V_{bi}	Built–in potential
V_{DD}	Supply voltage
V_{ds}	Drain–to–source voltage
$V_{ds,sat}$	Drain voltage at which the output characteristic saturates
V_{fb}	Flat–band voltage
V_{gs}	Gate–to–source voltage
V_{inv}	Gate voltage at which inversion layer is formed
V_{on}	Gate voltage at which BTBT starts
V_F	Forward bias
V_R	Reverse bias

V_t	Threshold voltage
$V_{tg(d)}$	Gate(drain) threshold voltage
W	Depletion width (also W_D)
w	BTBT barrier width
w_{min}	Minimum BTBT barrier width
χ	Electron affinity

List of Figures

Part I

Extended Abstract in Spanish

I.1 Introducción

El contenido de esta tesis se enmarca en el estudio y simulación de dos tipos de dispositivos nanoelectrónicos que actualmente se barajan como potenciales candidatos para reemplazar a los transistores MOSFET convencionales en aplicaciones a baja potencia, cuales son los transistores MOSFET de contactos metálicos —llamados Schottky barrier MOSFETs o, abreviadamente, SB–MOSFETs—; y los transistores de efecto campo basados en corriente túnel banda a banda, llamados Tunnel–FETs o, por su acrónimo en inglés, TFETs.

Como es bien sabido, los transistores MOSFET han sido la piedra angular de la industria de semiconductores durante los últimos cuarenta años. Su rendimiento ha ido progresivamente mejorándose a través de un incansable proceso de escalado, combinado con la llamada topología de MOS complementario (CMOS), que los ha llevado a dimensiones en rangos inferiores a los 30nm de longitud de puerta. Este proceso de escalado ha ido obedeciendo durante todo este tiempo unas reglas básicas esenciales cuyo objetivo era mantener constante el campo eléctrico en el interior del canal del transistor [4]. Para ello, se fueron reduciendo gradualmente tensiones y dimensiones de los dispositivos, y en paralelo se fueron incrementando en la misma proporción los dopados. Esta manera de proceder permitió que la potencia consumida por unidad de superficie permaneciera constante mientras se reducía el retraso de los circuitos y aumentaba su densidad.

Sin embargo, en la actualidad este proceso de escalado está alcanzando sus límites por la aparición de dos tipos de limitaciones. Las tecnológicas, que complican paulatinamente la continuación del proceso haciendo necesarias soluciones cada vez más complicadas desde el punto de vista ingenieril. Y las fundamentales, que atañen a límites de concepto puramente físicos que sencillamente no pueden sobrepasarse como es el límite de 60mV/decada para la inversa de la pendiente subumbral[1].

Las limitaciones tecnológicas hacen que comúnmente se hable de que

[1]Que es el cambio que debe aplicarse en el voltaje de puerta para conseguir un incremento de corriente de salida de un orden de magnitud.

se ha entrado en la llamada era de "escalado limitado por materiales" [5], puesto que los materiales que se han venido usando en la fabricación de los MOSFETs se han llevado ya hasta sus extremos de potencial rendimiento. Algunos de estos problemas tecnológicos que se plantean son: los efectos de canal corto (conocidos por sus siglas en inglés, SCE); el incremento de la corriente de pérdidas a través de la puerta; el incremento de las resistencias parásitas de fuente y drenador [6, 7]; o la llamada "crisis de potencia" basada en que si no se escala la tensión aplicada V_{DD} (como posible solución ante la reducción de la corriente cuando el dispositivo entra en modo de conducción), esto conduciría a un incremento de la potencia por unidad de superficie, de modo que la adición de más transistores por chip dispararía la potencia consumida.

Todos estos condicionantes plantean dos grandes líneas de actuación. Por una parte, aquélla basada en la búsqueda soluciones ingenieriles que permitan que los dispositivos se sigan comportando en cierta medida como los de canal largo, aún cuando se continúe con el proceso de escalado. En esta línea encontramos el empleo de aislantes de alta constante dieléctrica, materiales con gran movilidad, o nuevos diseños que permitan un mayor control electrostático de puerta como, por ejemplo, el FinFET [8]. Y una segunda línea, más a largo plazo, que busca una reinvención de los dispositivos desde un punto de vista más fundamental cambiando para ello incluso los mecanismos de inyección y transporte de portadores. Así, asumiendo un cambio en los procesos físicos constituyentes de los dispositivos, se abre también la vía para la superación de las limitaciones fundamentales que mencionábamos antes. Dispositivos de este tipo podrían en principio presentar pendientes subumbrales cuya inversa estuviera por debajo del límite de los 60mV/decada.

Los transistores en cuyo estudio se centra esta tesis se enmarcan dentro de esta segunda línea de actuación, por cuanto incorporan como elementos esenciales para su funcionamiento mecanismos de inyección basados en efecto túnel. Bien sea a través de barreras Schottky en uniones metal–semiconductor, como es el caso de los SB–MOSFETs; o bien sea en procesos de efecto túnel banda a banda como ocurre en los TFETs.

I.2 Planteamiento general

Esta tesis está enfocada en el análisis y simulación de aspectos concretos de los dispósitivos propuestos mediante el uso de las herramientas de simulación existentes en el simulador comercial ATLAS de Silvaco [9]. A lo largo de nuestro trabajo, prácticamente la totalidad de las simulaciones fueron desarrolladas con la versión 5.18.3.R, y una pequeña parte con la 5.17.47.C en las primeras etapas de obtención de resultados.

En el estudio de ambos transistores, fue común la necesidad de desarrollar herramientas y procedimientos no incluidos en el simulador y asimismo explotar al máximo su versatilidad para poder desarrollarlas dentro de sus actuales capacidades de cálculo. El resultado fue pues, en ambos casos, una mejora en el rendimiento del simulador que lo hace seguir siendo una valiosa herramienta de considerable recorrido en el estudio de los dispositivos analizados.

La estructura maestra que subyace y vertebra nuestro estudio de los SB–MOSFETs y de los TFETs consta de los siguientes puntos:

- Exposición de los fenómenos y mecanismos físicos que simular para la adecuada y precisa descripción de cada dispositivo.
- Identificación de las limitaciones de simulación existentes, así como de las eventuales incompatibilidades de resolución numérica y problemas de índole general que se plantean en el uso de ATLAS aplicado a nuestro estudio.
- Desarrollo de eficaces y sencillos algoritmos capaces de vencer los obstáculos planteados, y hacerlo en el marco de los límites del simulador.
- Presentación de resultados y conclusiones derivadas de la aplicación de los mecanismos propuestos de simulación.

En el caso de los SB–MOSFETs, el enfoque de su análisis vino marcado por la necesidad de ajuste de resultados experimentales, que puso de manifiesto que en ATLAS los llamados mecanismos de bajada de

barrera (BL por sus siglas en inglés) no son correctamente tenidos en cuenta en los procesos de inyección de portadores que involucran algún tipo de efecto túnel (emisión de campo y emisión de campo termoiónica). Por esta razón, y considerando que pequeñas variaciones en la altura de las barreras de las uniones metal–semiconductor rectificadoras (*barreras Schottky*, o SB) pueden tener un gran efecto sobre la corriente total dependiendo del tipo de mecanismo de inyección dominante en cada caso, resulta de gran importancia describir de la manera más precisa posible los procesos de bajada de barrera. En nuestro caso, desarrollamos un procedimiento de carácter iterativo combinado con una discretización vertical del canal que nos permitió tener en cuenta la dependencia de la altura de la barrera con la profundidad en el canal. Los ajustes obtenidos de los datos experimentales fueron notables, así como la capacidad de las simulaciones efectuadas para reproducir efectos de canal corto constatados experimentalmente, como por ejemplo la disminución de corriente al reducir la longitud de puerta hasta los 20nm.

Respecto a los Tunnel–FETs, y teniendo presente que en la actualidad la mayor parte de la investigación que se hace sobre ellos aún involucra aproximaciones semiclásicas, quisimos dar un paso adelante y llevar a caso un tratamiento más completo que supusiera la inclusión de los efectos cuánticos de confinamiento. Estos efectos, especialmente relevantes si se tienen en cuenta los actuales niveles de miniaturización de estos dispositivos, suponen en la práctica la sustitución de las bandas de conducción y de valencia por un espectro discreto de sub–bandas. La existencia de estos niveles energéticos o sub–bandas inevitablemente hace que se vean afectados los procesos de efecto túnel banda a banda característicos de estos dispositivos. El efecto de todo ello es una disminución de corriente al aumentar la separación efectiva entre los primeros estados excitados de las antiguas bandas de conducción y valencia (con el consiguiente aumento de la distancia de la barrera túnel que tienen que atravesar los portadores). Igualmente, nuestro estudio nos proporcionó argumentos para afirmar que la inclusión de los efectos de confinamiento condiciona el comportamiento y tendencias de los dos parámetros eléctricos más relevantes en la caracterización de los TFETs:

las tensiones umbrales y las pendientes subumbrales.

Al igual que con los SB–MOSFETs, en el caso de los TFETs, limitaciones del simulador hicieron necesario el desarrollo de mecanismos que soslayaran las dificultades con que nos encontramos. A saber: (*i*) incompatibilidades de resolución numérica en la aplicación simultánea del modelo autoconsistente de Schrödinger–Poisson junto con el modelo no local de corriente túnel banda a banda; y (*ii*) imposibilidad del modelo de corriente túnel banda a banda de inyectar portadores entre estados excitados resultantes del proceso de cuantización de las bandas de conducción y valencia.

I.3 Schottky barrier MOSFETs

Son, como hemos indicado anteriormente, unos de los posibles candidatos para reemplazar a los MOSFETs convencionales. Su estructura difiere de la de aquéllos en que las regiones de fuente y drenador son sustituidas por contactos metálicos [5, 10–12]. La correspondiente formación de uniones metal–semiconductor de carácter rectificador da lugar a la formación de las *barreras Schottky*.

I.3.1 Principales ventajas y limitaciones

Las ventajas de reemplazar las habituales regiones dopadas de fuente y drenador son múltiples y pueden ser sucintamente expuestas de la siguiente manera:

(i) El carácter abrupto de las uniones metal–semiconductor permite dotar a los SB–MOSFETs de una inherente escalabilidad que potencialmente los podría llevar a longitudes de canal inferiores a los 10nm [5].

(ii) La reducción de la resistencias de fuente y drenador. Esta constituye muy posiblemente la principal ventaja de estos dispositivos. En los MOSFETs convencionales, la reducción de tamaño del canal hacía que las resistencias de los contactos se hicieran proporcional-

mente cada vez más importantes, llegando incluso a contrarrestar los beneficios derivados del proceso de escalado [13].

(iii) La posibilidad de formación de transistores bipolares entre transistores MOSFET adyacentes dando lugar a procesos de conducción parásita (conocido en inglés como "parasitic bipolar latch-up") queda eliminada por la presencia de los contactos metálicos en la fuente y el drenador [14–16].

(iv) La ausencia de dopantes en la fuente y drenador previene en cierta medida a los SB–MOSFETs de fenómenos de variabilidad. Igualmente, las menores temperaturas necesarias para su elaboración hacen que la fabricación de los SB–MOSFETs sea un proceso, paradójicamente, más sencillo y con menos etapas.

(v) La eliminación del solapamiento de los contactos de fuente y drenador con el contacto de puerta contribuye a la eliminación de capacidades parásitas. En su lugar, se introduce una cierta separación ("underlap" en inglés) entre la fuente y el drenador y el canal de puerta.

(vi) La naturaleza de las uniones metal–semiconductor hace que la altura de las barreras Schottky sea independiente de la longitud del canal, con lo que se consigue un mejor control de la corriente de pérdidas en dispositivos de canal corto [5].

En lo concerniente a las limitaciones, podemos mencionar:

(i) La independencia de la altura de las barreras Schottky con la longitud del canal —que vimos que era una ventaja para el control de las corrientes de pérdidas—, puede también ser un inconveniente en el caso de barreras que resulten altas, por cuanto se ha demostrado que limitan la corriente de drenador y las pendientes subumbrales [17, 18]. Por ello, resulta muy atractivo el empleo de los llamados *siliciuros complementarios* que proporcionan barreras de alturas complementarias según el tipo de portador considerado [19–22].

(ii) El pinzamiento del nivel de Fermi en la interfase metal–semiconductor como consecuencia de la existencia de estados interfaciales [3]. Diversas investigaciones han planteado posibles alternativas para deshacer dicho pinzamiento a expensas de incrementar el número de etapas de fabricación [1, 23–31].

I.3.2 Procesos de transporte e inyección de portadores a través de barreras Schottky

Los principales procesos concurrentes en las uniones entre los contactos metálicos de fuente y drenador y el canal semiconductor —cuya importancia relativa dependerá de las condiciones de polarización del transistor— son esencialmente tres como puede verse en la Fig. I.1 para electrones:

- **Emisión termoiónica**. Se refiere a la emisión de portadores por encima de la barrera Schottky. Dependiendo de la situación puede estar limitada por procesos de difusión–deriva. Tradicionalmente en la bibliografía, ambos mecanismos se han tratado por separado [32–34]. En general, la corriente vendrá determinada por aquél de los dos que suponga un mayor impedimento al flujo de portadores. La emisión termoiónica es el proceso dominante en semiconductores de alta movilidad con dopados y temperaturas moderadas. La expresión que describe la densidad de corriente para electrones en este proceso viene dada por [3]

$$\begin{aligned} J_{n,te} &= \left[A^* T^2 \exp\left(-\frac{q\phi_{bn}}{kT}\right)\right]\left[\exp\left(\frac{qV}{kT}\right) - 1\right] \\ &= J_{TE}\left[\exp\left(\frac{qV}{kT}\right) - 1\right], \end{aligned} \tag{I.1}$$

con

$$J_{TE} = A^* T^2 \exp\left(-\frac{q\phi_{bn}}{kT}\right). \tag{I.2}$$

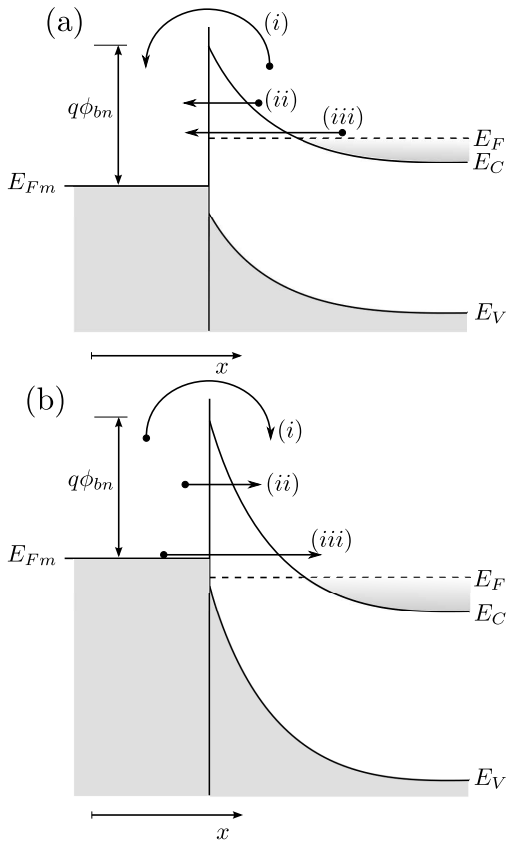

Figura I.1: Principales procesos de inyección de electrones del semiconductor al metal en una unión metal–semiconductor (degenerado de tipo n) polarizada en directa (a), y del metal al semiconductor al polarizar en inversa (b). Los procesos mostrados son: emisión termoiónica (i), emisión de campo termoiónica (ii), y emisión de campo desde el nivel de Fermi (iii).

Donde A^* es la constante de Richardson, ϕ_{bn} la barrera Schottky para electrones, T la temperatura y V la tensión aplicada.

- **Emisión de campo termoiónica**. Aquellos portadores con energías por encima del nivel de Fermi y por debajo de la altura de la barrera pueden atravesarla por efecto túnel. Este mecanismo es bastante dependiente de la temperatura ya que conforme mayor es la temperatura, mayor es la población de portadores con energías por encima del nivel del Fermi. La expresión correspondiente a este mecanismo para electrones en una polarización en directa [3]

viene dada por

$$J_{n,\,tfe} = \frac{A^* T \sqrt{\pi E_{00} q(\phi_{bn} - \phi_n - V_F)}}{k \cosh\left(\frac{E_{00}}{kT}\right)} \exp\left[\frac{-q\phi_n}{kT} - \frac{q(\phi_{bn} - \phi_n)}{E_0}\right] \exp\left(\frac{qV_F}{E_0}\right), \tag{I.3}$$

$$E_0 = E_{00} \coth\left(\frac{E_{00}}{kT}\right). \tag{I.4}$$

Con ϕ_n la distancia en el semiconductor entre el nivel de Fermi y el extremo de la banda de conducción, V_F la tensión de polarización en directa, y E_{00} una *energía reducida* función del dopado del semiconductor, de la masa efectiva y de la permitividad [3, 34].

- **Emisión de campo**. Se refiere a la inyección por efecto túnel de los portadores con energías próximas al nivel de Fermi. De manera análoga a los casos anteriores, su expresión para una polarización en directa y en el caso de electrones es [35]

$$\begin{aligned} J_{n,\,fe} &= \frac{A^* T \pi \exp\left[-q\frac{(\phi_{bn} - V_F)}{E_{00}}\right]}{c_1 k \sin(\pi c_1 kT)} \left[1 - \exp(-c_1 q V_F)\right] \\ &\approx \frac{A^* T \pi \exp\left[-q\frac{(\phi_{bn} - V_F)}{E_{00}}\right]}{c_1 k \sin(\pi c_1 kT)}, \end{aligned} \tag{I.5}$$

con

$$c_1 = \frac{1}{2E_{00}} \log\left[\frac{4(\phi_{bn} - V_F)}{-\phi_n}\right]. \tag{I.6}$$

I.3.3 Regímenes de operación de los SB–MOSFETs

Para ilustrar los distintos regímenes en los que pueden operar los SB–MOSFETs, consideraremos el funcionamiento de uno de tales disposi-

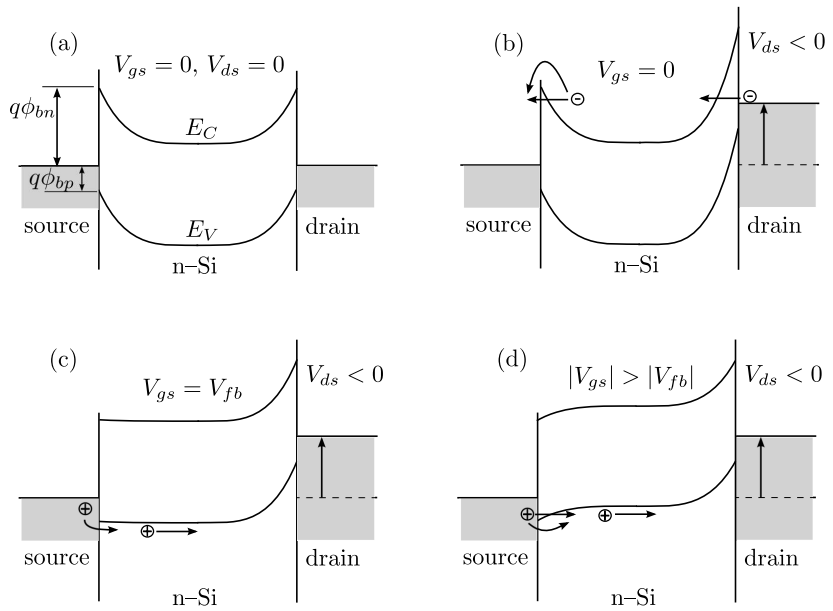

Figura I.2: Diagrama de bandas de energía correspondientes a un SB–MOSFET de canal p donde se muestran cuatro configuraciones distintas de operación: (a) estado OFF, sin tensión de drenador; (b) estado OFF, con tensión de drenador; (c) condición de banda plana (hasta este punto y viniendo de las situaciones anteriores sólo existe emisión termoiónica); (d) estado ON donde la inyección por efecto túnel está presente.

tivos de canal p (donde, por ejemplo, la fuente y el drenador podrían están fabricadas con PtSi). En el interior del canal, el transporte de huecos viene regido por los habituales procesos de difusión y deriva; mientras que en las uniones metal–semiconductor del canal con la fuente y el drenador, se tienen que tener el cuenta los procesos de inyección descritos en la sección anterior.

Por todo ello, y de acuerdo con los diagramas mostrados en la Fig. I.2, distinguimos las siguientes situaciones:

(a) Condición de equilibrio sin ninguna tensión aplicada.

(b) Corriente de pérdidas debida a electrones. En esta configuración, los electrones en el drenador poseen una probabilidad finita de atravesar por efecto túnel la barrera Schottky que se les presenta y entrar en el canal. Este efecto será especialmente pronunciado si el nivel

de Fermi en el drenador se eleva por encima de la banda de conducción en el canal, ya que esto posibilitará también la existencia de emisión de campo además de la emisión de campo termoiónica. Lo elevado de la barrera para electrones garantiza que esta componente de corriente de pérdidas es considerablemente menor que la corriente debida a huecos. Una vez en el canal, los electrones son arrastrados a hasta la unión fuente–canal donde deben remontar una segunda barrera de menor importancia para salir del dispositivo.

(c) En todo el abanico de configuraciones entre las mostradas en las Fig. I.2(b) y (c), el transistor se mueve en el régimen subumbral en el que la corriente de drenador es enteramente debida al mecanismo de emisión termoiónica de huecos por encima de la barrera de fuente a canal [36]. La banda de valencia permanece pues dentro de un rango de energías que no posibilita aún la aparición de corriente túnel alguna. Es justamente la Fig. I.2(c) la que marca el punto donde la componente termoiónica alcanza su máximo valor ya que el voltaje de puerta aplicado provoca una situación de banda plana en la unión fuente–canal.

(d) Cuando la banda de valencia sobrepasa la condición de banda plana y se eleva más, la contribución termoiónica no sigue aumentando —puesto que la barrera que deben remontar los huecos no disminuye— y, en cambio, es el momento de la aparición de la corriente túnel debida a la emisión de campo termoiónica. Si se continúa elevando la banda de valencia en el canal, eventualmente llegará un momento en que se alinee con el nivel de Fermi en la fuente y ello permitirá, además, la aparición de la contribución debida a la emisión de campo sin componente termoiónica.

I.3.4 Mecanismos de bajada de barrera

I.3.4.1 Efecto carga imagen

El efecto carga imagen (IFL por sus siglas en inglés) es un efecto dinámico que provoca una bajada en la altura de la barrera Schottky en presencia

de un campo eléctrico. Su nombre procede del efecto que un electron en presencia de un metal provoca en él induciendo una carga positiva sobre su superficie. Esta carga positiva ejerce sobre el electrón la misma fuerza que ejercería una carga de igual valor pero signo opuesto situada especularmente respecto a la superficie del metal. La expresión analítica que describe este efecto viene dada por

$$\Delta\phi_{ifl} = \sqrt{\frac{q|F_m|}{4\pi\varepsilon_s}}, \tag{I.7}$$

con F_m el valor del campo eléctrico en la unión, donde alcanza su máximo valor.

I.3.4.2 Efecto dipolo

El efecto dipolo (DL) se creyó originalmente ligado únicamente a la existencia de estados interfaciales en la superficie del semiconductor. Estados que podían tener su origen bien como resultado de la presencia de una capa de óxido entre el metal y el semiconductor [37–39]; o bien como consecuencia de imperfecciones de la interfase metal–semiconductor.

No obstante, la aparición de este efecto en superficies atómicamente limpias y abruptas hizo plantear la hipótesis de que el efecto dipolo tuviera alguna contribución adicional. Y, en efecto, tal contribución procede de la resolución mecanocuántica de la distribución de carga en equilibrio en una unión metal–semiconductor. En dicha resolución es posible observar una cierta penetración de las funciones de onda electrónicas — las llamadas *colas de Heine* [40]— del metal en la región prohibida (gap) del semiconductor [31, 41]. El resultado es la formación de los conocidos como *estados del gap inducidos por el metal*, comúnmente denominados por su acrónimo del inglés MIGS [40, 42], y la formación de un dipolo en la interfase que consiguientemente modifica la altura de la barrera Schottky.

Una expresión analítica habitual [9] para dar cuenta de este efecto es

$$\Delta\phi_{dl} \quad = \quad \alpha F_m^{\gamma}, \tag{I.8}$$

donde α y γ son parámetros ajustables a partir de resultados empíricos.

En un principio se pensó que la contribución de este efecto podía ser tan relevante como la del de carga imagen [43], pero en situaciones de uniones extremadamente limpias y abruptas demuestra ser de muy poca importancia al lado del anterior.

I.3.5 Estructura de dispositivo y datos experimentales

El tipo de dispositivos analizados en el trabajo aquí presentado son SB–MOSFETs basados en tecnología de silicio sobre aislante (SOI en su acrónimo inglés) con contactos de fuente y drenador de NiSi y $NiSi_2$ epitaxial. Las alturas de barrera Schottky para estos siliciuros son $\phi_{bn} = 0.65$V y $\phi_{bn} = 0.37$V, respectivamente. Los transistores fueron fabricados en el Forschungszentrum Jülich [44]. En la Fig. I.3 se puede ver una representación esquemática de ellos.

En las Figs. I.4–I.6, se muestran las características de transferencia y salida experimentales de estos dispositivos para distintas tensiones aplicadas y diferentes longitudes de puerta.

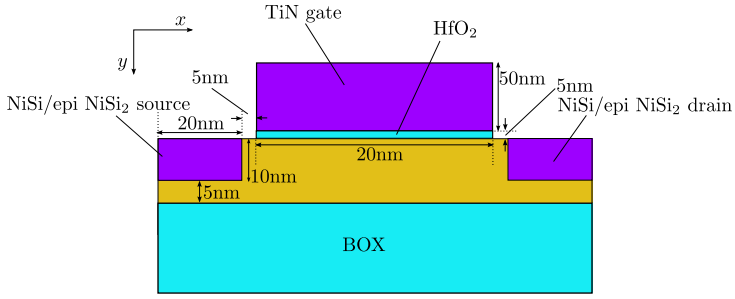

Figura I.3: Sección transversal y dimensiones (no representadas a escala) de los SB–MOSFETs con fuente y drenador de NiSi y $NiSi_2$ estudiados en esta tesis.

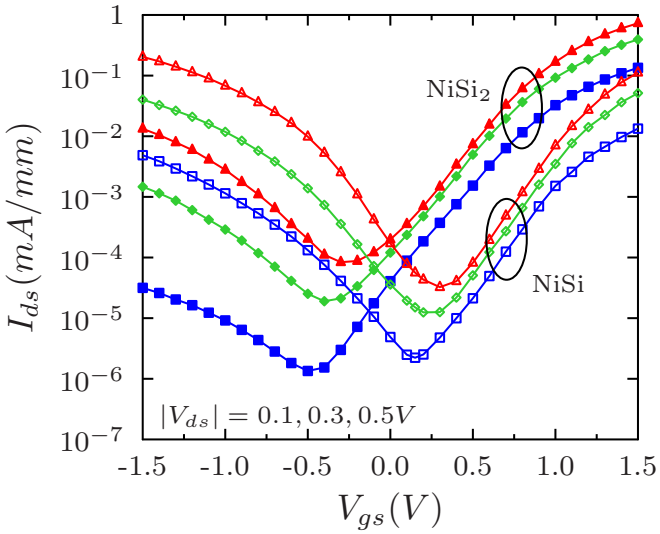

Figura I.4: Características de transferencia experimentales correspondientes a los SB–MOSFETs con NiSi y NiSi$_2$ epitaxial para una longitud de puerta de $L_g = 20$nm. $V_{ds} = 0.1$, 0.3 y 0.5V.

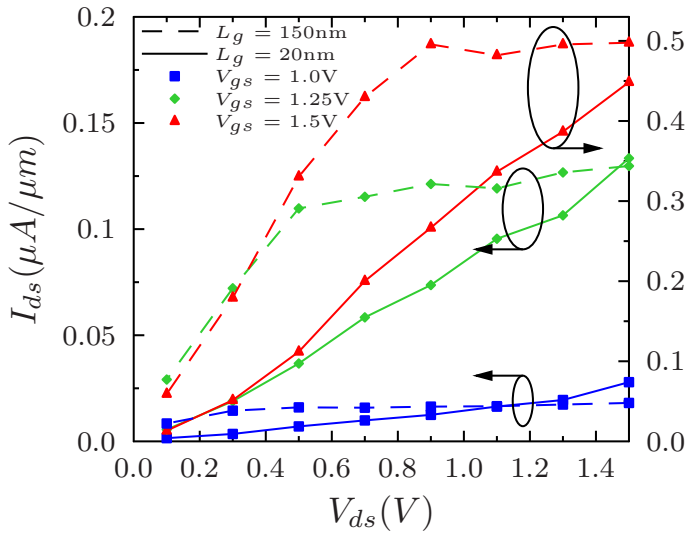

Figura I.5: Características de salida experimentales de los SB–MOSFETs con NiSi para longitudes de puerta $L_g = 20$ y 150nm. $V_{gs} = 1.0$, 1.25 y 1.5V.

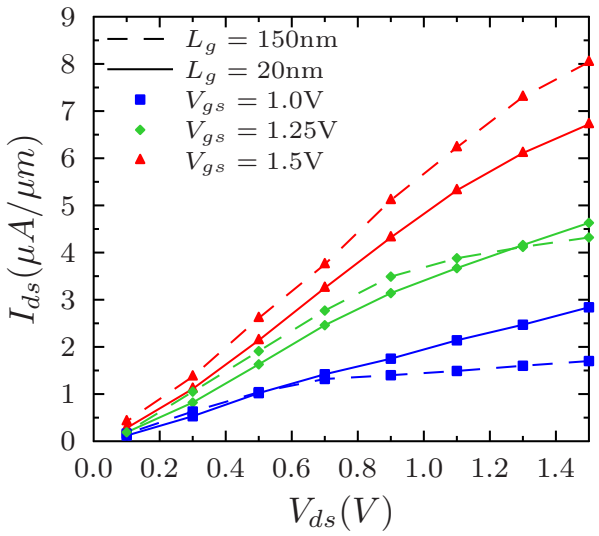

Figura I.6: Características de salida experimentales de los SB–MOSFETs con $NiSi_2$ epitaxial para longitudes de puerta $L_g = 20$ y 150nm. $V_{gs} = 1.0$, 1.25 y 1.5V.

I.3.6 Resultados de las simulaciones

I.3.6.1 Características de transferencia

Dada la limitación existente en ATLAS a la hora de incorporar correctamente los mecanismos de bajada de barrera aplicados a los distintos procesos de inyección de portadores, fue necesario el desarrollo de un mecanismo externo que aprovechara las propias capacidades de cálculo del simulador. Dicho mecanismo se basa en un proceso iterativo de extracción del campo eléctrico en las proximidades de la unión metal–semiconductor[2] y del consiguiente cálculo de la bajada de barrera asociada. La barrera así modificada vuelve a introducirse en el simulador y el proceso vuelve a iniciarse. Tras pocas iteraciones se alcanza la convergencia. Este proceso se acompaña con una discretización vertical del canal en diferentes capas horizontales que permite tener en cuenta la dependencia de la altura de la barrera con la profundidad. La potencia

[2]Problemas de convergencia ligados al propio diseño del mallado hacen inviable la extracción del campo eléctrico justo en la unión.

del mecanismo reside en gran medida en su sencillez, pues únicamente se utilizan como parámetros ajustables empíricamente las masas efectivas de efecto túnel de los portadores, y se continúa usando la aproximación WKB.

En el caso de las características de transferencia, las Figs. I.7 y I.8 muestran la comparación entre los resultados experimentales y los obtenidos de las simulaciones para los SB–MOSFETs con NiSi y $NiSi_2$, respectivamente.

I.3.6.2 Efectos de canal corto

Un aspecto interesante de los transistores SB–MOSFET estudiados es la disminución de corriente que experimentalmente se obtiene cuando se reduce la longitud de puerta desde 50nm a 20nm. Este comportamiento

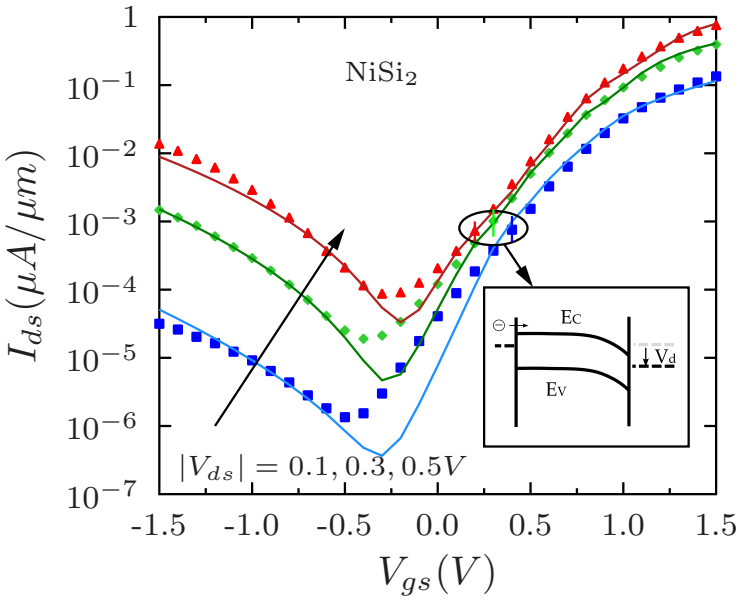

Figura I.7: Corrientes de drenador para $V_{ds} = 0.1, 0.3$ y 0.5V en el dispositivo con $NiSi_2$ epitaxial. Las líneas continuas corresponden a las características con la bajada de barrera calculada con nuestro método, los símbolos representan los datos experimentales. Las masas efectivas de efecto túnel consideradas son $m_h = 0.46$ y $m_e = 0.8$.

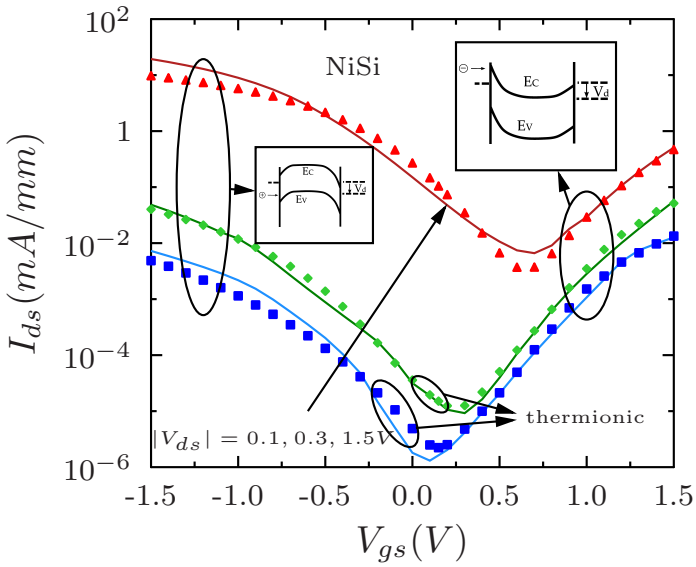

Figura I.8: Curvas $I_{ds} - V_{gs}$ correspondientes al SB–MOSFET con NiSi para $V_d = 0.1, 0.3$ y $1.5V$. Los símbolos marcan valores experimentales; las líneas continuas, los simulados. En este caso, $m_h = 0.8$ y $m_e = 0.4$.

es, en principio, contrario a las reglas habituales de escalado que predicen mayor nivel de corriente conforme se reduce el tamaño de la puerta. Sin embargo, predicciones apuntando en esta dirección de reducción de corriente también se pueden encontrar en la bibliografía [45]. En ellas se plantea que la reducción del tamaño de la puerta hasta dimensiones de $L_g \approx 20$nm podría hacer que el perfil de las barreras Schottky de los contactos de fuente y drenador afectaran a todo el canal, disminuyendo de esa manera la influencia de la tensión de puerta sobre la curvatura de las bandas de conducción y valencia.

En la Fig. I.9, se puede observar como los resultados obtenidos en nuestras simulaciones también reproducen este hecho experimental de reducción de corriente para $L_g = 20$nm. Igualmente, se aprecia como los resultados procedentes del cálculo propio de ATLAS para describir la bajada de barrera no sólo quedan lejos de los valores experimentales, sino que tampoco predicen concluyentemente una disminución de corriente al reducir la longitud de puerta.

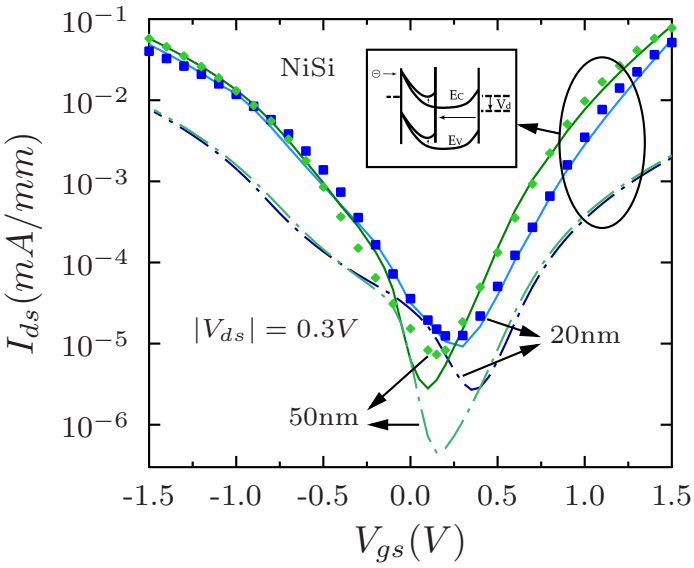

Figura I.9: Para valores crecientes de $|V_{gs}|$ obtenemos un mayor nivel de corriente para $L_g = 50$nm (rombos) que para $L_g = 20nm$ (cuadrados) en el SB–MOSFET con NiSi. Nuestras simulaciones ajustan con notable precisión (líneas continuas) los datos experimentales. Por contra, las simulaciones realizadas con los mecanismos de bajada de barrera internos de ATLAS no resultan satisfactorias para describir el comportamiento de estos dispositivos.

I.4 Tunnel FETs

Los transistores de efecto campo basados en corriente túnel banda a banda (TFETs) resultan especialmente atractivos a la hora de postularse como posibles sucesores de los MOSFETs. Los aspectos más reseñables de su funcionamiento son el bajo nivel de corriente en el estado OFF, y una inversa de la pendiente subumbral que puede situarse por debajo de los 60mV/decada. Ello convierte estos dispositivos en particularmente adecuados en aplicaciones de baja potencia.

Los TFETs basan su modo de operación en la inyección de portadores mediante efecto túnel banda a banda. Por tanto, al basarse en un proceso mecanocuántico fuertemente dependiente de la anchura de la barrera a atravesar, estos dispositivos permiten inversas de la pendiente subumbral extremadamente bajas. Esto es así ya que la corriente se incrementa

exponencialmente una vez que se alcanza la configuración que permite el inicio de la inyección banda a banda. Análogamente, mientras las condiciones para que aparezca el efecto túnel a traves de la barrera entre bandas no se den, es justamente esa barrera la que bloqueará radicalmente cualquier corriente haciendo que la corriente del estado OFF permanezca en niveles extremadamente bajos.

Es precisamente debido a que la corriente túnel banda a banda tiene lugar en una región muy limitada del espacio (de menos de 10nm), que el potencial escalado de estos dispositivos podría situarse en ese rango de tamaños.

I.4.1 Regímenes de operación de los TFETs

En los TFETs, a diferencia de lo que ocurre en los MOSFETs convencionales, las regiones de fuente y drenador poseen diferente tipo de dopado. El elemento constitutivo esencial, pues, de un TFET es típicamente un diodo p–i–n al que se incorpora una puerta. Un ejemplo de un TFET de una sola puerta puede verse en la Fig. I.10. Para que el transistor entre en modo de conducción, el diodo debe polarizarse en inversa y aplicarse un voltaje suficiente en la puerta. Un TFET de canal n como el de la figura requeriría una tensión positiva de puerta y de drenador.

En la Fig. I.11 se pueden ver los distintos regímenes de operación del TFET y su correspondencia con los estados ON y OFF. Vemos que

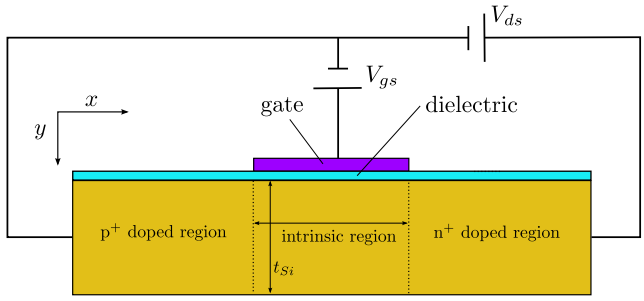

Figura I.10: TFET de puerta única. La región p$^+$ actúa como fuente mientras que la n$^+$ actúa como drenador.

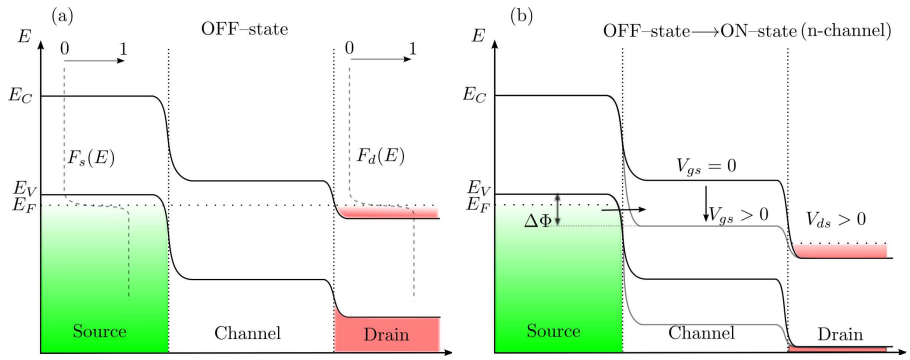

Figura I.11: Diagrama de bandas de energía del TFET correspondiente a un corte horizontal a lo largo del canal y próximo al aislante de puerta. (a) Estado OFF asociado a la situación de equilibrio donde no hay ninguna tensión aplicada en la puerta ni en el drenador. (b) Combinación de estados ON y OFF. Al aplicar una tensión de drenador, si no se aplica tensión alguna en la puerta, el efecto túnel banda a banda no puede ocurrir y sólo existiría corriente de pérdidas (estado OFF). Por contra, si aplicamos una tensión de puerta suficiente y positiva $V_{gs} > 0$, la banda de conducción en el canal desciende por debajo del nivel de la banda de valencia y se habilita la corriente túnel entre bandas (estado ON). $\Delta\Phi$ representa la diferencia entre el borde superior de la banda de valencia en la fuente y el extremo inferior de la de conducción en el canal, que es el rango de energías para el que está permitido el mencionado efecto túnel.

la sola aplicación de una tensión positiva en el drenador, $V_{ds} > 0$, no es suficiente para permitir el efecto túnel banda a banda. La corriente túnel no tendrá lugar hasta que la tensión de puerta no sea suficiente para que el extremo inferior de la banda de conducción en el canal quede alineado con el borde superior de la banda de valencia en la fuente. A partir de ese momento, los portadores (en el caso representado, electrones) pueden atravesar de la fuente al canal y el transistor entra en modo de conducción.

En aquellos TFETs diseñados con simetría entre las regiones dopadas, se puede obtener un comportamiento ambipolar siempre que se apliquen las tensiones adecuadas en los terminales.

I.4.2 Principios de funcionamiento

I.4.2.1 Modelos de corriente túnel banda a banda

Los modelos que describen el efecto túnel entre bandas buscan formalizar el cálculo de la probabilidad de transmisión a través de la barrera a partir del desarrollo de la expresion siguiente —en el caso unidimensional— mediante el uso de la aproximación WKB [3]

$$T(E) \approx \exp\left[-2\int_{x_{start}}^{x_{end}} k(x)dx\right], \tag{I.9}$$

siendo $k(x)$ el vector de onda del portador dentro de la barrera, y x_{start} y x_{end} los límites que determinan la anchura de dicha barrera. Los modelos se agruparán en locales y no locales dependiendo de si consideran o no, respectivamente, el campo eléctrico como constante en el proceso de integración para obtener la probabilidad de transmisión por efecto túnel. Los modelos locales permiten el manejo y obtención de expresiones analíticas, como por ejemplo (siguiendo con la descripción unidimensional)

$$T(E) \approx \exp\left(-\frac{4\sqrt{2m^*}E_g^{3/2}}{3qF\hbar}\right), \tag{I.10}$$

que permite obtener la siguiente expresión para la corriente túnel banda a banda [3, 46, 47]

$$I_{btbt} = \frac{Aq^2F}{36\pi\hbar^2}\sqrt{\frac{2m^*}{E_g}}D\exp\left(-\frac{4\sqrt{2m^*}E_g^{3/2}}{3qF\hbar}\right). \tag{I.11}$$

Donde E_g es el valor del gap, F el campo eléctrico, m^* la masa efectiva de túnel y D una cierta integral en la energía que involucra las distribuciones de Fermi para las probabilidades de ocupación en las bandas de conducción y valencia.

En nuestro caso, dada su mayor capacidad de descripción de la

situación real, nos centramos en los modelos no locales para los que la aplicación de cálculo numérico es obligada ante la imposibilidad de obtención de expresiones analíticas.

Extendiendo el formalismo a una configuración 2D, el cálculo de las probabilidades de transmisión por efecto túnel se efectúa en ATLAS mediante la expresión

$$T\left(E_{\|}, y\right) = \exp\left[-2\int_{x_{start}}^{x_{end}} k(x,y)dx\right], \tag{I.12}$$

donde y es la dirección perpendicular a aquella en la cual se asume que ocurre el efecto túnel, que se toma como x; y $E_{\|}$ representa cada una de las energías longitudinales para las que está permitida la inyección banda a banda via túnel. Una vez calculada esta probabilidad, la densidad de corriente vendrá dada por

$$\begin{aligned} J(y) = & \frac{q}{\pi\hbar}\iint T\left(E_{\|}, y\right)\left[f_l\left(E_{\|}+E_T\right) - f_r\left(E_{\|}+E_T\right)\right] \\ & \rho\left(E_T\right) dE_{\|}\, dE_T, \end{aligned} \tag{I.13}$$

con E_T la energía transversal, $f_{l(r)}$ la distribución de Fermi en el lado izquierdo (derecho) de la unión correspondiente a los portadores mayoritarios, y $\rho(E_T)$ la densidad de estados 2D correspondiente a las dos componentes transversales del vector de ondas.

I.4.2.2 Formación de la capa de inversión

De la misma forma que ocurre en los MOSFETs convencionales, la capacidad del semiconductor en los TFETs se ve incrementada cuando aparece una capa de inversión pegando a la interfase del dieléctrico como consecuencia del voltaje aplicado en la puerta. La formación de esta capa de inversión resulta de capital importancia ya que altera notablemente la relación existente entre la tensión de puerta, V_{gs}, y el potencial de

superficie del canal, ϕ_s[3], cuya modificación es a su vez responsable de la variación de la anchura de la barrera de túnel banda a banda.

La expresión que controla la correspondencia entre V_{gs} y ϕ_s es

$$\Delta\phi_s = \frac{C_{ox}}{C_{ox} + C_{semi}} \Delta V_{gs}. \tag{I.14}$$

De donde vemos claramente que cuando se forma la capa de inversión, C_{semi} crece y la tensión aplicada cae fundamentalmente en el óxido de puerta. En esa situación, el control de la puerta sobre ϕ_s disminuye drásticamente y, consecuentemente, también lo hará el control de ésta sobre la anchura de la barrera de túnel.

I.4.2.3 Pendiente subumbral

En los TFETs, la pendiente subumbral es significativamente dependiente del valor de tensión de puerta por cuanto estos transistores están basados en un mecanismo de inyección de portadores diferente al de los MOSFETs. Justamente este hecho hace que el límite 60mV/decada para la inversa de la pendiente subumbral (SS de sus siglas en inglés) pueda ser rebasado a la baja.

La dependencia de SS con V_{gs} puede obtenerse a partir de la definición

$$SS = \left[\frac{d\log(I_{ds})}{dV_{gs}}\right]^{-1} = \ln 10 \left[\frac{d\ln(I_{ds})}{dV_{gs}}\right]^{-1}, \tag{I.15}$$

que en los modelos locales [48] conduce a la siguiente expresión

$$\begin{aligned} SS &= \ln 10 \left[\frac{d\ln f(V_{gs})}{dV_{gs}} + \frac{d\ln F(V_{gs})}{dV_{gs}} + \frac{d}{dV_{gs}}\left(-\frac{C}{F\,(V_{gs})}\right)\right] = \\ &= \ln 10 \left[\frac{1}{f(V_{gs})}\frac{df(V_{gs})}{dV_{gs}} + \left(\frac{F(V_{gs}) + C}{F^2(V_{gs})}\right)\frac{dF(V_{gs})}{dV_{gs}}\right]^{-1}. \end{aligned} \tag{I.16}$$

[3]En los TFETs emplearemos para su cálculo el extremo de la banda de conducción en la interfase del dieléctrico y en el centro del canal.

Donde $f(V_{gs})$ es una cierta función que agrupa dependencias en V_{gs}, F es el campo eléctrico y C una constante.

El hecho de que SS no sea uniforme en la región subumbral motiva la definición de dos nuevas magnitudes a la hora de caracterizar estos dispositivos. Por un lado, *la inversa de la pendiente subumbral puntual*, S_{pt}, que es el mínimo valor de SS a lo largo de la curva $I_{ds} - V_{gs}$; y por otro, *la inversa de la pendiente subumbral promedio*, S_{av}, que definimos como la inversa de la pendiente entre el punto correspondiente a la tensión umbral y el punto donde se inicia la corriente túnel banda a banda.

I.4.3 Inclusión del confinamiento cuántico

Para incluir en el estudio de los TFETs el efecto cuántico de confinamiento, utilizamos el modelo autoconsistente de Schrödinger–Poisson que incorpora ATLAS. En nuestras simulaciones, el confinamiento ocurre en una dimensión (que tomaremos como y), de manera que la ecuación de Schrödinger que tendremos que solucionar será unidimensional. En el caso de electrones, tendremos que resolver para valores propios de energía E_{il} y funciones de onda Ψ_{il} en cada sección perpendicular a la dirección x y para cada valle l

$$-\frac{\hbar^2}{2}\frac{\partial}{\partial y}\left[\frac{1}{m_y^l(x,y)}\frac{\partial \Psi_{il}}{\partial y}\right] + E_C(x,y)\Psi_{il} = E_{il}\Psi_{il}. \qquad \text{(I.17)}$$

Con m_y^l la masa efectiva en la dirección y para el valle l, y $E_C(x,y)$ un extremo de la banda de conducción. Una expresión equivalente se debería plantear para huecos con las oportunas modificaciones.

El procedimiento autoconsistente calcula, a partir de los valores propios de energía y funciones de onda, la concentración de portadores y la sustituye en la ecuación de Poisson para obtener el correspondiente potencial. El potencial así obtenido vuelve a introducirse en la ecuación de Schrödinger, y vuelve a calcularse una nueva concentración de portadores modificada. Ésta, a su vez, permite obtener un nuevo potencial, repitiéndose el proceso sucesivamente hasta que se alcanza la convergencia. En nuestro caso, como ATLAS considera isótropas las masas efec-

tivas de túnel, consideramos también la simplificación de tomar como isótropas las masas efectivas convencionales.

No obstante lo anterior, la aplicación del modelo de Schrödinger–Poisson para incorporar los efectos de confinamiento presenta algunos problemas que fue necesario resolver:

- De un lado, nos encontramos la incompatibilidad numérica en ATLAS entre los métodos numéricos empleados al aplicar el modelo no local de corriente túnel banda a banda y el modelo de Schrödinger–Poisson. Para solucionar el problema, cada etapa de cálculo al resolver el dispositivo se divide en dos pasos. En el primero, se resuelve autoconsistentemente Schrödinger–Poisson para las condiciones de tensión elegidas y se obtienen el potencial, la distribución cuántica de carga y el espectro discreto de subniveles de energía en que quedan discretizadas las bandas de conducción y valencia. En el segundo, tomando como marco el potencial y las distribuciones de carga calculadas en el anterior paso, calculamos la inyección de portadores por efecto túnel banda a banda asumiendo que esta inyección no las modifica significativamente.

- Por otro lado, surge una limitación intrínseca al propio modelo no local de corriente túnel banda a banda. Este modelo asume invariablemente que los portadores son inyectados entre los extremos de las bandas de conducción y valencia, y no permite la posibilidad de que dicho proceso tenga lugar entre otros estados como, por ejemplo, los resultantes del proceso de discretización de bandas. Para superar este inconveniente, lo que era una dificulad en el punto anterior ahora se convierte en una nada desdeñable ventaja. Puesto que la ecuación de Poisson no se resuelve en el segundo paso de nuestro cálculo secuencial (el de inyección de portadores), ni el potencial ni la distribución de portadores se verán modificadas. Nuestra estrategia será, pues, modificar manualmente la separación entre bandas incrementando artificialmente el gap para hacerlo coincidente con el que existe entre los primeros estados

excitados. De esta manera, estaremos efectuando una especie de calibración que logra que el modelo simule la inyección entre los primeros estados excitados de las bandas.

I.4.4 Estructura de dispositivo

Nuestra atención se centrará en una estructura TFET de puerta única (SG) y otra de doble puerta (DG), como se ilustra en la Fig. I.12

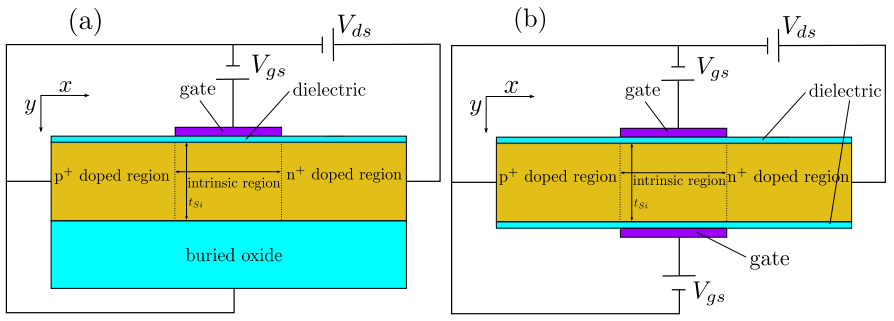

Figura I.12: Sección transversal esquemática (no a escala) del dispositivo de puerta única (a) y del doble puerta (b) de canal n considerados en nuestro trabajo.

En ellos, la puerta tiene un dopado p^+ de 10^{20}átomos/cm^3; el drenador, un dopado n^+ de 10^{20}átomos/cm^3; y la región intrínseca (que en realidad consideraremos ligeramente dopada), un dopado tipo n de 10^{17}átomos/cm^3. Por defecto, y salvo que se indique otra cosa, el dieléctrico será SiO_2 y el espesor de óxido de puerta 1nm. La longitud de las regiones de fuente y drenador es de 100nm, y la longitud de puerta, L_g, 20nm.

I.4.5 Resultados de las simulaciones

I.4.5.1 Efecto de la variación de espesor del canal

En nuestras simulaciones, consideramos tres espesores diferentes para el canal, t_{Si}, de 10, 5 y 3nm. Las características de transferencia obtenidas semiclásicamente —sin confinamiento cuántico—, y las que resultan de incluirlo se muestran en las Figs. I.13 y I.14 para tensiones de drenador de $V_{ds} = 0.1$V y $V_{ds} = 1.1$V, respectivamente.

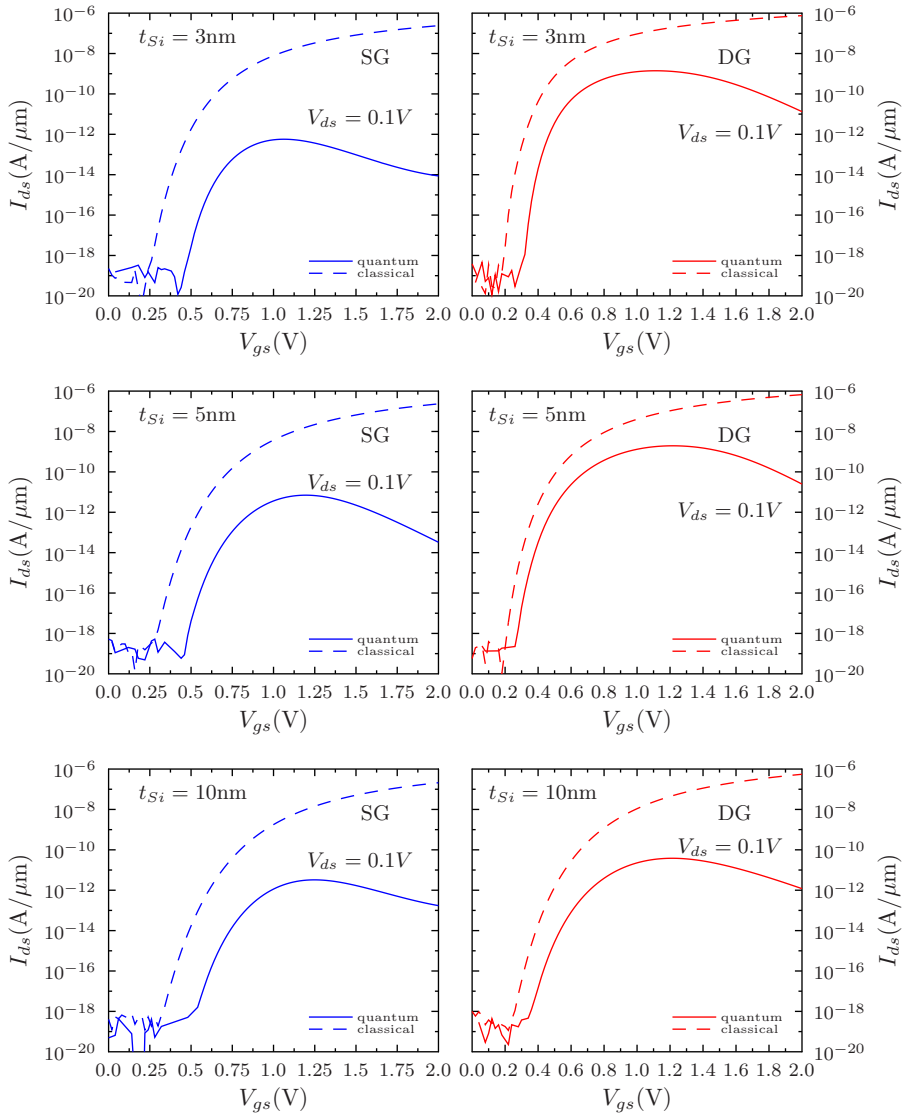

Figura I.13: Comparación entre las curvas $I_{ds} - V_{gs}$ semiclásicas y cuánticas para los TFETs de puerta única y de doble puerta con $t_{Si} = 3$, 5 y 10nm. $V_{ds} = 0.1$V.

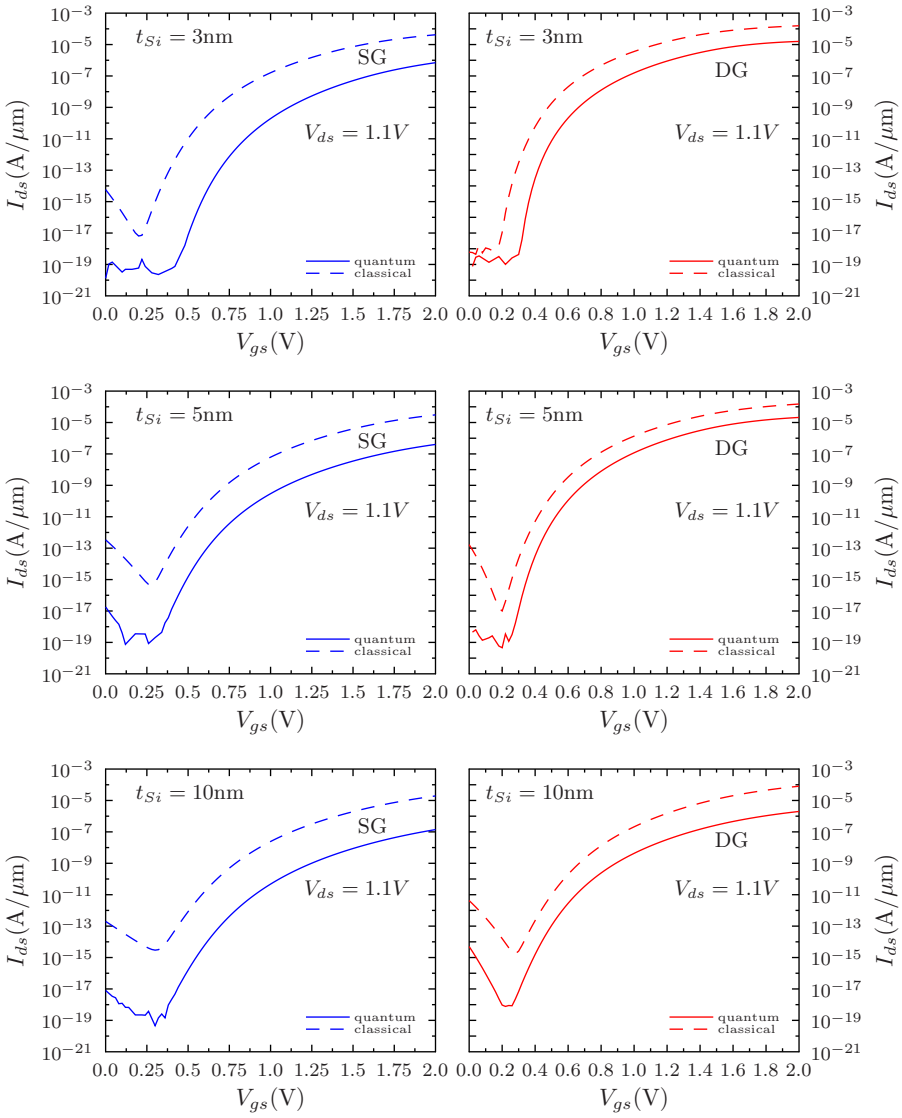

Figura I.14: Comparación entre las características de transferencia semiclásicas y cuánticas para los TFETs de puerta única y de doble puerta con $t_{Si} = 3$, 5 y 10nm. $V_{ds} = 1.1$V.

I.4.5.2 Efecto de la variación del dieléctrico de puerta

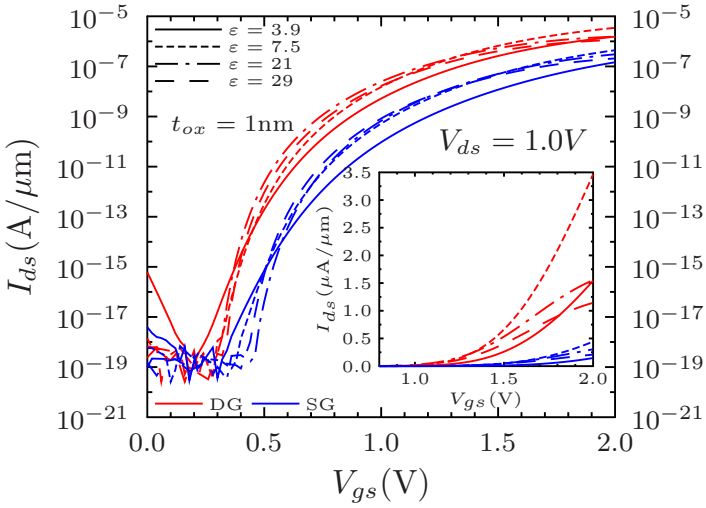

Figura I.15: Características de transferencia con efectos de confinamiento para $t_{Si} = 10$nm con diferentes dieléctricos de puerta para las estructuras SG y DG. $V_{ds} = 1.0$V. La gráfica interior muestra I_{ds} en escala lineal.

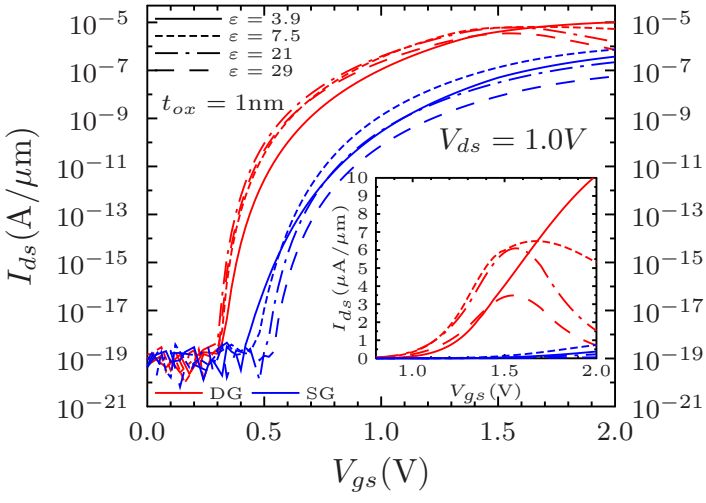

Figura I.16: Curvas $I_{ds}-V_{gs}$ con confinamiento para $t_{Si} = 3$nm con diferentes dieléctricos de puerta para los transistores SG y DG. $V_{ds} = 1.0$V. La gráfica interior corresponde a una escala lineal.

Cuando incrementamos la constante dieléctrica del óxido de puerta, encontramos que sorprendentemente para valores altos de ésta y conforme se reduce el espesor del canal, la corriente tiende a disminuir. Este efecto es más pronunciado cuando se incluye el confinamiento cuántico como puede observarse en las Figs. I.15 y I.16 para espesores $t_{Si} = 10$ y 3nm, con $V_{ds} = 1.0$V.

I.4.5.3 Efecto sobre la tensión umbral de puerta

El efecto más reseñable sobre la tensión umbral de puerta, V_{tg}[4], cuando se incluye el confinamiento cuántico es la aparente desaparición de la saturación de las curvas $V_{tg} - V_{ds}$ que parecía existir para tensiones de drenador altas en el escenario clásico.

Estos comportamientos, junto con los de las tensiones V_{inv} (tensión de puerta a la que se forma la capa de inversión), y V_{on} (tensión de puerta a la que se inicia la corriente túnel banda a banda) se muestran en las Figs. I.17–I.19 para el dispositivo de doble puerta.

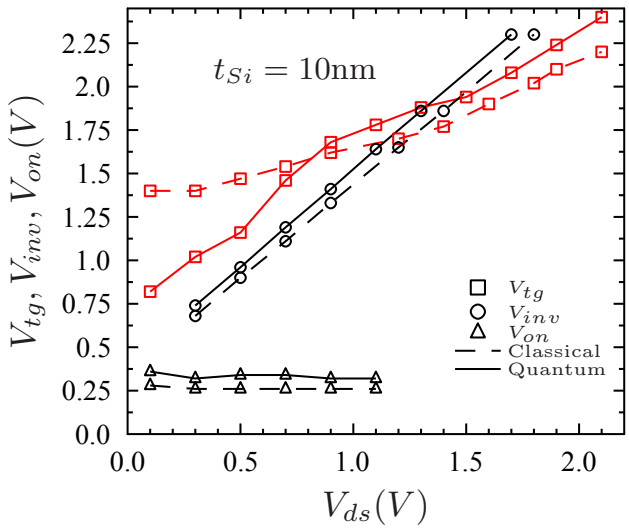

Figura I.17: Dependencia de V_{tg} con V_{ds} para la estructura DG con $t_{Si} =$ 10nm. Igualmente se muestran las tensiones V_{inv} y V_{on}.

[4]En los TFETs, también existe una tensión umbral de drenador cuyo comportamiento no se ha incluido en este resumen pero sí en la memoria principal.

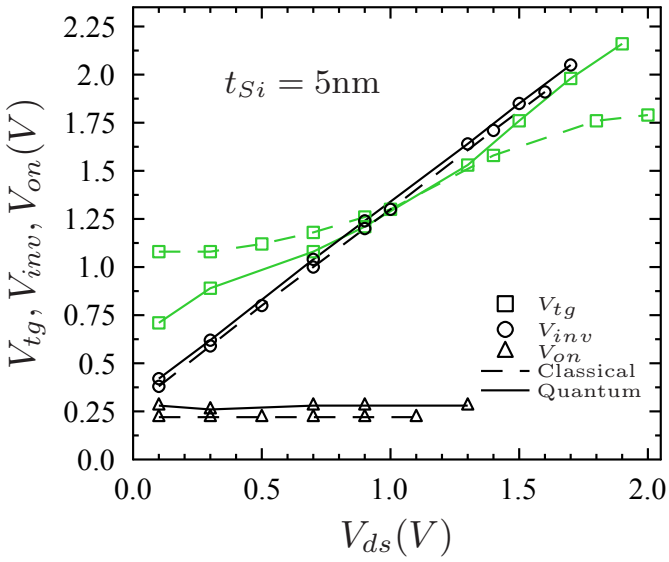

Figura I.18: Dependencia de V_{tg} con V_{ds} para la estructura DG con $t_{Si} =$ 5nm, junto con las tensiones V_{inv} y V_{on}.

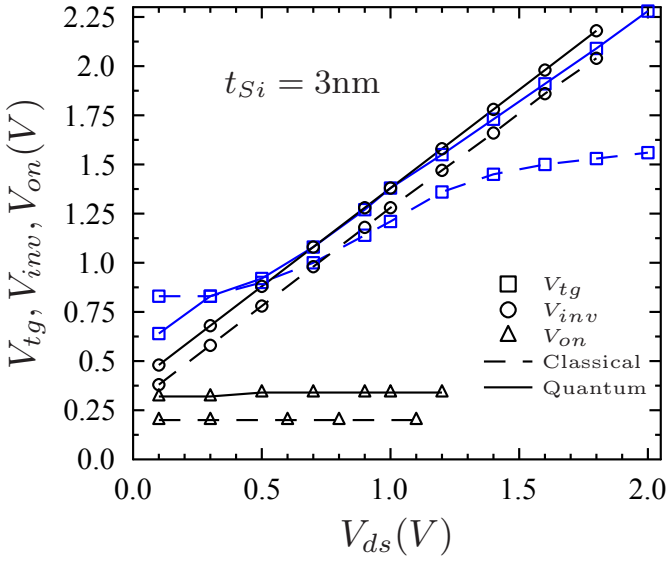

Figura I.19: Dependencia de V_{tg} con V_{ds} para la estructura DG con $t_{Si} =$ 3nm. También se incluyen las tensiones correspondientes a V_{inv} y V_{on}.

I.4.5.4 Efecto sobre la inversa de la pendiente subumbral

Debido a la existencia de los dos tipos de inversas que mencionamos en la Sec. I.4.2.3. Resultará interesante mostrar los resultados obtenidos para cada una de ellas. En nuestro caso, presentamos en la Fig. I.20 los resultados correspondientes al dispositivo DG en función del espesor de canal empleado según se incluya o no el efecto de confinamiento cuántico.

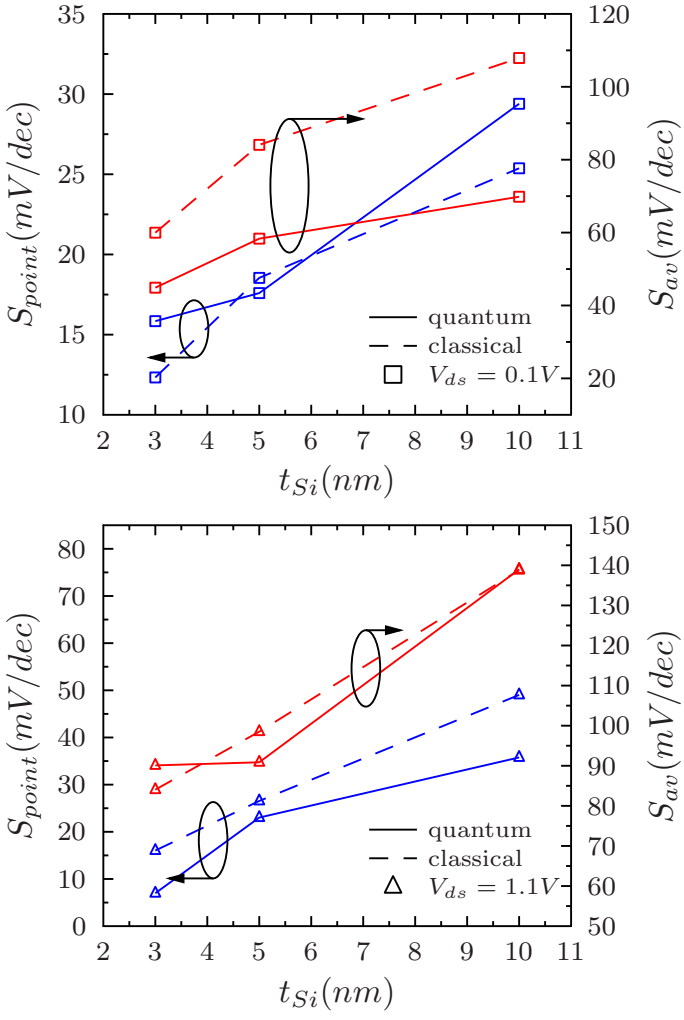

Figura I.20: S_{pt} y S_{av} como función de t_{Si} a $V_{ds} = 0.1$ y 1.1V para la estructura DG–TFET.

Los resultados más interesantes se refieren al comportamiento cuántico de S_{av} a tensiones bajas de drenador, que permitirían hablar de valores por debajo de 60mV/decada para espesores de $t_{Si} \leq 5$nm. Clásicamente, en cambio, estos niveles sólo eran alcanzables a partir de espesores de 3nm.

Part II

Main Document

Chapter 1

Introduction

Since the appearance of the metal–oxide–semiconductor field–effect transistor (MOSFET), it became the cornerstone of the development in the semiconductor industry by continuously improving its performance by means of tireless scaling efforts upheld during the past 40 years. Within this time it has been reduced from about $10\mu m$ to the sub–20nm regime. This dramatic downscaling has favoured the integration of MOSFETs as well as tremendous advances in microelectronics, which from now on must be more properly referred to as nanoelectronics.

The essentials of this scaling process were the progressive reduction of voltages and device dimensions by a certain factor κ and, at the same time, the increase of doping concentrations (N_A, N_D) by that same factor. The aim of such a way of proceeding was to keep a constant electric field throughout the channel length of the MOSFET [4]. By doing so, the power consumption per area (power density) remained constant while the circuit delay went down by κ and the circuit density increased with κ^2. However, in practice, the realistic implementation of the scaling process was more linked to constant voltage scaling, which was a specific way of implementing the general scaling method [49], which until recently was the main downscaling guideline. The idea of this voltage scaling was to keep the supply voltage at certain predetermined voltage nodes (5V, 3.3V, 1.5V...) while downscaling the device dimensions, and only switching to lower nodes when reliable operation due to increasing

electric fields was not possible.

Predicting this scaling race, the so–called Moore's law stated that the transistor number on an integrated circuit chip would double every 18 months and that the characteristic feature size would be in the 30nm regime by the year 2015. An accelerated device scaling allowed that regime to be reached much earlier than expected. But precisely for that reason, new technological issues have to be now faced. Some of them are:

(i) *The increasing source/drain parasitic resistance.* The scaling of the device dimensions implies a reduction of the intrinsic channel resistance. However, a similar reduction does not appear in the contact resistances thus making their relative contribution to the total resistance of the device increase. It has been stated that this contribution may become dominant as the gate length is reduced [6, 7].

(ii) *Short channel effects (SCEs).* The decrease of the channel length may cause deteriorated device performance affecting, for example, some parameters of the transistor such as the threshold voltage, or the reduction of the source–to–channel barrier due to the drain voltage (the so–called *drain induced barrier lowering*, DIBL).

(iii) *Gate oxide leakage.* The increase of the leakage current through the gate oxide limits the use of SiO_2 as insulator material for oxide thicknesses below 2nm. Consequently, the employment of high–κ materials becomes a quite advisable feature which is currently being taken into account in the commercialization of last generation devices.

Furthermore, as a result of the scaling process, the supply voltage (commonly referred to as V_{DD}) has been reduced to about 20% of its original value from the 1.4μm node to the 65nm node, whereas the threshold voltage only has gone down to approximately half of its original value (and not precisely as a result of the scaling). This difference in both reduction rates causes the gate overdrive also to go down and this, in

turn, produces on–current reduction which negatively affects device performance. One possible solution might be not to scale down V_{DD}, but that would increase power density and, subsequently, as more devices are added to a chip, the power consumption would also rise considerably. This is the so–called power crisis.

Finally, in addition to this technological challenges, fundamental limits are now being reached in conventional MOSFETs like the subthreshold swing limit arising from

$$SS = \frac{dV_{gs}}{d\left(\log I_{ds}\right)} = \frac{kT}{q}\ln 10\left(1+\frac{C_d}{C_{ox}}\right) > \frac{kT}{q}\ln 10 \approx 60\text{mV/dec}, \tag{1.1}$$

which simply cannot be surpassed.

It can be thus claimed that the semiconductor industry has entered the era of material–limited device scaling [5] after these four decades of straightforward scaling. A complete summary of these pending challenges to be faced is included in the *International Technology Roadmap for Semiconductors* (ITRS) [50].

All the above has led to a widespread interest across all fonts in novel FET designs and materials to obtain better performance as we scale to the sub–30nm regime. Two main work areas may be differentiated:

1. The first is mainly focused on novel engineeering solutions to create improved device architectures. For instance, the introduction of new gate dielectrics (high–κ materials), high mobility bulk materials (strained Si, SiGe, etc) and novel designs with great electrostatic control, like multigate devices. This development route wants to minimize short channel issues to create devices which prove to be more "long–channel" like in their behaviour.

2. The second approach, more long term focused, is that of the "new injection mechanisms". This area seeks the exploitation of new transport mechanisms and physical phenomena made possible due to new materials and small dimensions. FETs belonging to this camp should

therefore, in theory, not be limited by the aforementioned 60mV/dec SS barrier.

Among some of the most promising solutions, we find the two devices that we study in the present work: (1) the Tunneling Field–Effect Transistors (TFETs), as a clear example of novel injection mechanism devices (band–to–band tunneling) that may therefore overcome the SS limitation and replace conventional MOSFETs for low–power applications; and (2) the Schottky Barrier MOSFETs (SB–MOSFETs), which also incorporate injection mechanisms different from only diffusion or thermionic emission (thermionic field emission and field emission) and allow ease of integration in the existing CMOS semiconductor infrastructure. Further improvement of these new devices may also be accomplished by implementing some engineering solutions over them, thus combining to some extent the two prior approaches.

In parallel to the development of these new devices, the scaling process also leads to the appearance of quantum mechanical effects that should be carefully addressed and which are currently under deep study. As an example, the subband quantization of the conduction and valence bands in TFETs may degrade the band–to–band tunneling mechanism by increasing the tunneling barrier for carrier injection.

This thesis is focused on the study of different aspects of both proposed devices from the point of view of simulating possibilities given by the commercial simulator ATLAS from Silvaco [9]. The simulations were mostly performed with the 5.18.3.R version (a prior version, 5.17.47.C, was used for the initial simulations at the early stage of our work, which did not represent any change in the obtained results).

In the case of SB–MOSFETs, we dealt with a non completely solved problem such as the implementation of barrier lowering processes that affect the physical mechanisms of current transport in these transistors. Concerning TFETs, we focused on the inclusion of quantum confinement in devices with ultrathin body thicknesses and how this confinement modified —in some cases, quite significantly— their performance by means of new trends in threshold voltages and subthreshold swings.

This thesis is organized as follows:

- Detailed descriptions of SB–MOSFETs and TFETs along with their principles of operation are included in Chapters 2 and 4, respectively.

- In Chapter 3, we improve the capabilities of ATLAS by proposing a subtle iterative mechanism that makes possible the inclusion of barrier lowering mechanisms in SB–MOSFETs applied not only to thermionic emission but also to thermionic field emission and field emission. Simulation results are shown and compared to existing experimental data in the case of SB–MOSFETs on SOI.

- Chapter 5 explains how the inclusion of confinement in TFETs is incorporated in ATLAS along with the non–local band–to–band tunneling model for carrier injection. This treatment is in principle not allowed in the simulator due to software limitations. In our work, we propose a simple approach to circumvent this obstacle that (*i*) provides a potentially useful tool to gain an insight into quantum mechanical effects in TFETs, and (*ii*) allows to keep using ATLAS, which is an accessible commercial simulator for most researchers.

- In Chapter 6, we analyze the effect of quantum confinement over threshold voltages and subthreshold swings in TFETs, finding new behaviours of these parameters that should be taken into account in further developments of these devices.

- Finally, in Chapter 7, the main conclusions are drawn along with some future threads that naturally ensue from the work herein presented. A list of the publications yielded by this thesis can be found at the end.

Chapter 2

Schottky Barrier MOSFETs (SB–MOSFETs)

As a consequence of the fact that conventional MOSFETs are reaching their performance limitations, new devices are now being investigated to replace them. One of such promising devices to replace conventional MOSFETs is the so–called Schottky Barrier MOSFET (SB–MOSFET), which is based in the replacement of both doped source and drain regions by metallic contacts, thus creating rectifying metal–semiconductor junctions or *Schottky barriers*, SB. Usually, these metallic contacts are metal–silicides for fabrication process reasons.

2.1 Main benefits

There are numerous motivations for replacing doping with metal in the source/drain, S/D, regions. An initial and obvious reason is the necessity of new device structures that are better suited to handle scaling effects in the sub–100nm regime. In this context, and due to the atomically abrupt junctions formed at the silicide–silicon interface, it becomes clear that the inherent physical scalability of SB–MOSFETs would potentially allow to go down to sub–10nm gate length dimensions [5]. Recall that conventional MOSFETs required ever increasing doping concentrations in order to guarantee S/D–to–channel junction abruptness.

The reduction of S/D resistance, even for shallow junctions, is another major advantage of these devices. In MOSFETs, as junction depths are scaled to below 50nm, S/D series resistances become increasingly significant due to the reduction in cross–sectional area. This effect clearly counteracts the benefits that in principle may be offered by scaling. This is particularly true for thin–body structures (for example, ultrathin–body MOSFETs, FinFETs, or nanowire MOSFETs). It was shown [13] that scaling junctions to below 30nm results in little or no performance benefits, since the increase of the S/D resistance offsets any improvement gained by scaling the device.

The SB between metallic S/D and semiconducting channel provides a potential barrier to carriers in the OFF state, thus allowing greater control over the OFF–state leakage current in short channel devices [5].

Potential formation of parasitic bipolar latch-up is completely eliminated in SB–MOSFETs because of the presence of metallic S/D [14–16]. In conventional doped MOSFETs, bipolar transistors may be formed between adjacent devices and lead to parasitic conduction [12].

SB–MOSFETs are also valuable because they show an increased immunity to process variation due to the S/D absence of dopants, and because of the low temperature processing for S/D formation. Regarding this last feature, note that conventionally doped S/D devices require high temperature RTAs of about 900°C for dopant activation in source and drain implants. The problem is that this proves to be incompatible with proposed high–κ gate dielectrics required for further gate oxide scaling, which are damaged by such a high temperature treatment [5]. Silicides like PtSi or ErSi form at much lower temperatures of the order of 500°C and lower, and this allows to fulfill fabrication requirements for high–κ dielectrics. In general, it can be stated that fabrication of SB–MOSFETs require fewer processing steps.

One more issue concerns the elimination of the gate–to–S/D overlap that had to be introduced in doped MOSFETs to prevent current from spreading to lower doping locations in the S/D extensions. In fact, it was shown [13] that a reduction of the overlap caused saturation current to be degraded. Contrary to that, in SB–MOSFETs not only the overlap

is no longer required, but the presence of an underlap between the gate and the S/D electrodes is quite advisable. The absence of any overlap eliminates as well parasitic capacitances.

2.2 Limitations and improvements

The greatest strength of SB–MOSFETs may also be their greatest weakness. With metallic S/D regions, the SB height, SBH, proves to be independent of the channel length, L_{ch}. This may be regarded as a major advantage concerning OFF–state thermal leakage for some silicides; however, larger barrier heights have been demonstrated to limit the drive current and subthreshold swing, SS [17, 18]. The use of mid–gap silicides provides SBH of approximately half the silicon bandgap, nevertheless this results in extremely poor saturation drive current [51] and high subthreshold leakage current due to high gate–induced drain leakage (GIDL) and junction leakage. The employment of complementary silicides turns out to be much appealing, as they provide two different complementary SBH. In the case of n–channel and p–channel SB–MOSFETs, a low SB to electrons, ϕ_{bn}, and to holes, ϕ_{bp} is required, respectively. Silicides such as PtSi for p–channel SB–MOSFETs and rare–earth silicides for n–channel SB–MOSFETs such as $ErSi_x$ or $YbSi_x$ provide the lowest known SBH, with ϕ_{bp} of order 0.15–0.27eV [19, 20]; and ϕ_{bn} for $ErSi_x$ and $YbSi_x$ of order 0.27–0.36eV [21, 22].

Another feature that currently deserves a thorough study is the Fermi level pinning at the S/D–to–channel interface. This pinning comes from a neutral level, ϕ_0, that appears as a result of interfacial states [3]. Experimentally, ϕ_0 is generally located about one–third of the bandgap above the valence band edge [37]. Obviously, this results in relatively low SBH for holes and higher SBH for electrons (about twice the hole barrier height). For that reason, the performance of n–channel SB–MOSFETs has traditionally remained inferior to p–channel devices. Most of the current research involves possible solutions to *de–pin* the Fermi level, although that may imply some additional fabrication steps. An example could be the use of interfacial layers [1, 23], or the use of Group

VI valence–mending adsorbates [24–28] with promising results. However, the interfacial layer approach involves higher process complexity and also limits I_{on} due to the tunnel barrier imposed by the interfacial layer. The use of Group VI elements is equivalent to using dopants to form an S/D extension, SDE, at the metal–semiconductor interface since both have the same effect on the interface dipole [28, 29]. Using dopants at the interface has an added benefit of further reducing the SBH through barrier lowering, BL, processes if the SDE region is long enough or doped highly enough so that it is not fully depleted. There exist interesting works on SB–MOSFETs with SDE regions [30] and also the effects on them of random dopant fluctuations, which become more significant with decreasing contact area [31].

2.3 Historical overview

It was Y. Nishi who first proposed the idea of completely replacing doped S/D with metal in 1966 when he submitted a Japanese patent on this idea, which was later issued in 1970 [52]. However, the first publication was by Lepselter and Sze in 1968, focusing on a PMOS bulk device employing PtSi for the S/D regions [53]. The device was plagued by poor performance with room temperature drive current ten times lower than that of a conventional MOSFET. After that, the next publication to appear was in 1981, when Koeneke showed how the lateral gap between the edge of the S/D electrodes and the gate greatly affected the drive current of the device: the smaller the gap, the higher the performance [54, 55].

Later in the 1980s, a variety of SB–MOS structures were studied: the first SB–NMOS device by Mochizuki and Wise [56], devices using interfacial doping layers between the metallic S/D and the channel [14, 55, 57], and asymmetric devices in which the source is metal and the drain is doped silicon [58, 59]. This existing literature prior to 1994 manifestly suffered from low performance due to the device architecture and process–technology issues.

In 1994 and 1995, Tucker et al. [60, 61] and Snyder et al. [62] pub-

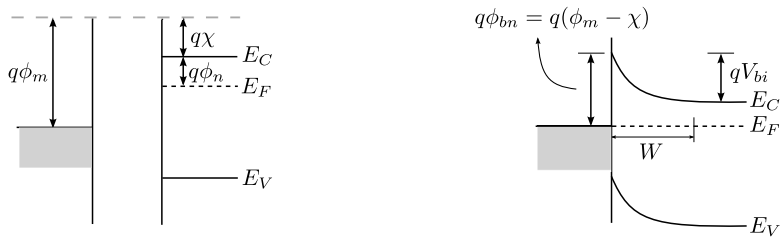

Figure 2.1: Energy band diagrams modification in an ideal metal–semiconductor junction. Left: before the contact formation. Right: in equilibrium with the SB for electrons formed.

lished on the advantages of SB–MOSFETs for device scaling. Since then, research on SB–MOSFETs has considerably increased. Electrical characteristics of p–channel [10, 19, 63–70] and n–channel devices [21, 71–74] with a variety of different S/D silicides have been investigated. The use of BH modification techniques —as mentioned in previous section— to improve device performance has also attracted a great interest [1, 23, 75–81], and a number of simulations have been performed [30, 31, 82–88].

2.4 Metal–semiconductor junction

As the cornerstone of SB–MOSFETs is the replacement of doped S/D by metallic contacts, the study of the SB formation at the metal– semiconductor junction as well as the effects that may lead to its modification (reducing it) are herein described in this section

2.4.1 Formation of the Schottky Barrier

2.4.1.1 Ideal framework

We first deal with the analysis of the ideal formation of a metal– semiconductor contact without taking into account the existence of possible interfacial states and other anomalies. In Fig. 2.1, we observe the changes that arise when putting in contact the metal (left side of the junction) and the semiconductor (right side). The resulting band structure de-

pends on the workfunction[1] of the metal, $q\phi_m$, and the workfunction of the semiconductor, $q(\chi+\phi_n)$ [34]. In this last expression, $q\chi$ stands for the electron affinity measured from the bottom of the conduction band, E_C, to the vacuum level; whereas $q\phi_n$ represents the separation between E_C and the Fermi level. When equilibrium is reached, the Fermi levels on both sides line up as a consequence of a charge transfer from one side of the junction to the other. In the formation of a SB for electrons, the Fermi level in the semiconductor is lowered by an amount which is equal to the difference between both workfuntions. The value of the built–in potential, V_{bi}, of Fig. 2.1 matches that potential difference

$$qV_{bi} = q\phi_m - q\left(\chi+\phi_n\right). \tag{2.1}$$

In Fig. 2.1, the SB height, SBH, is given by

$$q\phi_{bn} = q\left(\phi_m-\chi\right), \tag{2.2}$$

On the other hand, the ideal contact between a metal and a p–type semiconductor produces the corresponding SB for holes, which is given by

$$q\phi_{bp} = E_g - q\left(\phi_m-\chi\right). \tag{2.3}$$

In general, metals have a value for $q\phi_m$ which is within the range $2-6$eV. The ideality of these expressions is, however, deteriorated by the interplay of one or more factors: the existence of an interfacial layer between the metal and semiconductor, the appearance of interfacial states at the junction, or the consideration of the so–called barrier lowering processes that will be later explained.

In a metal–semiconductor junction, the formation of the depletion layer resembles a p^+–n junction. When the metal and the semiconductor

[1] Which is the energy diference between the vacuum level and the Fermi level.

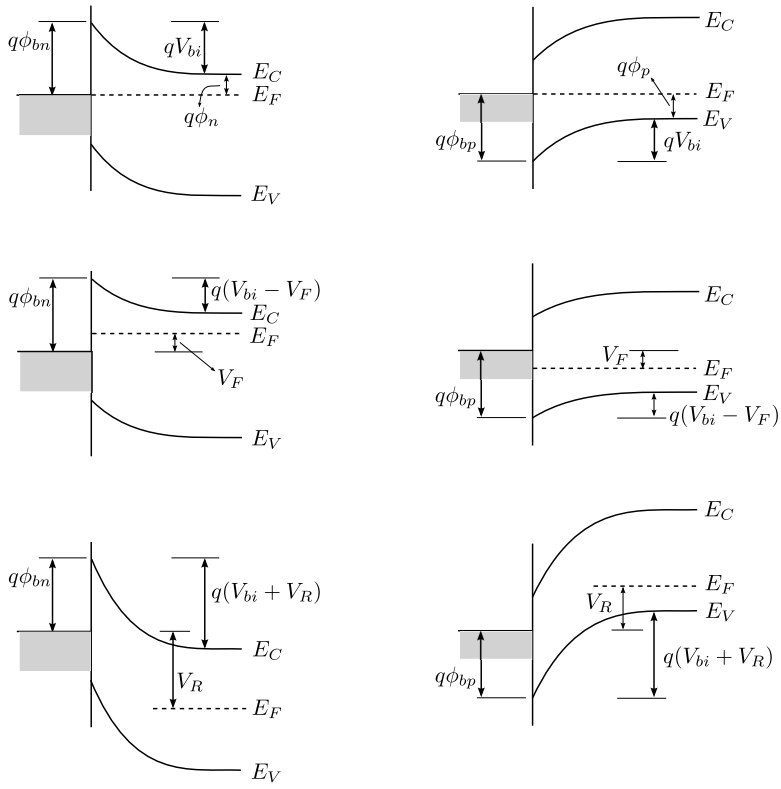

Figure 2.2: Energy band diagrams of an ideal metal–semiconductor junction depending on the bias conditions. Left: with a n–type semiconductor. Right: with a p–type semiconductor..

are brought into intimate contact, the conduction and valence bands of the semiconductor bend according to how the Fermi level is placed in the metal. No band bending occurs inside the metal. Once the relationship between energies is established, it serves as a boundary condition to the solution of the Poisson equation inside the semiconductor. The energy band diagrams for junctions with n–type and p–type semiconductors along with different biasing conditions can be seen in Fig. 2.2.

For contacts involving n–type semiconductors, if we consider that

$$\begin{cases} \rho \approx qN_D & \text{for} \quad x < W_D \\ \rho \approx 0,\ F \approx 0 & \text{for} \quad x > W_D \end{cases}, \tag{2.4}$$

with ρ the charge density, N_D the concentration of donor impurities, F

the electric field and W_D the depletion width; we obtain [3]

$$W_D = \sqrt{\frac{2\varepsilon_s}{qN_D}\left(V_{bi} - V - \frac{kT}{q}\right)}, \tag{2.5}$$

$$|F(x)| = \frac{qN_D}{\varepsilon_s}(W_D - x) = F_m - \frac{qN_D x}{\varepsilon_s}, \tag{2.6}$$

$$E_C(x) = q\phi_{bn} - \frac{q^2 N_D}{\varepsilon_s}\left(W_D x - \frac{x^2}{2}\right), \tag{2.7}$$

where the term kT/q arises from the contribution of the majority–carrier distribution tail (see [3]), V_{bi} is the built–in potential, and F_m is the maximum of the electric field which occurs at $x = 0$

$$F_m = F(x=0) = \sqrt{\frac{2qN_D}{\varepsilon_s}\left(V_{bi} - V - \frac{kT}{q}\right)} = \frac{2\left[V_{bi} - V - (kT/q)\right]}{W_D}. \tag{2.8}$$

The space charge Q_{sc} per unit area of the semiconductor and the depletion layer capacitance C_D per unit area are given by

$$Q_{sc} = qN_D W_D = \sqrt{2q\varepsilon_s N_D\left(V_{bi} - V - \frac{kT}{q}\right)}, \tag{2.9}$$

$$C_D = \frac{\varepsilon_s}{W_D} = \sqrt{\frac{q\varepsilon_s N_D}{2\left[V_{bi} - V - (kT/q)\right]}}. \tag{2.10}$$

2.4.1.2 Interfacial states

The next step in the description of the metal–semiconductor junction is to consider the existence of interfacial states that affect the SBH. This influence is based on two assumptions:

(i) With intimate contact between the metal and the semiconductor, and with an interfacial layer of atomic dimensions (δ), this layer will be transparent to electrons but can withstand potential across

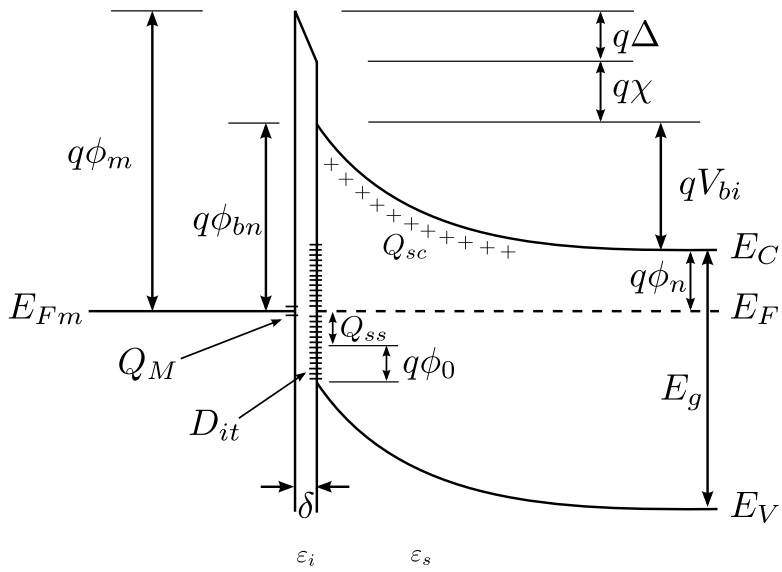

Figure 2.3: Energy band diagrams of a metal–semiconductor (n–type) junction with an interfacial layer and interface states on the semiconductor surface.

it.

(ii) The interface states per unit area per energy are a property of the semiconductor surface and are independent of the metal.

A detailed diagram of a metal–semiconductor (n–type) junction illustrating this situation is depicted in Fig. 2.3 [37, 89]

The first quantity of interest is the energy level $q\phi_0$ above E_V at the semiconductor surface. It is called the neutral level above which the states are acceptor type (thus neutral when empty, and negatively charged when full), and below which the states are donor type (neutral when full, positively charged when empty). Therefore, when the Fermi level at the surface coincides with this neutral level, the net interface–trap charge (Q_{ss} in the figure) is zero [90]. This energy level also tends to pin the semiconductor Fermi level at the surface.

Following the description of Fig. 2.3, Δ is the potential drop across the interfacial layer; Q_M the surface charge density on the metal; ε_i and ε_s the permittivities of the interfacial layer and the semiconductor, respectively; D_{it} the interface–trap density and $q\phi_{bn}$ is the SB for elec-

trons, which is the barrier that must be surmounted for those electrons coming from the metal. The interfacial layer is assumed to have a thickness of a few angstroms (δ in the figure) and is essentially transparent to electrons.

In the proposed situation of Fig. 2.3, the Fermi level of the semiconductor is above the neutral level, which means that some of the acceptor interface traps are full and therefore a negative charge density Q_{ss} is present at the semiconductor surface. Assuming that D_{it} (with units $\mathrm{cm}^{-2}\mathrm{eV}^{-1}$) is constant over the energy range from $E_V + q\phi_0$ to the Fermi level, Q_{ss} is given by

$$Q_{ss} \quad = \quad -qD_{it}\left(E_g - q\phi_0 - q\phi_{bn}\right). \tag{2.11}$$

The space charge that forms in the depletion layer of the semiconductor at thermal equilibrium is given as

$$Q_{sc} \quad = \quad qN_D W_D = \sqrt{2q\varepsilon_s N_D \left(\phi_{bn} - \phi_n - \frac{kT}{q}\right)}. \tag{2.12}$$

The total equivalent surface charge density on the semiconductor side is given by the sum of Eqs. 2.11 and 2.12. In the absence of any space–charge effects in the interfacial layer, an exactly equal and opposite charge (what we labelled as Q_M) develops on the metal surface

$$Q_M \quad = \quad -\left(Q_{ss} + Q_{sc}\right). \tag{2.13}$$

The potential drop across the interfacial layer can be obtained by applying Gauss' law to the surface charge on the metal and semiconductor

$$\Delta \quad = \quad -\frac{\delta Q_M}{\varepsilon_i}. \tag{2.14}$$

Another expression for Δ may be obtained by inspection of the energy band diagram which yields

$$\Delta \;=\; \phi_m - (\chi + \phi_{bn})\,. \tag{2.15}$$

If Δ is eliminated from Eqs. 2.14 and 2.15, and we use Eq. 2.13 to substitute for Q_M, we get

$$\begin{aligned}\phi_m - \chi - \phi_{bn} \;=\; & \sqrt{\frac{2q\varepsilon_s N_D \delta^2}{\varepsilon_i^2}\left(\phi_{bn} - \phi_n - \frac{kT}{q}\right)} - \\ & - \frac{qD_{it}\delta}{\varepsilon_i}\left(E_g - q\phi_0 - q\phi_{bn}\right).\end{aligned} \tag{2.16}$$

Let us now define the following parameters

$$c_1 = \frac{2q\varepsilon_s N_D \delta^2}{\varepsilon_i^2} \quad , \quad c_2 = \frac{\varepsilon_i}{\varepsilon_i + q^2\delta D_{it}}, \tag{2.17}$$

which contain all the interfacial properties. If, for example, we consider normal values of $\varepsilon_s \approx 10\varepsilon_0$, $\varepsilon_i = \varepsilon_0$ and $N_D < 10^{18}\mathrm{cm}^{-3}$; then c_1 is small ($c_1 < 0.003\mathrm{V}$), and the square–root term of Eq. 2.16 can be neglected. In that case, ϕ_{bn} proves to be [91]

$$\phi_{bn} \;=\; c_2\left(\phi_m - \chi\right) + (1 - c_2)\left(\frac{1}{q}E_g - \phi_0\right). \tag{2.18}$$

Or, rearranging this expression, we can make more explicit the dependence on ϕ_m and group the rest of the terms involving constants in a new parameter, c_3,

$$\phi_{bn} \;=\; c_2\phi_m + c_3. \tag{2.19}$$

If c_2 and c_3 are estimated from experimental results, the interfacial properties may be derived as

$$\phi_0 = \frac{1}{q}E_g - \frac{c_2\chi + c_3}{1 - c_2} \tag{2.20}$$

$$D_{it} = \frac{(1 - c_2)\varepsilon_i}{c_2\delta q^2}. \tag{2.21}$$

Some examples of experimental values for c_2, c_3 and χ when considering different materials are shown in Table 2.1. In Table 2.2, their extracted interfacial parameters are also listed. Observe that the values of $q\phi_0$ for Si, GaAs and GaP are very close to one–third of the bandgap. Similar behaviour may be found in other semiconductors [92].

In light of Eq. 2.18, there are two extreme cases that are easily parametrized using c_2:

(1) If $c_2 \to 0$, then

$$q\phi_{bn} = E_g - q\phi_0 \tag{2.22}$$

In this case, the Fermi level at the interface is pinned by the surface states at the value $q\phi_0$ above the valence band. The SBH is independent of the metal workfunction [92] and is determined entirely by the surface properties of the semiconductor.

(2) If $c_2 \to 1$, then

$$q\phi_{bn} = q\left(\phi_m - \chi\right), \tag{2.23}$$

Semiconductor	c_2	c_3(V)	χ(V)
Si	0.27 ± 0.05	-0.52 ± 0.22	4.05
GaAs	0.07 ± 0.05	0.51 ± 0.24	4.07
GaP	0.27 ± 0.03	0.02 ± 0.13	4.0
CdS	0.38 ± 0.16	-1.17 ± 0.77	4.8

Table 2.1: Summary of barrier height data for different materials (extracted from [91]).

Semiconductor	$D_{it}(10^{13}\text{cm}^{-2}\text{eV}^{-1})$	$q\phi_0(\text{eV})$	$q\phi_0/E_g$
Si	2.7 ± 0.7	0.30 ± 0.36	0.27
GaAs	12.5 ± 10.0	0.53 ± 0.33	0.38
GaP	2.7 ± 0.4	0.66 ± 0.2	0.294
CdS	1.6 ± 1.1	1.5 ± 1.5	0.6

Table 2.2: Interfacial properties extracted from experimental results.

which corresponds to the ideal case of Eq. 2.2 where surface states are neglected.

2.4.2 Carrier transport

Current transport in metal–semiconductor junctions is essentially determined by majority carriers. Our aim in this section is to analyze which are the main mechanisms for carrier transport under forward bias (analogous treatment may be applied for reverse biasing by inverting those mechanisms). Depending on the bias conditions, one or more of them will dominate. Let us briefly outline them first:

(i) Thermionic emission of electrons from the semiconductor over the potential barrier into the metal. Depending on the case, this process may be limited by drift–diffusion mechanisms. For that reason, separate treatments for both processes (thermionic emission and drift–diffusion) were developed in the literature [32–34], as well as a synthesis of both that intended to be an unified approach [93]. As a general rule, the current is determined by whichever mechanism that causes the larger impediment to carrier flow.

(ii) For electrons with energies below the barrier and over the Fermi level (i.e. thermally excited), it becomes possible to traverse the barrier by quantum mechanical tunneling. This process is known as thermionic–field emission [34], and is highly dependent on the temperature because as we increase the temperature, so does the number of carriers at high energies.

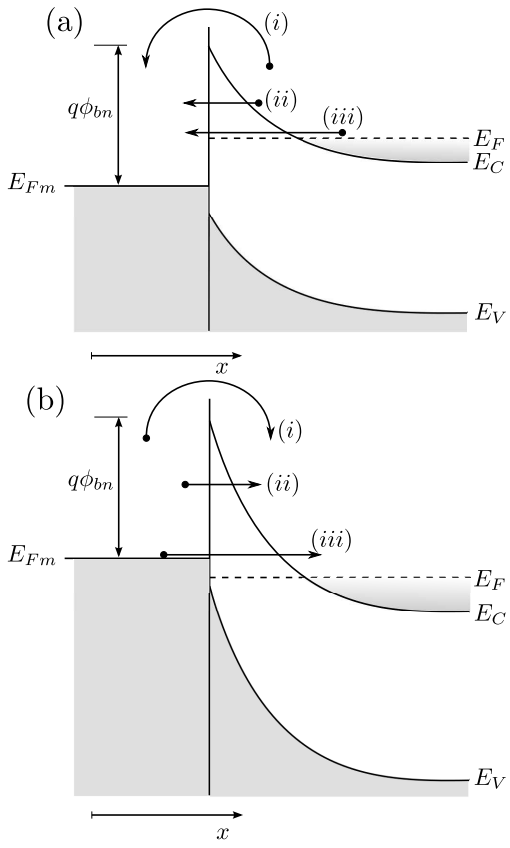

Figure 2.4: Main processes for electrons to be injected into the metal from the semiconductor in a metal–semiconductor (degenerated n–type) junction forward biased (a), and into the semiconductor from the metal when reverse biased (b). The depicted processes are: thermionic emission (i), thermionic–field emission (ii), and field emission from the Fermi level (iii).

(iii) Electrons with energies close to the Fermi level can also tunnel through the barrier into the metal. We refer to this mechanism simply as field emission.

These three mechanisms are summarized in Fig. 2.4. Notice that field emission might only happen if the semiconductor is degenerated.

2.4.2.1 Thermionic emission

Thermionic emission turns out to be the dominant process for common high–mobility semiconductors (e.g., Si and GaAs) with moderate dopings and operating in a range of moderate temperatures. When thermionic emission dominates, the shape of the barrier profile is immaterial and the current flow depends solely on the barrier height. The current density from the semiconductor to the metal, $J_{s\to m}$, is given by the concentration of electrons with energies sufficient to overcome the potential barrier and traversing in the x–direction

$$J_{s\to m} = \int_{E_{Fm}+q\phi_{bn}}^{\infty} qv_x dn, \tag{2.24}$$

where $E_{Fm}+q\phi_{bn}$ is the minimum energy required for thermionic emission into the metal, v_x is the carrier velocity in the direction of transport, and dn is the electron density in an incremental energy range. The evaluation of the integral [3] leads to

$$J_{s\to m} = A^*T^2 \exp\left(-\frac{q\phi_{bn}}{kT}\right) \exp\left(\frac{qV}{kT}\right), \tag{2.25}$$

with

$$A^* = \frac{4\pi q m^* k^2}{h^3}, \tag{2.26}$$

the so–called Richardson constant, where m^* is the effective mass for electrons.

Since the SBH for electrons moving from the metal into the semiconductor remains the same under bias, the current flowing into the semiconductor is thus unaffected by the applied voltage. It must therefore be equal to the current flowing from the semiconductor into the metal when thermal equilibrium prevails (i.e. when $V = 0$). By doing so, we obtain from Eq. 2.25

$$J_{m\to s} = -A^*T^2 \exp\left(-\frac{q\phi_{bn}}{kT}\right). \quad (2.27)$$

And summing both contributions we get

$$\begin{aligned} J_{n,\,te} &= \left[A^*T^2 \exp\left(-\frac{q\phi_{bn}}{kT}\right)\right]\left[\exp\left(\frac{qV}{kT}\right) - 1\right] \\ &= J_{TE}\left[\exp\left(\frac{qV}{kT}\right) - 1\right], \end{aligned} \quad (2.28)$$

with

$$J_{TE} = A^*T^2 \exp\left(-\frac{q\phi_{bn}}{kT}\right). \quad (2.29)$$

2.4.2.2 Drift–diffusion

In the drift–diffusion model, the expression for the current density in the transport direction (that we assumed to be the x–direction) depends on the local field and the concentration gradient, and is given by

$$\begin{aligned} J_x = J_{n,\,d} &= q\left(n(x)\mu_n F(x) + D_n\frac{dn}{dx}\right) \\ &= qD_n\left(\frac{n}{kT}\frac{dE_C}{dx} + \frac{dn}{dx}\right), \end{aligned} \quad (2.30)$$

where μ_n is the electron mobility, D_n is the diffusion coefficient and $F(x)$ is the electric field. Under steady–state condition, the current density is independent of x, and Eq. 2.30 may be integrated using $\exp[E_C(x)/kT]$ as an integrating factor [3]

$$J_{n,\,d}\int_0^{W_D} \exp\left[\frac{E_C(x)}{kT}\right] dx = qD_n\left\{n(x)\exp\left[\frac{E_C(x)}{kT}\right]\right\}\Big|_0^{W_D}, \quad (2.31)$$

with the appropriate boundary conditions and the expression for $E_C(x)$ obtained in Eq. 2.7. The evaluation leads to

$$\begin{aligned} J_{n,\,d} &\approx q\mu_n N_C F_m \exp\left(-\frac{q\phi_{bn}}{kT}\right)\left[\exp\left(\frac{qV}{kT}\right)-1\right] \\ &= J_D\left[\exp\left(\frac{qV}{kT}\right)-1\right]. \end{aligned} \tag{2.32}$$

It can be noticed that the current densities of the drift–diffusion and thermionic emission theories are basically very similar. They differ in the saturation current densities.

2.4.2.3 Tunneling current

The aforementioned thermionic field emission and field emission processes involve the tunneling of carriers through a potential barrier. The tunneling current from semiconductor to metal, $J_{s\to m}$ is proportional to the transmission probability (tunneling probability) multiplied by the occupation probability in the semiconductor and the unoccupied probability in the metal [94]

$$J_{s\to m} = \frac{A^* T^2}{kT}\int_{E_{Fm}}^{q\phi_{bn}+E_{Fm}} f_s(E)T(E)\left[1-f_m(E)\right]dE, \tag{2.33}$$

with f_s and f_m the Fermi distribution functions for the semiconductor and metal, respectively. $T(E)$ is the tunneling probability which depends on the height and the width of the barrier at a particular energy. A similar expression can be given for the current in the opposite direction $J_{m\to s}$. In that case, Fermi distributions would be interchanged. The resulting total current is the algebraic sum of those two components.

As we previously indicated, while field emission is a pure tunneling process, thermionic field emission is tunneling of thermally excited carriers which, consequently, see a thinner barrier than those with energies close to the Fermi level. The relative contributions of these components depend on both temperature and doping level. A rough criterion

can be set by comparing the thermal energy, kT, to a reduced energy E_{00} which is a function of the semiconductor doping, effective mass and permittivity given by [3, 34]

$$E_{00} = \frac{q\hbar}{2}\sqrt{\frac{N}{m^*\varepsilon_s}}. \tag{2.34}$$

When $kT >> E_{00}$, thermionic emission dominates over tunneling contributions. When $kT << E_{00}$, field emission represents the main contribution to the current. Finally, when $kT \approx E_{00}$, thermionic field emission is the main mechanism.

Under forward bias, the current due to field emission can be expressed as [35]

$$\begin{aligned} J_{n,\,fe} &= \frac{A^*T\pi \exp\left[-q\frac{(\phi_{bn}-V_F)}{E_{00}}\right]}{c_1 k \sin(\pi c_1 kT)}\left[1-\exp(-c_1 qV_F)\right] \\ &\approx \frac{A^*T\pi \exp\left[-q\frac{(\phi_{bn}-V_F)}{E_{00}}\right]}{c_1 k \sin(\pi c_1 kT)}, \end{aligned} \tag{2.35}$$

with

$$c_1 = \frac{1}{2E_{00}}\log\left[\frac{4(\phi_{bn}-V_F)}{-\phi_n}\right]. \tag{2.36}$$

Note that ϕ_n is negative for degenerate semiconductors. The corresponding expression for thermionic field emission is [3]

$$\begin{aligned} J_{n,\,tfe} &= \frac{A^*T\sqrt{\pi E_{00}q(\phi_{bn}-\phi_n-V_F)}}{k\cosh\left(\frac{E_{00}}{kT}\right)}\exp\left[\frac{-q\phi_n}{kT}-\frac{q(\phi_{bn}-\phi_n)}{E_0}\right] \\ &\quad \exp\left(\frac{qV_F}{E_0}\right), \end{aligned} \tag{2.37}$$

$$E_0 = E_{00} \coth\left(\frac{E_{00}}{kT}\right). \tag{2.38}$$

For thermionic field emission, note that population of thermally excited electrons decrease as we consider higher energies. In that sense, one would expect the contribution of thermionic field emission to monotonically decrease with energy. However, that is not the case because the barrier that electrons have to traverse is also reduced for increasing energies. As a result, there exist an energy, E_m, for which thermionic field emission roughly peaks. This energy is measured from E_C of the neutral region and proves to be

$$E_m = \frac{q(\phi_{bn} - \phi_n - V_F)}{\cosh^2\left(\frac{E_{00}}{kT}\right)}. \tag{2.39}$$

Under reverse bias, the tunneling current can be much larger because a high voltage is possible. In that case, the currents corresponding to field emission and thermionic field emission are given by

$$J_{n,\,fe} = A^* \left(\frac{E_{00}}{k}\right)^2 \left(\frac{\phi_{bn} + V_R}{\phi_{bn}}\right) \exp\left(-\frac{2q\phi_{bn}^{3/2}}{3E_{00}\sqrt{\phi_{bn} + V_R}}\right) \tag{2.40}$$

$$J_{n,\,tfe} = \frac{A^*T}{k}\sqrt{\pi E_{00} q\left[V_R + \frac{\phi_{bn}}{\cosh^2\frac{E_{00}}{kT}}\right]} \exp\left(-\frac{q\phi_{bn}}{E_0}\right) \exp\left(\frac{qV_R}{\tilde{\varepsilon}}\right), \tag{2.41}$$

with

$$\tilde{\varepsilon} = \frac{E_{00}}{\frac{E_{00}}{kT} - \tanh\left(\frac{E_{00}}{kT}\right)}. \tag{2.42}$$

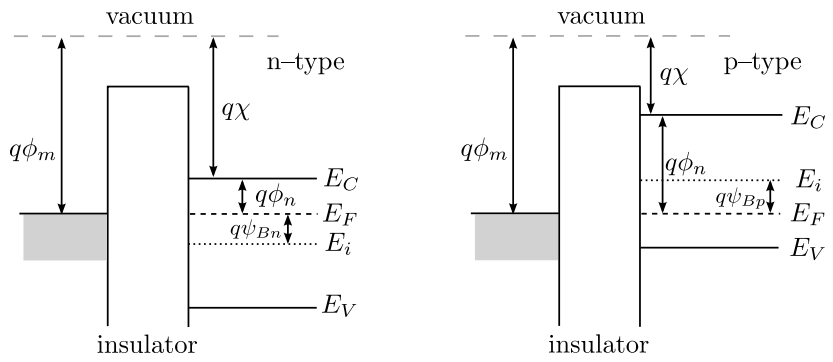

Figure 2.5: Band diagrams of ideal MOS capacitors at equilibrium ($V = 0$). Left: n–type semiconductor; right: p–type.

2.5 The MOS capacitor

The metal–insulator–semiconductor structure is schematically presented in Fig. 2.5 for both n–type and p–type semiconductors. For the sake of simplicity, we assume that the metal has been chosen such that there is no difference between the Fermi levels of the metal and the semiconductor.

And ideal MOS structure satisfies: (*i*) The only charges that may exist under any biasing condition must be located within the semiconductor, and on the metal surface adjacent to the insulator (i.e., no interface charged traps nor any kind of oxide charge); (*ii*) There is no carrier transport through the insulator under dc biasing conditions.

Given that the metal was chosen to have the same workfunction as the semiconductor, their difference must be zero and therefore the following relationships between the magnitudes of Fig. 2.5 hold

$$\phi_m - \left(\chi + \frac{E_g}{2q} - \psi_{Bn}\right) = \phi_m - (\chi - \phi_n) = 0 \qquad \text{for n–type,} \tag{2.43}$$

$$\phi_m - \left(\chi + \frac{E_g}{2q} + \psi_{Bp}\right) = \phi_m - \left(\chi + \frac{E_g}{q} - \phi_p\right) = 0 \qquad \text{for p–type.} \tag{2.44}$$

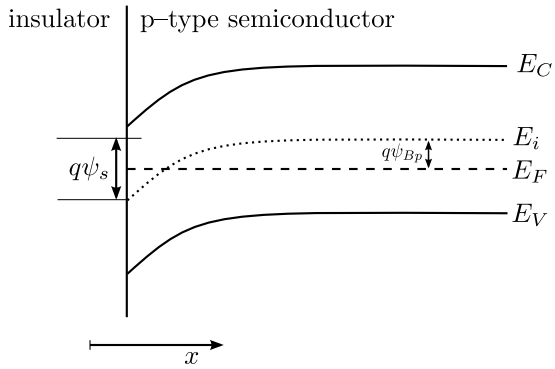

Figure 2.6: Energy band diagrams illustrating the potential ψ_s with respect to the intrinsic Fermi level in the bulk, and the potential ψ_{Bp}.

The Fermi potentials ϕ_n and ϕ_p were already shown in Fig. 2.2 as the difference between the Fermi level and the band edges (conduction band for n–type semiconductors, and valence band for p–type). In addition to those, the potentials ψ_{Bn} and ψ_{Bp} are now introduced [3] to describe, inside the bulk, the difference of the Fermi level with respect to the midgap. When a voltage V is applied to the gate, and as a consequence of the band bending, new position–dependent potentials, $\psi_n(x)$ and $\psi_p(x)$, may be defined. They account for the potential with respect to the bulk of the semiconductor. For example (see Fig. 2.6), for a p–type semiconductor

$$\psi_p(x) \quad = \quad -\frac{[E_i(x) - E_i(\infty)]}{q}, \tag{2.45}$$

with the shortened notation $\psi_p(x = 0) = \psi_s$.

When the MOS capacitor is biased with positive or negative voltages, three main cases may arise at the semiconductor surface. They are shown in Fig. 2.7 for both n–type and p–type semiconductors. Let us consider the p–type semiconductor to illustrate them:

(i) When a negative voltage ($V < 0$) is applied to the gate, the valence band edge E_V bends upward near the surface and is closer to the Fermi level. Since the carrier density depends exponentially

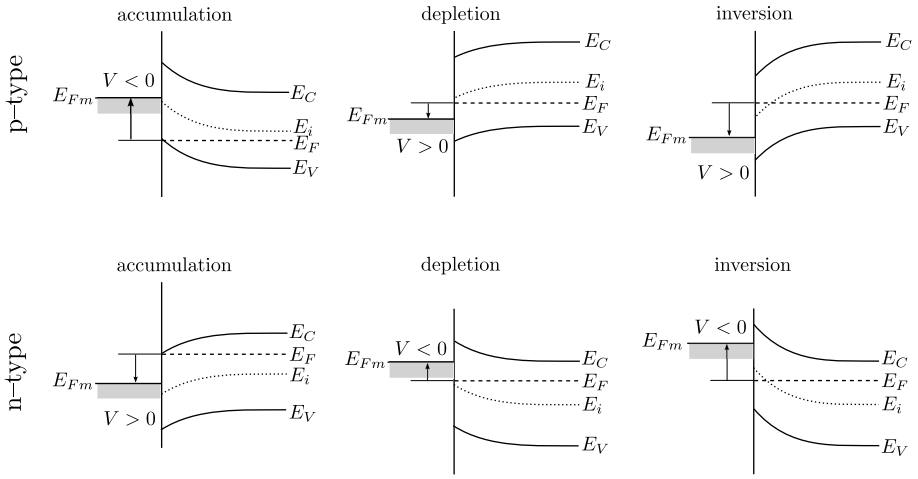

Figure 2.7: Energy band diagrams of the MOS structure for n–type and p–type substrates under different bias conditions. Three cases are depicted: accumulation, depletion and inversion.

on the energy difference $E_F - E_V$, this band bending causes an accumulation of holes near the semiconductor surface. This is the *accumulation* case and corresponds to $\psi_s < 0$.

(ii) When a small positive voltage ($V > 0$) is applied, the bands bend downward, and the majority carriers are depleted. This is the *depletion* case, and for it $\psi_{Bp} > \psi_s > 0$.

(iii) When a larger positive voltage is applied, the band bend even more downward so that the intrinsic level E_i at the surface crosses over the Fermi level E_F. At this point, the number of electrons at the surface is larger than that of the holes, the surface is thus inverted. This is the *inversion* case. It corresponds to $\psi_s > \psi_{Bp}$. In this last case, if the applied voltage is big enough, we reach what is commonly known as strong inversion. It happens when $\psi_s > 2\psi_{Bp}$.

Similar reasoning can be followed for the n–type semiconductor in light of the bottom figures of Fig. 2.7 in order to differentiate between the three previous cases.

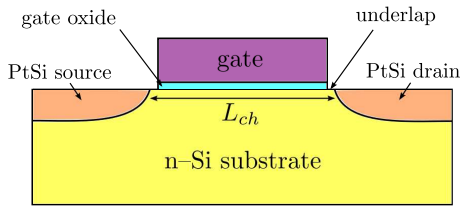

Figure 2.8: Schematic cross–section of a p–channel SB–MOSFET where the source and the drain have been formed using PtSi.

2.6 SB–MOSFETs operating regimes

Let us illustrate the different operating regimes of a p–channel SB–MOSFET in which, for example, the source and the drain are fabricated with PtSi. A cross section of the device is shown in Fig. 2.8. When traversing the device, the holes must travel through metal (source), then semiconductor, then metal again (drain). In the channel, hole motion is governed by drift–diffusion equations, whereas in the source–to–channel and channel–to–drain junctions, the previously discussed processes for metal–semiconductor contacts have to be taken into account.

In Fig. 2.9, we show the band diagrams corresponding to four states of operation of the considered device depending on the bias conditions:

(a) Equilibrium situation with no bias applied.

(b) Reverse leakage current. In this situation, electrons in the drain possess a finite probability of tunneling through the large electron SB into the channel. This effect will be especially pronounced if the Fermi level at the drain is raised above the conduction band in the channel, as this also enables field emission in addition to thermionic field emission. The high electron barrier guarantees that this leakage current is considerably smaller than the hole current. Once in the channel, electrons drift to the source–to–channel interface where they must surmount a second barrier to exit via the source. Nevertheless, since this barrier poses a smaller hindrance to carriers, it is the drain–to–channel junction which controls the reverse leakage current.

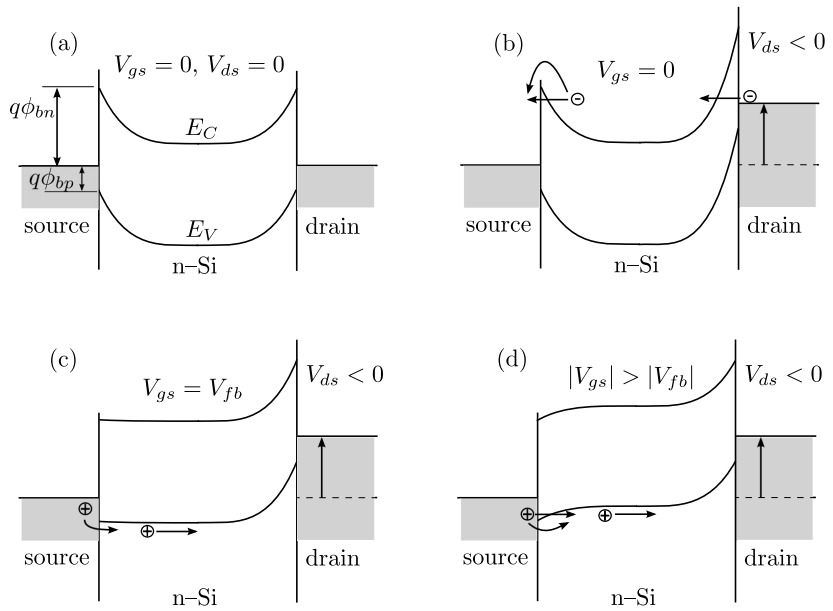

Figure 2.9: Energy band diagrams for a p–channel SB–MOSFET illustrating four states of operation: (a) OFF–state, no drain bias; (b) OFF–state, with drain bias; (c) Flat band condition (only thermionic emission of holes is present before this point is reached); (d) ON–state with tunneling taking place.

(c) In the situation between Figs. 2.9(b) and (c), we deal with the subthreshold regime in which drain current is entirely due to thermionic emission of holes over the source–to–channel barrier [36]. The valence band in the channel remains in a range that prevents any tunneling current from taking place. Fig. 2.9(c) marks the limit where thermionic emission reaches its maximum contribution, as the applied gate voltage causes flat band condition to appear at the source junction.

(d) Raising the valence band beyond flat band condition ($|V_{gs}| > |V_{fb}|$) does not lead to further increase in thermionic emission, but to the beginning of thermionic field emission. In this case, the source–to–channel junction becomes reverse biased and the SB for holes is gradually thinned by the raising of the channel potential. If we continue increasing $|V_{gs}|$, we would reach a point where the valence band in the channel is raised above the Fermi level in the source. At

that moment, purely field emission would also begin to contribute and adds to the thermionic and thermionic field emission.

As carrier injection grows, eventually one has to consider the channel resistance as a limiting factor to current flow. That means that in that case, carrier drift and scattering in the channel would pose a greater hindrance to the drain current than the source–to–channel SB.

2.7 Bulk and SOI SB–MOSFETs

In SB–MOSFETs —as well as conventional MOSFETs— on bulk silicon, the continuous downscaling and the loss of electrostatic control by the gate deteriorate the device performance. In this context, the employment of Silicon–on–Insulator (SOI) technology provides a significant reduction of most parasitic effects observed in bulk silicon devices. Additionally, because the dielectric constant of the buried insulator (BOX, usually SiO_2) is three times smaller than that of silicon, the parasitic capacitances between the S/D junctions and the substrate are strongly reduced.

Analogously to conventional MOSFETs, SB–MOSFETs on SOI may be of two types: fully depleted, in which the silicon film thickness t_{Si} in the channel region is smaller than depletion depth at the threshold voltage; and partially depleted, in which t_{Si} is larger than the depletion depth. For thick SOI films, the devices approach the bulk limit.

In Fig. 2.10, we show an example of the distribution of depletion charges in a short channel SB–MOSFET on partially and fully depleted SOI. Obviously, with decreasing channel length, the gate controlled depletion charge is in both cases reduced. However, the ratio of the trapezoidal hatched area to the whole depletion charge is larger in the fully depleted transistor than in the partially depleted one [95]. Therefore, partially depleted SOI SB–MOSFETs suffer more from short channel effects (SCE) than those fully depleted. Furthermore, the floating body in partially depleted devices can degrade the device performance because of an uncontrolled lowering of the threshold voltage caused by a parasitic bipolar transistor action due to the charges generated by impact

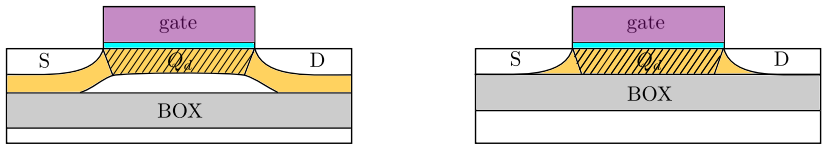

Figure 2.10: Partially depleted SB–MOSFET on SOI (left) vs. fully depleted (right). Q_d represents the depletion charge distribution controlled by the gate.

ionization at high S/D voltages [95]. Fully depleted SB–MOSFETs on SOI are immune to the floating body effect.

In general, ultra–thin body (UTB) SOI offers improved electrostatic gate control and enhanced device performance. However, if the SOI film thickness is reduced below 10nm, quantum mechanical effects become important so that the energy bands split into a discrete set of energy subbands. As a result, the energy of the first subband rises, which increases the effective barrier heights. The conclusion that may be extracted from this is that in both bulk and SOI SB–MOSFETs the silicon thickness has to be chosen with care in order to find the optimum trade–off betweeen the increase of the barrier heights caused by quantum band–splitting and the improved electrostatics with decreasing t_{Si}.

In the case of fully depleted SB–MOSFETs on SOI, extensive research has been done [96–99] into their tunneling subthreshold behaviour, and the following analytical formula has been derived for the SS

$$SS = \frac{kT}{q}\ln 10 \frac{1}{1-\exp\left(-\frac{d}{\lambda}\right)}, \tag{2.46}$$

where λ stands for the screening length (see Eq. 4.15 of Sec. 4.3.2), and d a fitting parameter, or the so–called tunneling distance. In this expression, it is assumed that for dimensions beyond this distance the tunneling probability $T(E)$ of Eq. 2.33 is set to 0 and below this distance to 1 (Fig. 2.11). Regard that for a constant d, a small λ leads to better gate control which ultimately allows to approach the 60mV/dec limit of conventional MOSFETs. The approximations made in the literature to obtain Eq. 2.46 do not allow an exact conclusion to be drawn as to

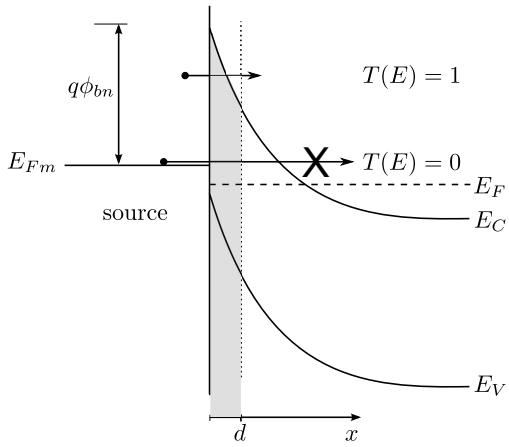

Figure 2.11: Example of d as tunneling distance limit in the framework of Eq. 2.46. Carriers may tunnel through the barrier if the distance that they have to traverse is smaller than d, otherwise tunneling is forbidden.

the true origin of this 60mV/dec limit. Nevertheless, from the available experimental results [99], it is safe to assume that one should not expect SS lower than 60mV/dec when using SB–MOSFETs.

Chapter 3

Implementation of Barrier Lowering in SB–MOSFETs

Apart from the fact that in a metal–semiconductor junction the SBH can be technologically modified by the controlled mechanism of introducing a thin layer (≈10nm or less) of dopants on the metal–semiconductor interface[1], there exist two other mechanisms that alter the height and shape of the barrier which —for different reasons— cannot be controlled and need to be accounted for in order to obtain a detailed description of the carrier transport processes across the SB.

The first of them is the so–called *image force lowering* (IFL) and is a dynamic effect that appears as a consequence of the presence of charge carriers in the vicinity of a metallic surface. It thus has nothing to do with possible imperfections or impurities at the metal–semiconductor junction. It is an intrinsic physical process that the SB undergoes whenever a carrier approaches it, and implies a certain barrier lowering, BL. The second, much less important in practice —unlike some authors originally claimed that it could be as important as IFL [43]— is known as *dipole lowering* (DL). The name comes from the formation of a dipole layer at intimate metal–semiconductor interfaces, which appears because of the states that the metal induces in the semiconductor

[1]For example, by silicidation induced dopant segregation or ion implantation [76, 100–102].

bandgap, known as MIGS (*metal induced gap states*) [42].

3.1 Image force lowering

As mentioned above, IFL is a dynamic effect that lowers the barrier energy for charge carrier emission in the presence of an electric field. When an electron is at a distance x from a metal, a positive charge will be induced on its surface. The force of attraction between the electron and the induced positive charge is equivalent to the force that would exist between the electron and an equal but opposite charge located at a distance $-x$ from the interface. This charge is referred to as the image charge. The attractive force towards the metal and the potential energy of such electron are given by

$$\mathcal{F} = \frac{-q^2}{16\pi\varepsilon_0 x^2}, \qquad E(x) = \int_{\infty}^{x} \mathcal{F}\, dx = -\frac{q^2}{16\pi\varepsilon_0 x}. \tag{3.1}$$

When an external field F is applied (in this case in the $-x$ direction), the total potential energy as a function of distance is

$$E(x) \;=\; -\frac{q^2}{16\pi\varepsilon_0 x} - q|F|x. \tag{3.2}$$

This equation has a maximum value, which turns out to be the original barrier lowered by an amount $\Delta\phi$ given by

$$\Delta\phi_{ifl} \;=\; \sqrt{\frac{q|F|}{4\pi\varepsilon_0}}, \tag{3.3}$$

and located at a distance x_m from the metal surface that proves to be

$$x_m \;=\; \sqrt{\frac{q}{16\pi\varepsilon_0|F|}}. \tag{3.4}$$

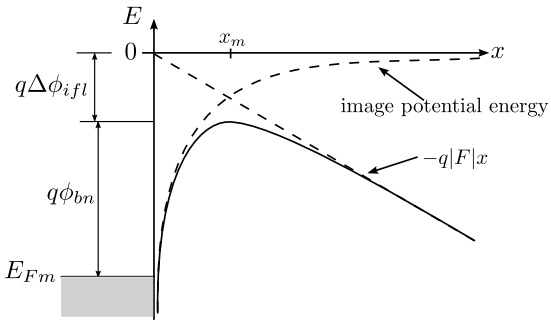

Figure 3.1: Energy band diagram between a metal surface and a vacuum when an electric field is applied. The barrier is effectively lowered by an amount $\Delta\phi_{ifl}$ as a result of IFL.

This BL effect is schematically depicted in Fig. 3.1. If these results are now to be applied to a metal–semiconductor junction, the uniform electric field must be replaced by the appropriate field at the interface, and the dielectric constant of the semiconductor should be used instead of that of the vacuum

$$\Delta\phi_{ifl} = \sqrt{\frac{q|F_m|}{4\pi\varepsilon_s}}, \tag{3.5}$$

with F_m standing for the electric field at the junction, where it reaches its maximum value. Using the depletion approximation, it can be estimated as

$$F_m = \sqrt{\frac{2qN|\psi_s|}{\varepsilon_s}}, \tag{3.6}$$

where N is the substrate doping density. Note that in a metal–semiconductor contact, the electric field is not zero even without applied bias due to the built–in potential. Although the BL may be small, it does have a profound effect on current transport processes given their exponential dependence on the SBH.

If we apply a forward bias, the field and the IFL are reduced, so

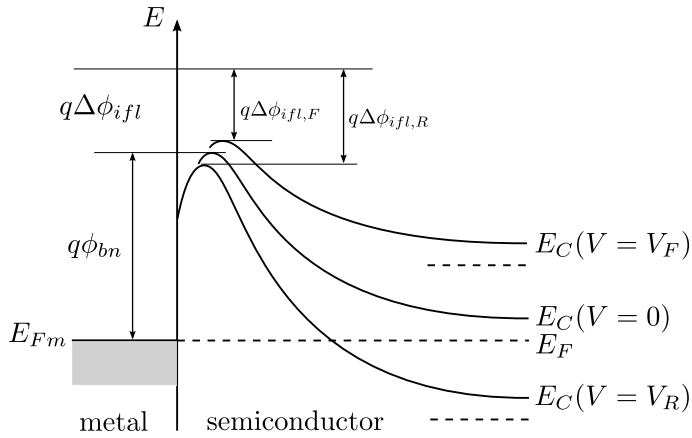

Figure 3.2: Energy band diagram incorporating IFL for a metal–semiconductor (n–type) contact under different biasing conditions. The reduced SBH at thermal equilibrium is $q\phi_{bn}$. The lowerings under forward and reverse bias are $\Delta\phi_{ifl,F}$ and $\Delta\phi_{ifl,R}$, respectively.

that the SBH is slightly larger than that corresponding to zero bias. Conversely, for reverse bias, the SBH is slightly smaller. These effects have been depicted in Fig. 3.2. As a result, it is obvious that the SBH becomes bias dependent.

3.2 Dipole lowering

This additional effect, which may be labelled as dipole lowering, was originally related to the existence of interface states at the semiconductor surface as a result of an oxide layer between the metal and the semiconductor [37–39], or defects associated with the physical interface non–idealities. However, this effect also appears in atomically clean and abrupt interfaces indicating that DL has another contribution coming from the quantum–mechanical solution to the equilibrium charge distribution of an ideal metal–semiconductor junction. This charge distribution shows a certain penetration —the so–called *Heine tails* [40]— of electronic wave functions from the metal into the forbidden gap of the semiconductor [31,41]. The result is the formation of the aforementioned MIGS [40,42], and the appearance of a dipole layer at the interface which

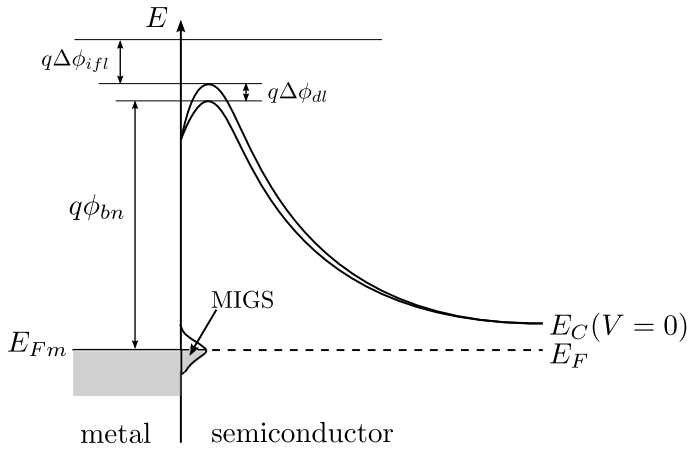

Figure 3.3: Schematic band diagram of a metal–semiconductor junction in which the metal in direct vicinity of the (n–type) semiconductor creates MIGS [1]. The figure includes both IFL and DL through $\Delta\phi_{ifl}$ and $\Delta\phi_{dl}$, respectively.

consequently modifies the SBH.

Fig. 3.3 illustrates how MIGS would be incorporated in a schematic band diagram showing that those states below E_F are filled and above E_F are empty. It should be mentioned that MIGS can either be donor or acceptor states.

An initial approach [43] suggested that DL could be directly proportional to the interface electric field as

$$\Delta\phi_{dl} = \alpha F_m. \tag{3.7}$$

More recently [9], this empirically based dependence has been thought to be better described by

$$\Delta\phi_{dl} = \alpha F_m^{\gamma}, \tag{3.8}$$

where α and γ have to be fitted from experimental results.

An alternative formulation [30] based on the formalism of Heine tails accounts for DL as

$$\Delta\phi_{dl} = \frac{\beta Q_{sc}\lambda_H}{\varepsilon}\exp\left[-\frac{x_m}{\lambda_H}\right], \tag{3.9}$$

with x_m the position of the maximum potential, β is the fraction of ionized dopants on the silicon side of the junction that contributes to the effect, λ_H is the Heine tail length, and Q_{sc} is the areal charge density on the silicon side. Realistic values for λ_H and β have to be fitted from experimental data.

3.3 Device structure

We have studied the behaviour and performance of SB-MOSFETs on SOI with NiSi and epitaxial $NiSi_2$ S/D. These silicides have a measured SBH for electrons of $\phi_{bn} = 0.65$V and $\phi_{bn} = 0.37$V on n-Si(100), respectively [103], which leads to an ambipolar switching behaviour, greater in the case of NiSi. The devices were fabricated at Forschungszentrum Jülich [44] and are schematically depicted in Fig. 3.4.

The fabrication process is reported in [2]. The transistor consists of 15nm SOI channel, a 5nm thick HfO_2 gate dielectric layer, and a TiN gate. The source/drain contacts were formed with 10nm thick NiSi or epitaxial $NiSi_2$ layers by annealing of a 5nm Ni layer at 500°C, and a 3nm Ni layer at 750°C, respectively [44]. The fast diffusion of Ni induces a large encroachment of NiSi after annealing which causes serious variabil-

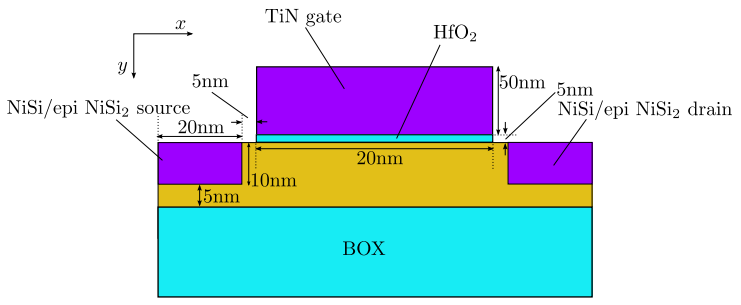

Figure 3.4: Schematic cross-section (not to scale) of the NiSi/epitaxial $NiSi_2$ SB-MOSFETs simulated in this work along with their dimensions.

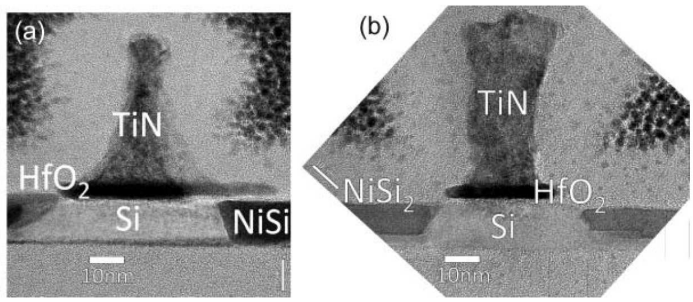

Figure 3.5: Cross–sectional TEM image of 20nm gate length SB–MOSFETs with NiSi and epitaxial $NiSi_2$ source and drain, respectively [2].

ity of the devices when the silicide layer is thick. The ultra–thin silicide formation using very thin Ni layer avoids this problem. In Fig. 3.5, the NiSi and $NiSi_2$ devices are shown with cross–sectional transmission electron microscopy (XTEM) images, which reveal good gate alignment with no encroachment of the silicide into the channel.

A small misalignment is caused due to the shadowing effect of the gate during Ni deposition, resulting in a small gap of 5nm between the channel and the silicide —as schematically shown in Fig. 3.4— which will be considered in our simulations. A limiting factor that may hamper the potential scaling of SB–MOSFETs is the line edge roughness of the

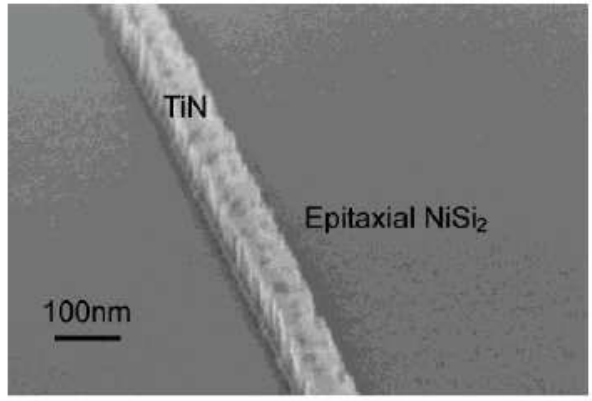

Figure 3.6: Tilted SEM image of epitaxial $NiSi_2$ SB–MOSFET [2]. A straight line edge along the gate is perfectly formed.

silicide along the gate. The employment of ultra–thin silicides helps to overcome this stumbling block, as it can be seen in Fig. 3.6 which shows a scanning electron microscopy (SEM) image of a device structure with epitaxial $NiSi_2$ S/D.

3.4 Experimental data

The experimental results that we will handle hereunder were obtained and reported in [2]. We briefly present some of them. In Fig. 3.7, we depict the experimental transfer characteristics of the studied SB–MOSFETs with both type of silicide contacts for $V_{ds} = 0.1$, 0.3 and 0.5V and a gate length of $L_g = 20$nm. The aforesaid ambipolar behaviour is more clearly marked in the case of the NiSi device as long as its SB for electrons is more similar to that of a mid–gap silicide. Obviously, the specific choice of the silicide influences the device performance. In general, the employment of mid–gap silicides provides poorer saturation

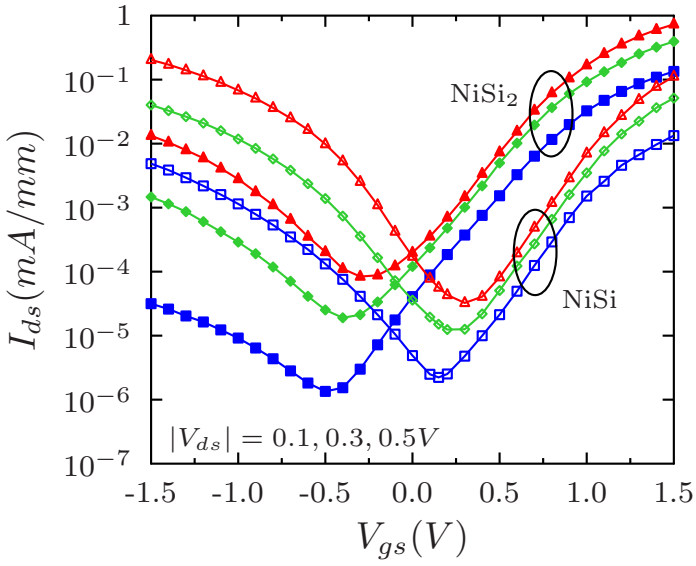

Figure 3.7: Experimental data corresponding to the transfer characteristics of NiSi and epi–$NiSi_2$ S/D SB–MOSFETs with $L_g = 20$nm. $V_{ds} = 0.1$, 0.3 and 0.5V.

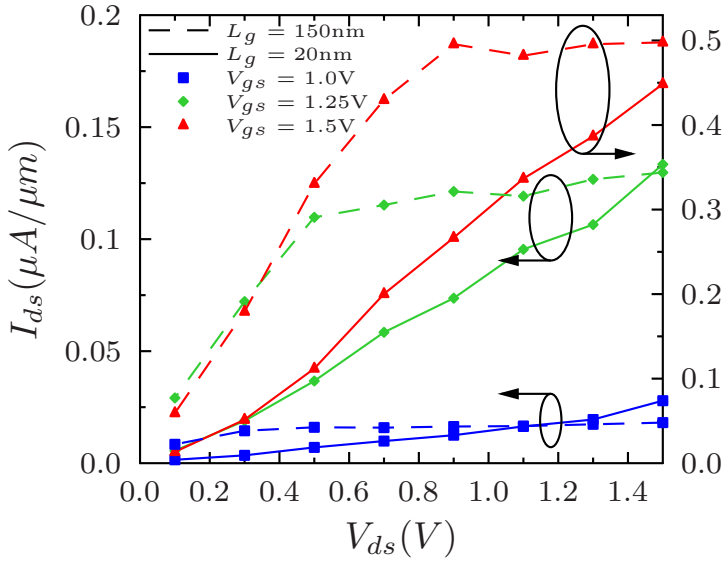

Figure 3.8: Experimental output characteristics of NiSi S/D SB–MOSFETs with $L_g = 20$ and 150nm. $V_{gs} = 1.0$, 1.25 and 1.5V.

drive current [5, 51] and high subthreshold leakage.

Output characteristics for the NiSi and epitaxial $NiSi_2$ devices for gate lengths of 150nm and 20nm are shown in Figs. 3.8 and 3.9, respectively. The obtained current levels suffer from the increased parasitic resistance caused by the small underlap between the silicide and the TiN gate. Short channel effects —especially the DIBL— make the $I_{ds} - V_{ds}$ curves increase without saturation for the case with reduced gate length ($L_g = 20$nm). Note that a quite remarkable issue is that the $L_g = 20$nm devices show smaller currents in the linear region compared to those with $L_g = 150$nm. This current reduction for a gate length of 20nm can also be noticed in Fig. 3.10, where experimental $I_{ds} - V_{gs}$ curves for the NiSi device with $L_g = 20$ and 50nm are shown. This feature will be later elucidated and reproduced by our simulations.

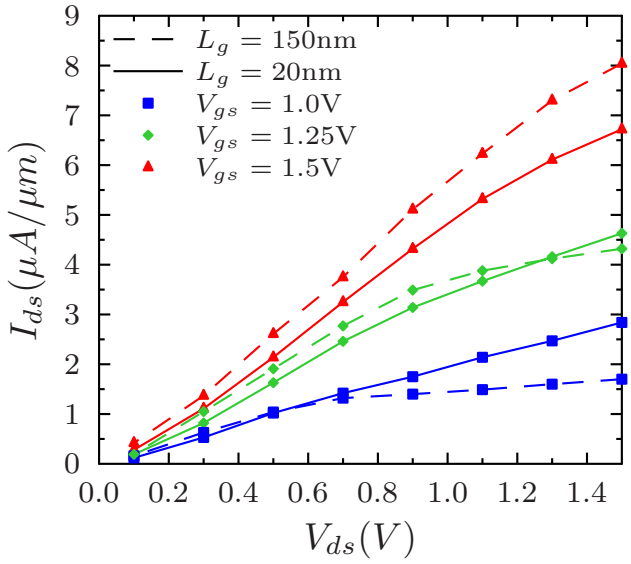

Figure 3.9: Experimental output characteristics of epi–$NiSi_2$ S/D SB–MOSFETs with $L_g = 20$ and 150nm. $V_{gs} = 1.0$, 1.25 and 1.5V.

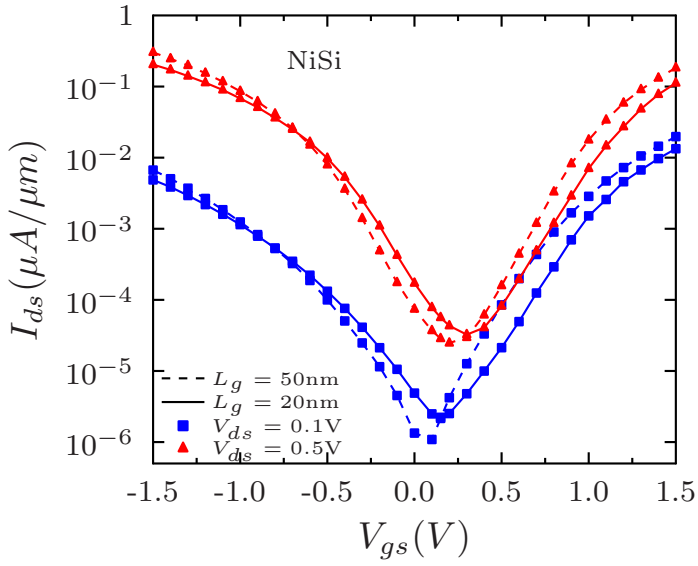

Figure 3.10: Experimental transfer characteristics of NiSi S/D SB–MOSFETs with $L_g = 20$ and 50nm. $V_{ds} = 0.1$ and 0.5V.

3.5 Simulation scheme and methodology

3.5.1 Existing framework and limitations

Both IFL and DL have an electric field dependence; however, the electric field at the interface depends on the shape of the SB which is, in turn, controlled by the gate and influenced by short–channel effects such as the overlap between source and drain potential profiles. Thus, a self–consistent solution to the potential at and near the contacts is required. Furthermore, as the barrier modulation affects its thickness, the tunneling probability used when field emission begins to appear would be also modified and that would force this consistency to account for it too. Unfortunately, as explained in [30] such a self-consistent BL calculation including field emission is not currently implemented in commercial TCAD software (for example, [9, 104]).

Recent simulations [31] using *Sentaurus device* implement BL mechanisms using the model presented in [41] and slightly modified in [30]. In our case, we use ATLAS version 5.18.3.R with field emission current described by the universal Schottky tunneling (UST) approach presented in [105, 106] where the tunneling probability is calculated using the Wenzel-Kramers-Brillouin (WKB) approximation, which —as will be explained in Sec. 4.3.2 when modeling the tunneling probabilities in TFETs— assumes a triangular potential profile. The use of this approximation has been commonly accepted [105, 107, 108]. In our case, BL estimations applied to both thermionic and field emission are externally calculated to the ATLAS simulator and described in Sec. 3.5.2 (a detailed discussion of a similar approach involving external calculations can be found in [30]).

It is also worth noting that for high gate biases the “wide barrier” assumption, inherent in the WKB approximation, begins to be less precise [109]. As a possible solution, it has been proposed [85] that the accuracy of the WKB model could be extended if BL were excluded, or if included at the expense of excluding thermal current. However, these solutions are hardly justifiable from a physical point of view. Instead of that, an alternative Airy-transfer-matrix (ATM) formalism [85, 110] was

shown to be more precise. This formalism has a higher complexity and is not currently implemented in commercial simulators.

3.5.2 Proposed simulation approach

We propose a tradeoff that allows to keep the widely used WKB model by correcting its deviations through an adequate fitting of the carrier masses that we use for the estimation of the tunneling probabilities (the so–called tunneling effective masses), along with a vertical discretization of the channel (it will be divided in several horizontal layers parallel to the semiconductor/insulator interface). This vertical discretization makes it possible to account for the SBH dependence on the depth inside the channel, which may be of considerable importance in relative terms especially for small barrier values.

Our simulation approach is based on the idea of estimating BL externally to the Silvaco ATLAS simulator through an iterative procedure that recalculates for each iteration the SBH using the electric field value extracted from the previous iteration. This BL calculation is performed for every horizontal layer in which the channel is vertically divided. The electric field needs to be extracted very close to the contacts (ideally at the metal–semiconductor interface). This process is repeated until variations from one iteration to the next (in terms of $\Delta\phi$) are not significant. Fig. 3.11 shows a diagram illustrating this iterative calculation for a bias configuration allowing, for example, hole tunneling across the SB.

The role of the mesh used in the simulations is crucial. Next to the contacts, it has to be extremely thin because otherwise small fluctuations in the slope of the potential may occur from one iteration to the next, making the resultant value of the electric field not to converge in the iterative procedure indicated above. A tradeoff has to be assumed between the refining of the mesh, that determines the minimum distance of the contacts at which the electric field can be extracted assuring convergence (in our case, typically of order 1nm) and the computational cost in time that it implies. An example of the convergence process is

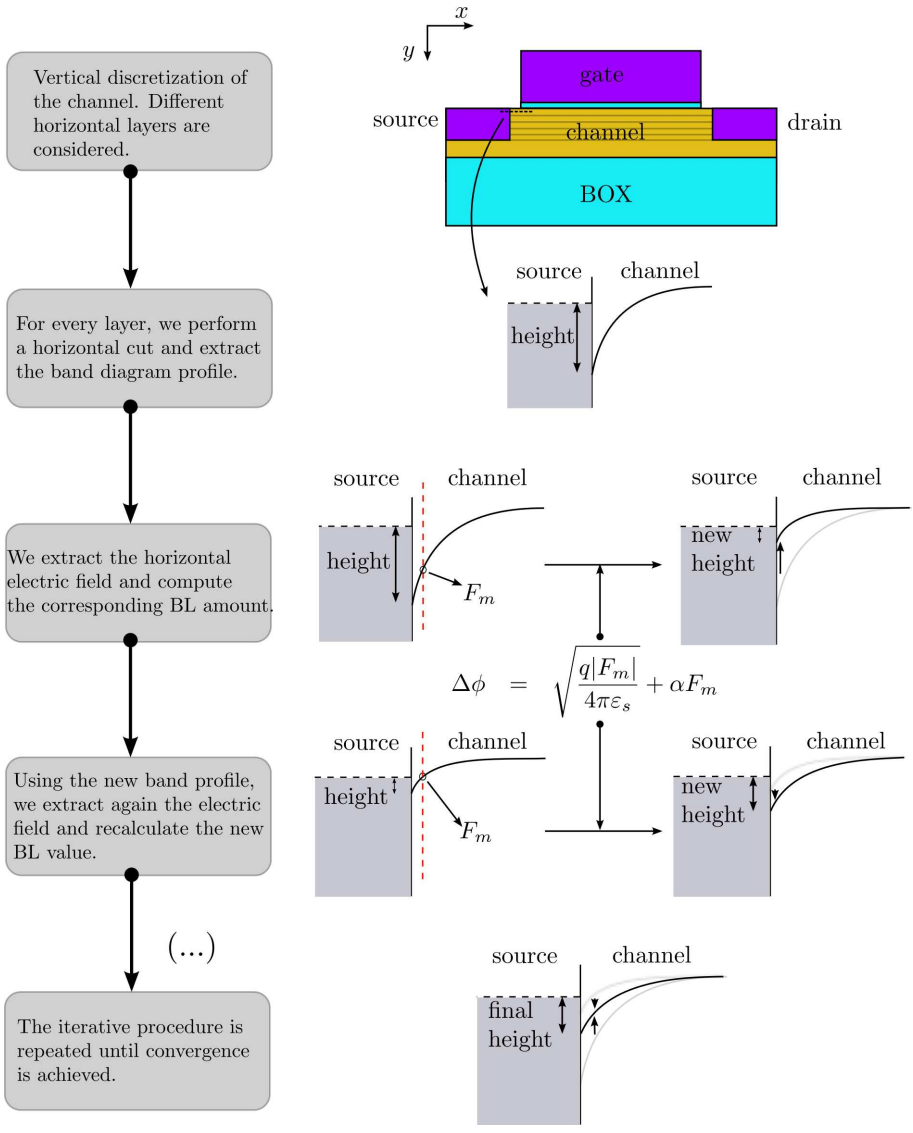

Figure 3.11: Diagramatic illustration of the iterative external procedure to estimate BL. The depicted diagrams correspond to the situation where BL is calculated for holes.

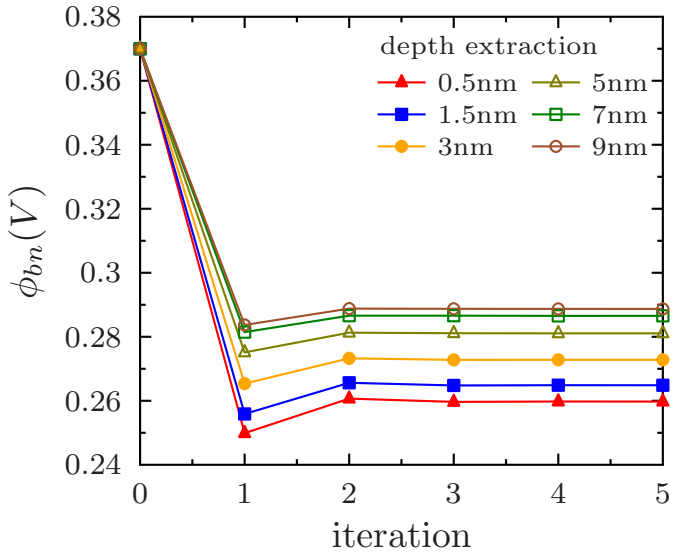

Figure 3.12: Convergence example of the external iterative procedure to estimate the resulting effective SBH for electrons at every horizontal layer of the discretized channel. The displayed values correspond to the $NiSi_2$ S/D SB–MOSFET at $V_{ds} = 0.5$V and $V_{gs} = 1.5$V.

displayed in Fig. 3.12, where we show the simulated SBH for electrons at the source–to–channel junction for the different discretization layers corresponding to the $NiSi_2$ S/D SB–MOSFET. The gate and drain voltages were chosen to be $V_{gs} = 1.5$V and $V_{ds} = 0.5$V, respectively. If the process is repeated for the drain–to–channel junction, we can obtain for electron injection a 3D profile of the effective conduction band at that bias configuration, as shown in Fig. 3.13.

Preliminary simulations including both IFL and DL, seemed to suggest that the contribution of the latter to the total BL was considerably lower than that due to the first. An explanation to this may be found in the abruptness of the metal–semiconductor junctions that eliminates the contribution of interfacial states to the DL effect. For this reason, in what follows DL will be assumed to be included through the fitting of the tunneling effective masses (thus eliminating the two fitting parameters of Eq. 3.8), without losing too much accuracy.

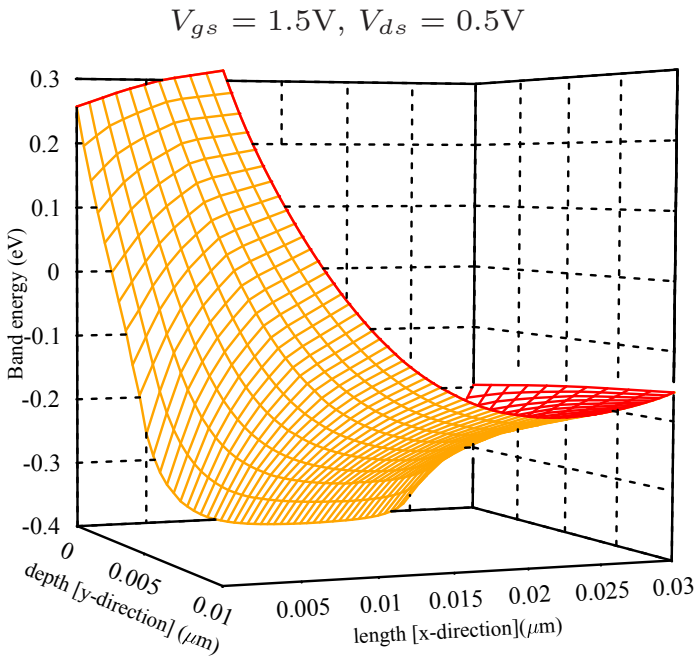

Figure 3.13: 3D effective profile of the conduction band for the $NiSi_2$ device once BL has been taken into account. $V_{ds} = 0.5\text{V}$ and $V_{gs} = 1.5\text{V}$.

3.6 Simulation results and experimental data

3.6.1 Transfer characteristics for $L_g = 20\text{nm}$

In Fig. 3.14(a) we show the transfer characteristics of the epitaxial $NiSi_2$ S/D SB–MOSFETs with $L_g = 20\text{nm}$. We see that the n–channel current is higher than the p–channel one which agrees with the barrier heights for electrons reported in Sec. 3.3. Solid lines stand for the simulated currents with tunneling masses of $m_h = 0.46$ and $m_e = 0.8$ which are higher than those reported for example in [30, 31, 85]. Fig. 3.14(b) illustrates how our BL procedure modifies the source/drain barriers for electrons and holes. Note that as the BL treatment is external to the simulator, this estimation is not fully self–consistent, leading to a consideration of the barriers as abrupt changes in potential which was already mentioned

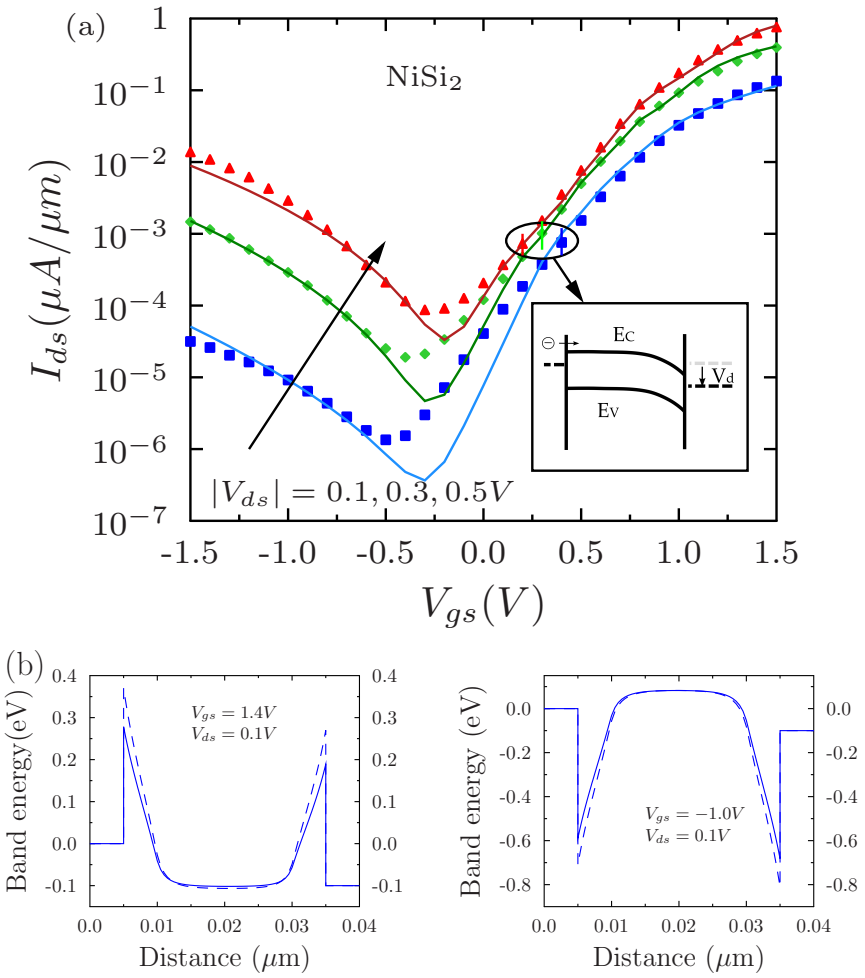

Figure 3.14: (a) Drain currents for $V_{ds} = 0.1, 0.3$ and 0.5V in the device with epitaxial $NiSi_2$. Solid lines represent simulated characteristics including BL, symbols stand for experimental data. (b) Profiles of the source/drain Schottky barriers at 0.5nm from the channel interface corresponding to electrons (left, $V_{gs} = 1.4$V, $V_{ds} = 0.1$V) and holes (right, $V_{gs} = -1.0$V, $V_{ds} = 0.1$V) without BL (dashed lines) and with our BL estimation (solid lines).

in [30]. Therefore the resulting potential profiles after BL slightly divert from the theoretical, more rounded, ones. This issue, along with the current overestimation arising from the WKB approximation, may help to understand the higher values for effective tunneling masses used in

our simulations.

In the right branch of Fig. 3.14(a), that of electrons, we have marked the gate voltage values corresponding to flat band condition in the source edge and at the bottom of the channel ($y = 10nm$ going down from the insulator, where flat bands first occur when increasing gate bias) for different values of V_{ds}. These points represent the limit where tunneling begin to appear when we increase V_{gs}. Note how the simulations (including BL) fit very well the region dominated by tunnel current, but show certain deviation from experimental results in the case of pure thermionic emission for electrons. If we now focus on the left side, hole current has tunneling and thermionic contribution in the entire range of V_{gs} and thus we do not observe significant deviations in the simulated currents.

In order to quantify the relative importance of both thermionic and

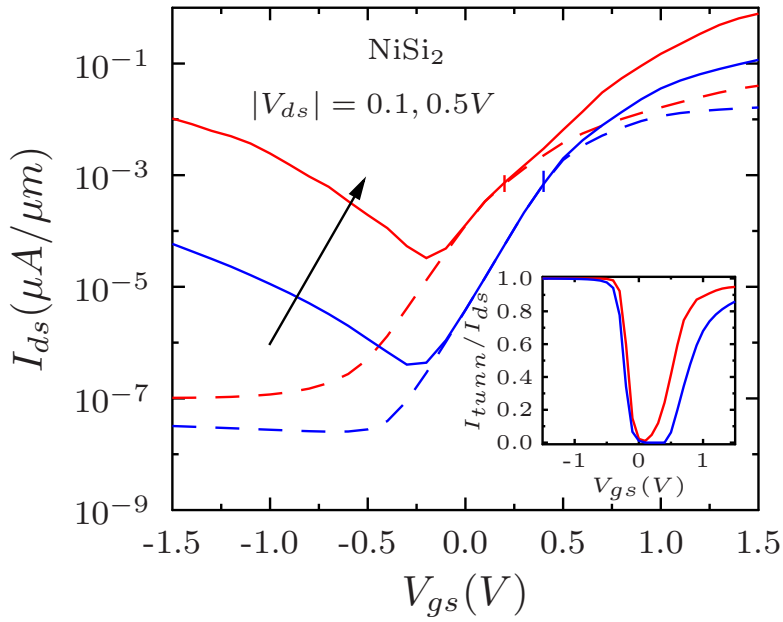

Figure 3.15: Simulated transfer characteristics with BL of the epitaxial $NiSi_2$ SB-MOSFET for $V_{ds} = 0.1, 0.5$V including only thermionic emission (dashed lines) and both thermionic and field emission (solid lines). Inset shows relative importance of tunneling current over total current.

field emission contributions when BL is included, and to show explicitly where tunneling begins to appear in the n–branch, we present in Fig. 3.15 a comparison between previous characteristics corresponding to epitaxial $NiSi_2$ for $V_{ds} = 0.1$ and 0.5V including thermionic, thermionic field emission and field emission (solid lines displayed in Fig. 3.14(a)), and those including only thermionic emission (dashed lines). The inset shows the relative importance of tunneling current over total current. In light of these curves, importance of tunneling becomes apparent for increasing values of $|V_{gs}|$, specially for the p–branch. Differences of up to five orders of magnitude can be observed for $V_{gs} = -1.5$V and $V_{ds} = 0.5$V.

Let us now focus on the role that discretization plays. If the channel were not vertically discretized, the same lowered value for the SB would be present in the whole metal–semiconductor surfaces. In Fig. 3.16

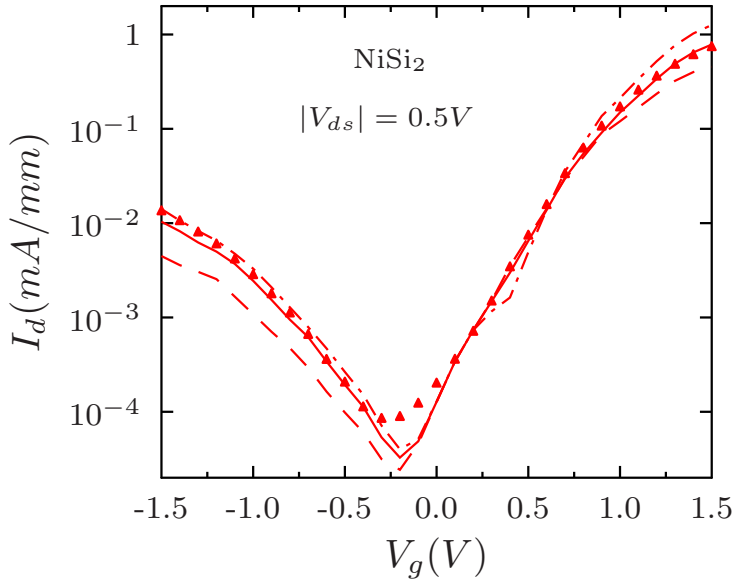

Figure 3.16: Comparison between simulated transfer characteristics of the epitaxial $NiSi_2$ SB-MOSFET for $V_d = 0.5$V with discretization (solid line) and without it (dashed and dashed-dotted lines). Triangles stand for experimental data.

we display simulations for $NiSi_2$ SB–MOSFET at $V_{ds} = 0.5$V without discretization and compare them to the discretized curve (solid line) and experimental data (triangle markers). In the n–channel with no discretization, if we use the value of the electric field in the uppermost part of the junctions to calculate BL, the resultant curve (dashed-dotted line) presents an unphysical kink around $V_{gs} = 0.4$V. The appearance of this kink relies on the fact that, at the source, flat band condition first arises ($V_{gs} \approx 0.3$V) at the bottom of the channel (see how in Fig. 3.17 for the $NiSi_2$ SB–MOSFET, the deeper we go down the contact, the sooner the barrier height for electrons coincides with the zero-bias height for increasing gate voltages). However, in this case BL does not begin until flat band is reached at the top ($V_{gs} \approx 0.45$V) because that is the point where we are extracting the electric field. Therefore, as no BL is being considered for $V_{gs} = 0.3 - 0.45$V, current is clearly underestimated in this range. Contrary to this underestimation, as gate voltage gets higher, the electric field at the top of the junction increases faster (see in Fig. 3.17 that for high gate voltages, BL is greater at the top), thus producing an overestimation of the current. On the other hand, if we choose the value of the electric field at the bottom of the junctions, no kink appears (because BL is incorporated since it first occurs at the bottom) but underestimated current results for increasing gate voltages (dashed line in Fig. 3.16). In the p–channel, as tunneling is present in its whole branch, no unphysical kinks are observed. Note that in the case of holes, the discretized curve lies very close to the undiscretized one using the value of the electric field at the top of the contacts.

Simulations and experimental data for the NiSi SB–MOSFET with $L_g = 20$nm can be seen in Fig. 3.18. Tunneling effective masses were taken as $m_h = 0.8$ and $m_e = 0.4$. For $V_{ds} = 0.1$V there is a small region in the p–branch of purely thermionic emission that gets narrower for increasing drain voltage. When $V_{ds} = 1.5$V this region has disappeared and tunneling is present in the whole range of V_{gs}. Unlike what happened in the epitaxial $NiSi_2$ device, there is no significant deviation between simulated curves and experimental results in those regions where tunnel current is absent.

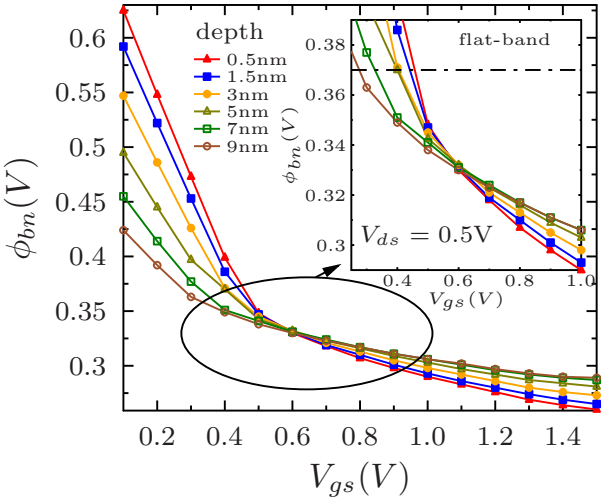

Figure 3.17: Simulated electron barrier height dependence on V_{gs} at $V_{ds} = 0.5$V for several depths going down the metal-semiconductor junction at the source in the $NiSi_2$ SB–MOSFET. Recall that for this device zero–bias BH was $\phi_{bn} = 0.37$eV. Note how flat band condition is first reached at the bottom of the contact.

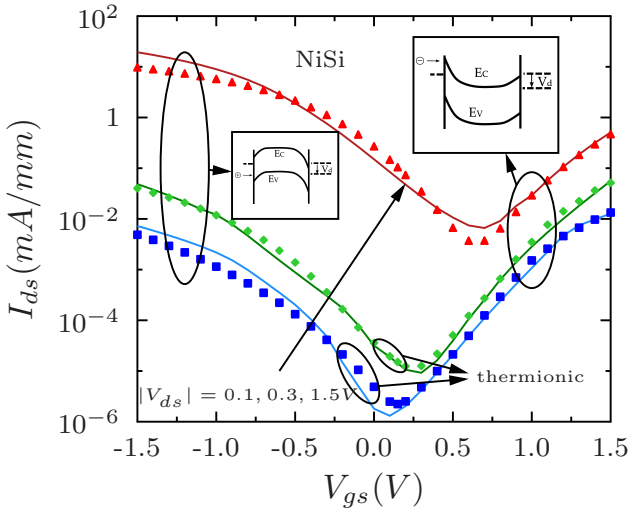

Figure 3.18: Drain currents corresponding to the NiSi SB–MOSFET for $V_d = 0.1, 0.3$ and $1.5V$. Symbols stand for experimental data. Solid lines represent simulated results.

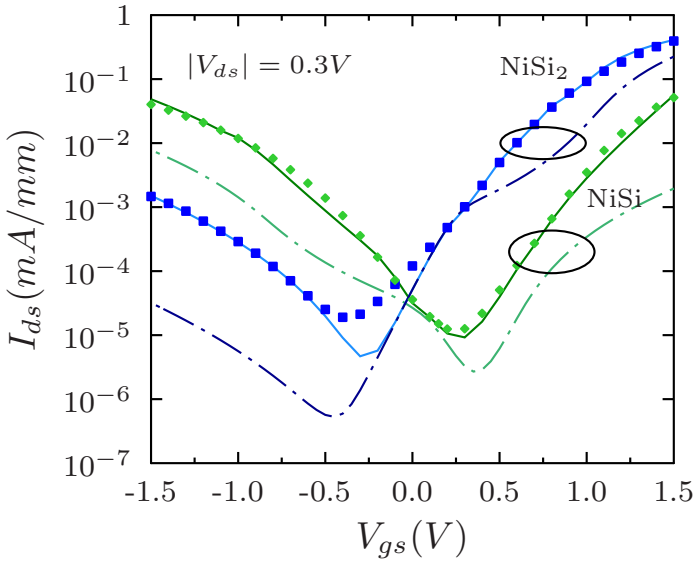

Figure 3.19: Simulated characteristics including internal (dashed–dotted lines) Silvaco ATLAS BL calculation (only applied to thermionic emission) and our external BL estimation (solid lines) for epitaxial $NiSi_2$ and NiSi SB–MOSFET. Both curves clearly differ when field emission contributes to total current.

It is worth noting that for both the NiSi and epitaxial $NiSi_2$ SB–MOSFETs the simulated characteristics including tunnel current using the WKB approximation with BL and the fitted tunneling effective masses remain very close to experimental data even for high $|V_{gs}|$ contrary to what was shown in [85] where simulations using the WKB model and including BL differed up to one order of magnitude from experimental results. This strengthens the usefulness of our proposed simulation mechanism that allows considerable accuracy with only two fitting parameters.

In Fig. 3.19 we show the comparison between the simulated characteristics (including thermionic and field emission) obtained using internal ATLAS estimation for BL (dashed–dotted lines) which only applies BL to thermionic emission, and those including our external BL calculation procedure (solid lines) that account for it for both contributions. Misleading results arising from the first curves clearly show a systematic

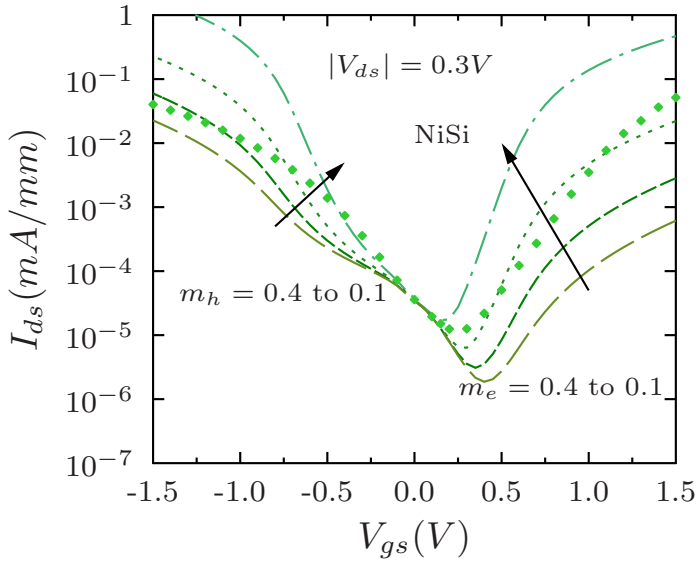

Figure 3.20: Transfer characteristics for different effective tunneling masses using internal ATLAS estimation of BL (only implemented for thermionic emission) for NiSi SB–MOSFET. Diamonds stand for experimental data.

underestimation of total current in the range where tunneling is present. If one attempts to modify the effective tunneling masses trying compensate that lack of current when using internal ATLAS BL, it can be seen in Fig. 3.20 for NiSi that there are no values of m_h and m_e neither producing a good fit to the experimental data nor reproducing the shape of experimental curves.

3.6.2 Short channel effects (SCE)

Another interesting point is the apparent reduction of tunnel current obtained for increasing values of $|V_{gs}|$ when the gate length is reduced from 50nm to 20nm, contrary to what one would expect considering the scaling behaviour. This experimental discovery [2] is also revealed in our simulations and is consistent, for example, with the predictions made in [45], where it was suggested that for single gate SB–MOSFETs the downsizing of the devices, starting at $L_g \approx 20$nm, would make the influence of Schottky contacts become significant across the entire length

of the channel, making the conduction band much stiffer to bending by the gate voltage. In Fig. 3.21 we show this current reduction for $L_g = 20$nm in the NiSi SB–MOSFET as a result of the overlapping between source and drain barriers and the subsequent increase of the total potential inside the channel. The inset in Fig. 3.21 schematically illustrates this effect in the range of positive V_{gs}. This observed difference between characteristics is higher for NiSi in the case of electrons because in the n–channel branch tunneling occurs for both source and drain in the whole range of voltages. Fig. 3.22 shows the simulated conduction band profiles for $L_g = 20$nm (solid line) and $L_g = 50$nm (dashed line) at $V_{gs} = 0.7$V, $V_{ds} = 0.3$V. In the p–channel branch, however, tunneling begins to appear in both interfaces at $V_{gs} \approx -0.5$V while for $V_{gs} > -0.5$V only thermionic emission occurs at the source. This causes this overlapping to be less dramatic for holes as displayed in Fig. 3.23 for $V_{gs} = -1.0$V,

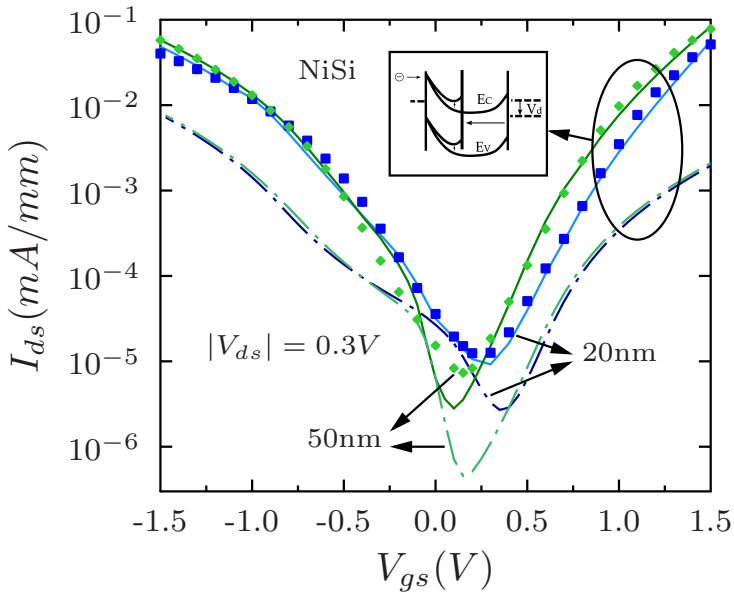

Figure 3.21: For increasing values of $|V_{gs}|$ we get higher current for $L_g = 50$nm (diamonds) than for $L_g = 20nm$ (squares) in the NiSi SB–MOSFET. This behaviour is reproduced in our simulations (solid lines). Simulated characteristics using internal ATLAS estimation of BL (dashed–dotted lines) do not follow this pattern conclusively.

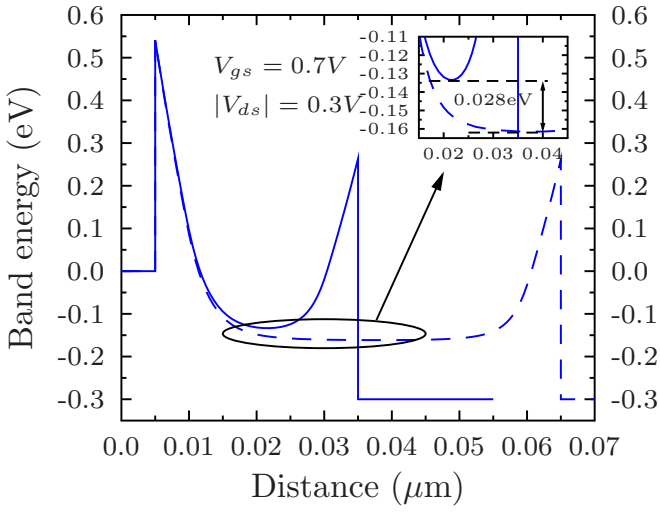

Figure 3.22: Conduction band profiles for the NiSi SB–MOSFET with $L_g =$ 50nm (dashed line) and $L_g = 20nm$ (solid line) after BL. A certain overlap of source and drain barriers produces an increase of the potential inside the channel as depicted in the inset.

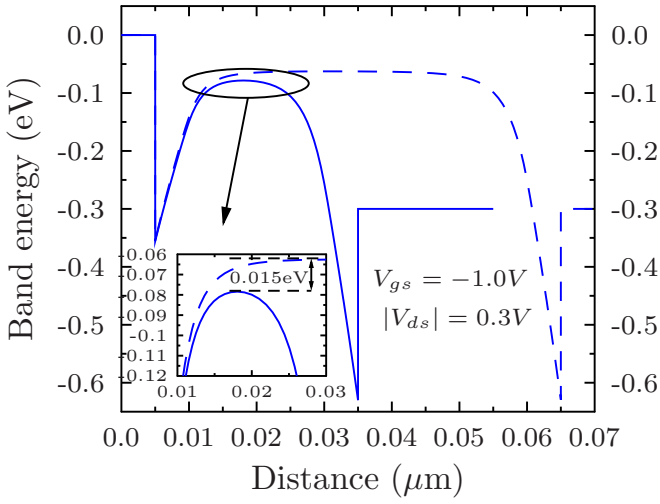

Figure 3.23: Valence band profiles for the NiSi SB–MOSFET with $L_g = 50$nm (dashed line) and $L_g = 20nm$ (solid line) after BL. As SB for holes in NiSi is lower than for electrons, the potential inside the channel is less affected by the overlap as shown in the inset.

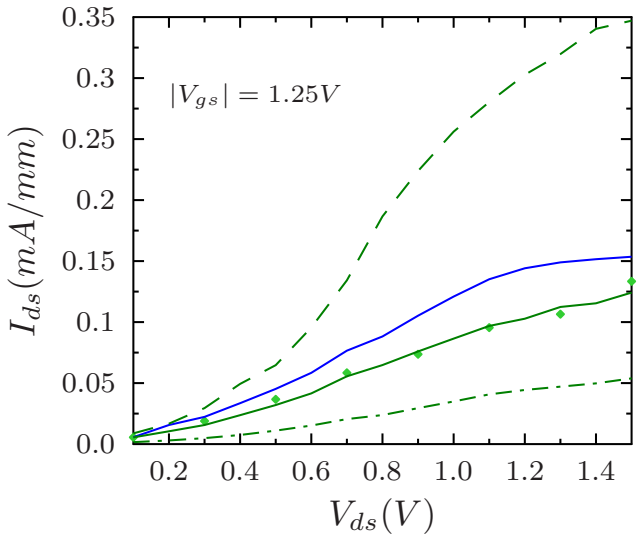

Figure 3.24: Simulated output characteristics of NiSi SB-MOSFET. Dashed and dashed-dotted lines correspond to $L_g = 20nm$ with underlaps of 4nm and 6nm respectively. Solid lines correspond to 5nm underlap, top most stands for $L_g = 50$nm. Diamonds represent experimental data for $L_g = 20$nm and 5nm underlap.

thus moderating to some extent the aforementioned overlapping.

Another effect which also appears in these devices is the *drain induced barrier lowering* (DIBL). It causes I_{ds} to increase with V_{ds} without appreciable saturation. This behaviour is indeed observed in Fig. 3.24 for the three curves corresponding to $L_g = 20$nm. Observe that, however, the curve with $L_g = 50$nm seems to show certain saturation for $V_{gs} > 1.2$V. Simulations changing the underlap between the gate and the silicide for $L_g = 20$nm have been performed showing, for increasing values of it, an apparent degradation of performance manifested in a reduction of drain current, thus indicating the impact that this underlap has on parasitic resistance.

3.6.3 Underlap variation

Although no experimental data are available concerning variations in the underlap between the gate contact and the S/D regions, we may

use the tunneling effective mass calibration so far obtained for the case corresponding to 5nm. In Fig. 3.25, we depict the transfer characteristics for the epitaxial $NiSi_2$ S/D SB–MOSFET at $V_{ds} = 0.5$V, where the underlap has been modified from 6nm to 1nm using a symmetrical configuration, i.e. assuming the same underlap for the source and the drain.

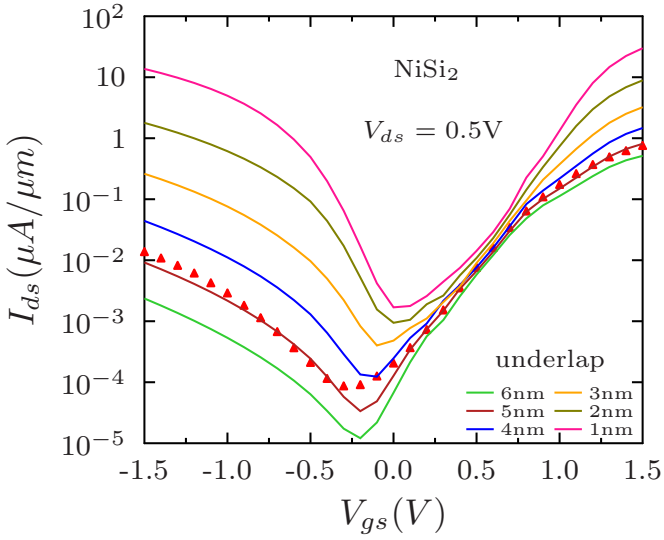

Figure 3.25: Simulated transfer characteristics for different values of the underlap of the $NiSi_2$ SB-MOSFET. Triangle markers stand for experimental data in the case of a 5nm underlap. $L_g = 20$nm and $V_{ds} = 0.5$V.

Observe how the reduction of the underlap produces for the p–branch of the curve a current increase of nearly four orders of magnitude at $V_{gs} = -1.5$V. However, the corresponding current level for the n–brach at $V_{gs} = 1.5$V is only raised by less than a factor of 100. The explanation of such a behaviour lies in the fact of the different SBH for electrons and holes when using $NiSi_2$. If we reduce the underlap, the influence of the gate over the SB potential profile near de junction increases, hence producing in the case of holes stronger electric fields (because of their bigger zero–bias SBH). And this, in turn, causes a bigger BL for these carriers with the subsequent current level rise, greater than that obtained for electrons.

Chapter 4

Tunneling Field Effect Transistors (TFETs)

Tunneling Field–Effect Transistors are one of the most promising devices to replace conventional MOSFETs. Their low off–current and steeper subthreshold slope overcoming the 60mV/dec limit of MOS transistors make them enormously attractive for low–power applications. One of the main problems arising in MOSFETs is that when they are scaled down, so has to do their power supply voltage in order to reduce power density. The subthreshold swing limit of 60mV/dec present in conventional MOSFETs imposes a severe roadblock to reduce the supply voltage plateau of 1V and maintain high ON–state currents along with low OFF–state leakages.

TFETs, on the other hand, are based in the so called band–to–band tunneling (BTBT) mechanism which makes the carrier injection into the channel essentially dependent on the quantum process of tunneling across an energy barrier. This fact allows extremely low subthreshold swings when the device turns on due to the quasi–exponential dependence of the current on the barrier width. Likewise, when the transistor is off, the tunneling barrier keeps the leakage current extremely low. In MOSFET scaling, tunneling phenomena from heavily doped junctions resulted in parasitic leakage currents. However, as this process is precisely the working principle of TFETs, it is no longer an unwanted

parasitic effect. Furthermore, since tunneling only takes place in a very small region of these devices, this may allow significant gate scaling up to the distance of BTBT which in Silicon represents less than 10nm. As source–to–drain direct tunneling is negligible for channel lengths greater than that value [111], TFETs could in principle be scaled to very small dimensions without relevant leakage current degradation.

4.1 Origins and history

Since the discovery of BTBT in 1957 by Esaki [112] when studying very narrow germanium p–n junctions, this phenomenon based on the tunneling injection of carriers from occupied states in the valence(conduction) band to empty states in the conduction(valence) band has been demonstrated in many devices. As indicated above, such a mechanism has been shown for example in MOSFETs (both lateral and vertical). The first gated p–i–n structure was proposed in 1978 at Brown University [113] suggesting it for spectroscopy. Transistors based on it (like B^2T–MOSFETs [114], or others replacing the i–region under the gate by a p^-–region [115]) were investigated, showing the lack of V_t rolloff and temperature dependence of the device characteristics when scaling. However, the first gated p–i–n diodes operating as Surface Tunnel Transistors were proposed on III–V compounds [116, 117]. Similar tunneling transistor operation was developed in silicon at Cambridge [118], and later at Toshiba [119].

The interest of these first results was limited until the experimental results presented in 2000 by W. Hansch and I. Eisele in Munich [120], which led to a series of works over vertical p–i–n diodes [121–123]. In 2004, a lateral gated p–n junction diode (without intrinsic region) on silicon–on–insulator was fabricated [124] at Brown University. In this last case, the lack of intrinsic region reduced gate capacitance but did not significantly improved ON–current —which was still very low— and also produced an increase in leakage current. Also in 2004, Appenzeller et al. [125] reported for the first time a subthreshold swing under 60mV/dec in carbon nanotube (CNT) FETs. A back gate and a

top gate were employed to achieve the neccesary band configuration to trigger BTBT. In 2005, the same authors [126] compared several CNT transistors concluding that the single gate configuration presented the best performance.

Despite the obtained results for CNT FETs, the research on silicon based FETs offered in principle a more immediate possibility to industrial applications due to the greater development of this technology [127–129]. In 2009 at Forschungszentrum Jülich, new results showing experimental data for p–type TFETs varying source and drain doping, gate dielectric thickness and channel length were published [130]. The same year at Zurich, Moselund et al. [131] fabricated TFETs based on wrapped silicon nanowires around gate using two different dielectrics: SiO_2 and HfO_2. Works combining and comparing different substrates [132], or focusing on germanium were also carried out like, for example, TFETs on thin GeOI [133] which suffered from very high leakage currents.

With the interest renewed in these last years, many groups have studied different aspects of TFET performance such as channel dimensionality [134], power consumption [135], phonon scattering [136], temperature dependence [137], gate overlap [138], threshold voltages [139], high-κ gate dielectric [111, 140, 141], performance comparison to CMOS [142], heterostructure TFETs [143–146], strain [147], quantum confinement [148, 149] and general modeling [48, 150–153].

4.2 Structure and operation

The device structure of a TFET essentially differs from that of the MOSFET in the nature of the dopants used in the source and the drain. While MOSFETs have the same type of dopants, in TFETs, source and drain are of opposite types. The basic constituent of TFETs is thus a gated p–i–n diode, or less frequently —as previously mentioned— a gated p–n diode [124, 154]. The name of the terminals is chosen to resemble the MOSFET biasing. To switch the device ON, the diode has to be reverse biased and a voltage applied to the gate. Therefore, a n–type TFET

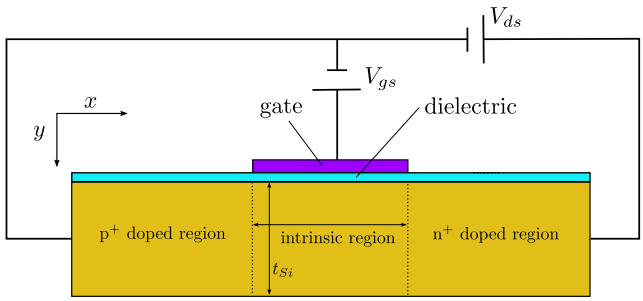

Figure 4.1: Single gate n–type TFET. The p$^+$–region acts as the source and the n$^+$–region as the drain.

would require positive voltage in the gate and also in the n–doped region, which would play the role of the drain if one recalls the analogy with the NMOS. The other p$^+$ region would act as the source and the intrinsic region as the channel.

In Fig. 4.1 we present the structure of a lateral single gate n–type silicon TFET where the dielectric covers the source, the drain and the channel. In a p–type TFET, the dopings would be the opposite: the source would be n$^+$ and the drain, p$^+$.

4.2.1 Tunnel diodes

Prior to study the operation regimes of these transistors, and in order to better understand them, it may result very useful to analyze the qualitative behaviour of the p–n tunnel diodes, in which p–i and n–i junctions of the TFET are based. Tunnel diodes consist of a p–n junction in which both p and n sides are degenerate (i.e. very heavily doped). To illustrate this, let us consider the tunnel diode configurations depicted in Fig. 4.2 along with their corresponding points in the $I-V$ curve, and discuss the different processes taking place in them at absolute zero temperature. Each of the different figures correspond to:

(a) When the tunnel diode is reverse biased (p–side negative bias with respect to n–side), the BTBT current increases monotonically and indefinitely.

(b) At thermal equilibrium, no voltage is applied and the Fermi levels

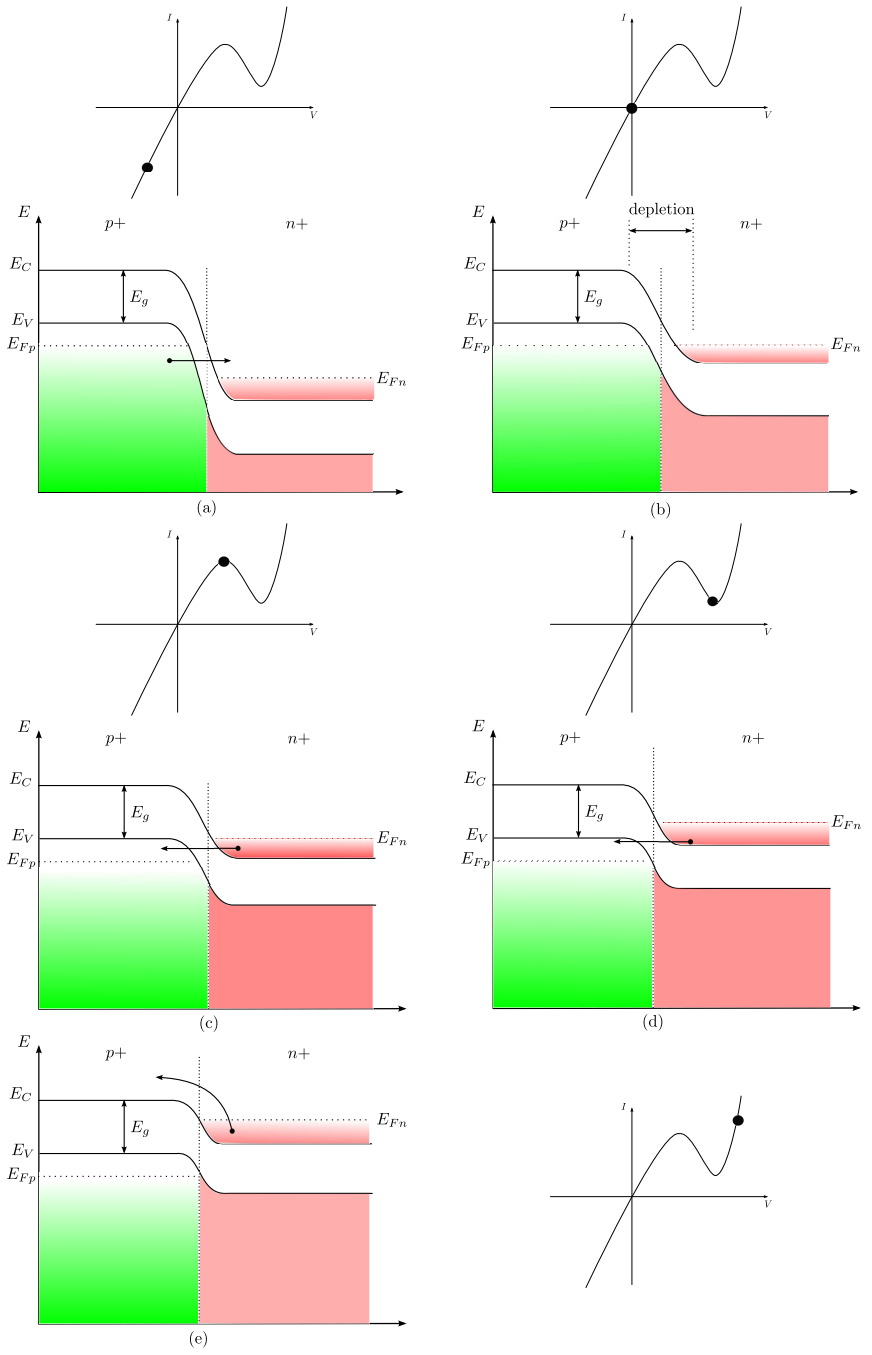

Figure 4.2: Energy band diagrams and $I-V$ characteristics of a tunnel diode at: (a) reverse bias with increasing tunneling current; (b) thermal equilibrium, zero bias; (c) forward bias V such that peak current is obtained; (d) forward bias approaching valley current; and (e) forward bias with diffusion current and no tunneling.

are aligned. That means that above the Fermi level there are almost no filled states on either side of the junction, and below it there are almost no empty states available on either side of the junction. Hence, net tunneling current at zero bias is zero.

(c) In the forward direction (positive voltage at the p–side with respect to n–side), the current first increases to a maximum because electrons can tunnel from the conduction band to the valence band. Tunneling is possible as there is a common band of energies with filled states on the n–side and unoccupied states on the p–side.

(d) If forward voltage is further increased, this range of common energies decreases and so does the current until the bands are uncrossed and there are no available states aligned with filled states.

(e) Once tunneling current becomes zero, normal diffusion current begins to dominate and current increases again exponentially.

However, in Fig. 4.2 (e) there also exists another type of current contribution called *excess current*. The excess current arises from a BTBT process that takes place indirectly through energy states within the forbidden gap. Several possible routes followed by carriers can be seen in Fig. 4.3 [155]. As an example, an electron could drop down from A to an empty level at B, decreasing its energy, from which it might tunnel to the final state D in the valence band. Alternative trajectories would be ACD, or even a staircase route formed by several tunneling transitions followed by vertical losses of energy. This last one is by far less probable and requires a sufficiently high concentration of intermediate states.

4.2.2 Operating regimes of the TFET

In Fig. 4.4, we show the OFF and ON states of the TFET. There are two configurations that correspond to the OFF state. The first is when the band structure is in equilibrium and no bias applied, Fig. 4.4 (a), which corresponds to a situation as that shown in Fig. 4.2 (b). In this initial state, the built–in potentials of the p–i and n–i junctions determine

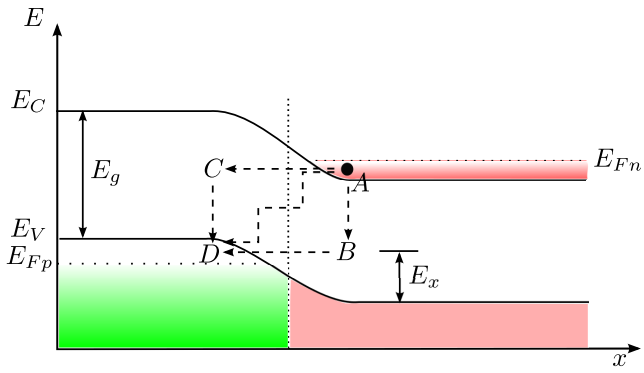

Figure 4.3: Example of different BTBT trajectories through intermediate states in the forbidden gap. E_x is the energy through which the electron must tunnel in B [3].

the characteristic staircase–like profile that can be observed. When we only apply a positive voltage to the drain, $V_{ds} > 0$, the current flow is not allowed through the device in absence of gate bias because electron and hole currents are blocked by the built–in barriers (Fig. 4.4 (b) with $V_{gs} = 0$). In this situation, only reverse biased p–i–n diode leakage current flows between the source and the drain. This leakage current is extremely low (may result of order fA/μm).

When we apply positive voltage to the gate, $V_{gs} > 0$, the conduction band inside the channel is pushed down until it is aligned with the top of the valence band of the source. From that point onwards, BTBT begins to be possible and carriers (in this case, electrons) are injected from the source into the channel (Fig. 4.4 (b) with $V_{gs} > 0$). This operating mode is the n–channel ON–state with surce/channel junction resembling that of Fig. 4.2 (a). We clearly see how the gate controls the band bending inside the channel and, consequently, the BTBT mechanism. TFETs designed with symmetry between the p and n–doped regions, may show ambipolar behaviour provided that adequate voltages are applied to the terminals. In our case, if we change the sign of the voltage applied to the gate, $V_{gs} < 0$ —while keeping $V_{ds} > 0$—, the bands inside the channel move up and carriers can tunnel through the drain/channel junction as soon as the valence band inside the channel is lifted above the bottom

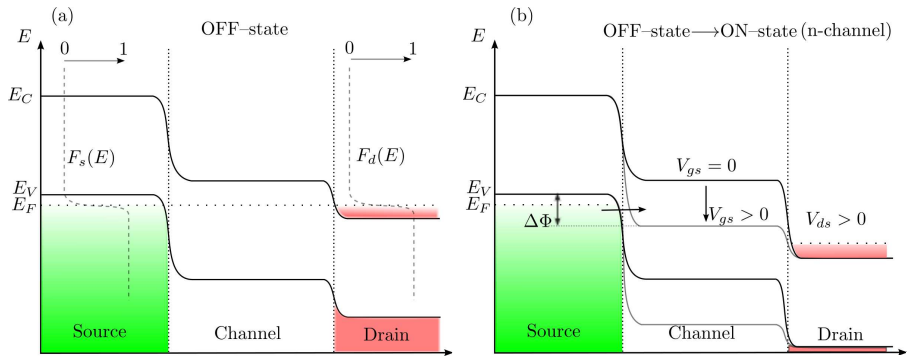

Figure 4.4: Energy band diagrams of the TFET taken horizontally along the channel close to the dielectric interface. (a) OFF–state showing the equilibrium configuration of energy bands when no bias is applied at the gate and the drain. (b) Combined ON and OFF–states. When $V_{gs} = 0$, BTBT can not take place and only p–i–n diode leakage current exists: OFF–state. If a big enough $V_{gs} > 0$ is applied, the conduction band in the channel is pushed down and BTBT may appear: ON–state. $\Delta\Phi$ represents the difference between the top of the valence band in the source and the bottom of the conduction band in the channel.

of the conduction band in the drain. This can be seen in Fig. 4.5. In this case, we again recall the situation of Fig. 4.2 (a), but now between the drain and the channel. Obviously, this ambipolar bahaviour would imply a reassignation of "source" and "drain" labels for the p–channel ON–state if one wants to keep the analogy with the MOSFET operation.

To complete the description of the operating modes of the n–type TFET, let us consider $V_{ds} < 0$ and sufficiently large. In that case, the p–i–n structure is forward biased which means that carriers can flow with $V_{gs} = 0$ and result in exponential diode characteristics. The application of gate bias may block either the electrons or the holes by means of a potential barrier but not both. This configuration is not appropriate for switching purposes.

Hence, while the drain bias switches the device characteristics from that of a forward–biased p–i–n diode to that of a TFET, the gate bias switches the TFET characteristics from a n–channel to a p–channel mode of operation, when the diode is reverse–biased.

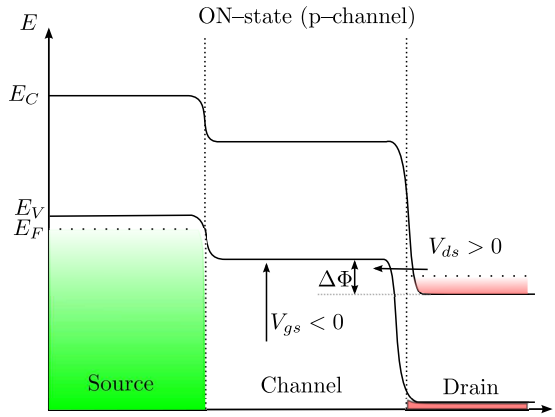

Figure 4.5: Band diagram configuration of the p–channel ON–state of the TFET. Carriers are injected from the drain into the channel once the top of the valence band inside the channel is raised over the bottom of the conduction band in the drain.

4.3 Working principles

4.3.1 Band–to–band tunneling current

TFETs are essentially based on tunneling rather than thermal emission. In this sense, they clearly differ from the normal operation pattern of conventional MOSFETs. In the subthreshold regime, we can use the following expressions [3, 156] to describe the BTBT current in a tunnel diode assuming direct tunneling where the momentum is conserved in direct bandgap

$$I_{C\to V} = C_1 \int_{-\infty}^{\infty} f_C(E) N_C(E) T(E) \left[1 - f_V(E)\right] N_V(E)\, dE \quad (4.1)$$

$$I_{V\to C} = C_1 \int_{-\infty}^{\infty} f_V(E) N_V(E) T(E) \left[1 - f_C(E)\right] N_C(E)\, dE, \quad (4.2)$$

C_1 is a constant, $N_{C(V)}(E)$ are the density of states in the conduction and valence bands respectively, $T(E)$ is the transmission probability across the energy tunneling barrier width —which is assumed to be equal for both directions—, and $f_{C(V)}(E)$ are the occupation probabilities of the bands described by the Fermi distribution functions

$$f_{C(V)}(E) = \frac{1}{1+\exp\left(\frac{E-E_{F_{C(V)}}}{kT}\right)}, \tag{4.3}$$

where $E_{F_{C(V)}}$ are the Fermi energy levels. If the junction is forward biased, I_{btbt} is given by

$$I_{btbt} = I_{C\to V} - I_{V\to C} = C_1 \int_{E_{Cn}}^{E_{Vp}} [f_C(E) - f_V(E)]\, T(E) N_C(E) N_V(E)\, dE \tag{4.4}$$

Observe that the limits of integration go from E_C of the n-side (E_{Cn}) to E_V of the p-side (E_{Vp}).

Several scenarios then arise depending on the value of the factors in Eq. 4.4. If $T(E)$ is significantly greater than zero for a certain range of energies, which means that BTBT is not negligible, carriers can flow through the device only at those energies of the previous range corresponding to a non–zero value of $[f_C(E) - f_V(E)]$. And, conversely, that difference being non–zero does not guarantee the existence of current if the bands profile does not fulfill the BTBT requirements. For the n–channel TFET, namely: (i) bottom of the conduction band inside the channel at or below the level of the top of the valence band at the source, (ii) tunneling barrier width, w, of the order of a few nanometers. This analysis automatically excludes the high and low energy parts of the Fermi distributions which therefore do not contribute to current flow. This can be regarded as a band–pass filter.

The band–pass filtering may in principle lead to steeper subthreshold slopes, enabling further down–scaling of the operation voltage. However, to accurately quantify the subthreshold swing, we need to model the expression for the transmission probability. If we did not and assumed that it jumps from 0 (carriers blocked by the barrier) to 1 (w appropriately small to trigger BTBT), we would get an infinite slope.

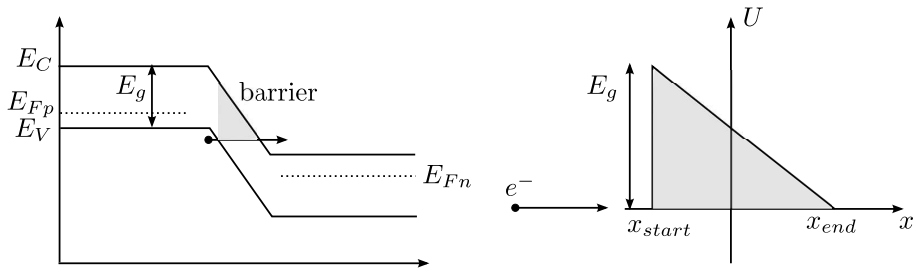

Figure 4.6: BTBT process approximated by a triangular barrier with the electron tunneling at the base of the triangle.

4.3.2 Modeling the transmission probability

Let us focus on the tunneling process that happens at the source/channel junction in the n–channel mode. The bands profile is similar to that shown in Fig. 4.2 (a). An accurate way to describe the transmission probabilities through the barrier was developed by Sze [3] using the WKB approximation and a triangularly shaped potential barrier, as depicted in Fig. 4.6.

The 1D expression for the tunneling transmission probability is given by

$$T(E) \approx \exp\left[-2\int_{x_{start}}^{x_{end}} k(x)dx\right] \tag{4.5}$$

with $k(x)$ the wave vector of the electron inside the barrier, which is given by (using the parabolic band approximation)

$$k(x) = \sqrt{\frac{2m^*}{\hbar^2}\left(U(x) - E\right)}. \tag{4.6}$$

where m^* is the electron effective mass. If we assume that, according to Fig. 4.6, the electron approaches the barrier at the bottom of the triangle, then $E = 0$ for it. The linear equation for the potential energy reads as

$$U(x) = \frac{E_g}{2} - qF(x)x. \tag{4.7}$$

We substitute this last expression into Eq. 4.6 and then into Eq. 4.5, obtaining

$$\begin{aligned} T(E) &\approx \exp\left\{-2\int_{x_{start}}^{x_{end}} \sqrt{\frac{2m^*}{\hbar^2}\left[\frac{E_g}{2} - qF(x)x\right]}dx\right\} = \\ &= \exp\left[\frac{4}{3}\frac{\sqrt{2m^*}}{qF\hbar}\left(\frac{E_g}{2} - qFx\right)^{3/2}\right]\Bigg|_{x_{start}}^{x_{end}}. \end{aligned} \tag{4.8}$$

To perform the integration we have assumed that the electric field, $F(x)$, is uniform along the integration path ($F(x) \equiv F$) which is consistent with the approximated shape of the barrier of Fig. 4.6. This assumption is made in the so–called semiclassical local models and allows to obtain an analytical expression for the transmission probability. A more realistic approximation is incorporated in non–local models, which explicitly regard F as a function of x given its dependence on the band bending at every point inside the barrier. As a consequence, in non–local models, $T(E)$ is more accurately described. Nonetheless, unlike what happens in local models, the integration can not be solved analytically but only numerically.

Evaluating Eq. 4.8, we get[1]

$$\left.\begin{array}{l} \frac{E_g}{2} - qFx_{end} = 0 \\ \frac{E_g}{2} - qFx_{start} = E_g \end{array}\right\} \implies T(E) \approx \exp\left(-\frac{4\sqrt{2m^*}E_g^{3/2}}{3qF\hbar}\right). \tag{4.9}$$

The electric field may be replaced [150] by $F = E_g/qw$ in the case of the situation shown in Fig. 4.6 (with $w = x_{end} - x_{start}$), thus leaving $T(E)$ as a function of the width and height of the barrier

[1] In the case of *excess current* reported in Sec. 4.2.1, E_g should be replaced by E_x in Eq. 4.9.

$$T(E) \approx \exp\left(-\frac{4w\sqrt{2m^*E_g}}{3\hbar}\right). \tag{4.10}$$

The BTBT current associated with this transmission probability can be calculated replacing Eq. 4.9 into Eq. 4.4, which yields the result obtained in [46, 47] and also reported in [3]

$$I_{btbt} = \frac{Aq^2F}{36\pi\hbar^2}\sqrt{\frac{2m^*}{E_g}}D\exp\left(-\frac{4\sqrt{2m^*}E_g^{3/2}}{3qF\hbar}\right), \tag{4.11}$$

where the integral D is

$$D = \int \left[f_C(E) - f_V(E)\right]\left[1 - \exp\left(-\frac{2E_S}{\bar{E}}\right)\right]dE, \tag{4.12}$$

with $\bar{E}$ and E_S given by

$$\bar{E} = \frac{\sqrt{2}q\hbar F}{\pi\sqrt{m^*E_g}} \tag{4.13}$$

$$E_S = \min\left(E_{Vp} - E, E - E_{Cn}\right). \tag{4.14}$$

An approximate solution to Eq. 4.4 can also be found in [157] with the introduction of the screening length λ. The idea is that an ultrathin body in combination with thin gate oxides gives rise to an exponential screening of potential variations on the length scale λ given by

$$\lambda = \sqrt{\frac{\epsilon_{Si}}{\epsilon_{ox}}t_{ox}t_{Si}} \tag{4.15}$$

for a single gate device. Double gate devices need to replace t_{Si} by $t_{Si}/2$. Originally, these equations for λ were introduced to be applied to MOSFETs (recall, for example, Eq. 2.46), but have also been used with

TFETs. If we assume that the spatial extent of the tunneling region in the ON–state is of order of λ, and use the $\Delta\Phi$ parameter shown in Fig. 4.4 (a) and Fig. 4.5 as a measure of the energy range for which tunneling is possible, the following expression is derived

$$I_{btbt} \propto \exp\left[-\frac{4\lambda\sqrt{2m^*}E_g^{3/2}}{3\hbar\left(E_g+\Delta\Phi(V_{gs})\right)}\right]\Delta\Phi. \tag{4.16}$$

Note that the gate voltage enters indirectly in this expression since $\Delta\Phi$ is a function of V_{gs}. However this approximation has a few shortcomings as pointed out in [150]:

- It assumes that w at the onset of tunneling equals λ. Nevertheless, at the onset of BTBT in TFETs, w is first very high and then — only as V_{gs} increases— it becomes comparable to λ.
- In TFETs, the potential profile must be determined as accurately as possible given the exponential dependence of tunnel current on w. Hence, regarding w as constant along the available BTBT range of energies may lead to spurious results.
- The screening length picture does not consider a dependence on the source doping, although it strongly impacts the TFET performance.

However, this rough estimation based on λ may be valuable to extract some conclusions:

(i) The smaller the values of m^* and λ, the higher the BTBT current.

(ii) In principle, a reduction of E_g should also lead to an increase of tunneling probabilities. Nevertheless, a small energy gap also produces an increase of the OFF-state leakage current because of thermal emission. A suitable value for E_g must be chosen in order to get a desired I_{ON}/I_{OFF} ratio.

4.3.3 Direct and indirect tunneling

The tunneling process can be either direct or indirect depending on the band structure of the considered material. In direct materials, such as GaAs and GaSb, electrons can tunnel from the vicinity of the bottom of the conduction band to the vicinity of the valence band maximum without a change of momentum (in k–space). This is possible because the conduction band minimum and valence band maximum have the same momentum. On the other hand, in indirect materials —those for which conduction band minimum and valence band maximum are not aligned at the same momentum—, like Si and Ge, the tunneling process implies a change of momentum of the carriers that needs to be supplied by either phonon or impurities scattering [3]. For phonon scattering [158], both energy and momentum have to be preserved. That means that the sum of the phonon energy and that of the electron must equal the final energy of the electron after tunneling. And the same must be satisfied for the momentum. In Fig. 4.7, we superimpose to the band diagram the $E - k$ relationship of the bands at the points where tunneling takes place.

Direct tunneling may also occur in indirect materials if the supplied voltage is large enough to allow carriers to tunnel from the secondary higher conduction band minimum (Γ point) rather than from the

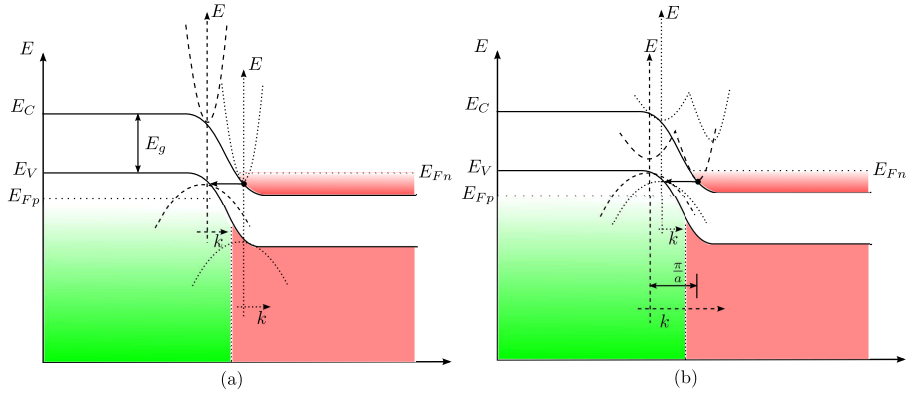

Figure 4.7: Direct (a) and indirect (b) tunneling processes along with the $E - k$ relationships of conduction and valence bands.

lower principal minimum [159]. Normally, indirect tunneling has a much lower probability to occur than direct tunneling when both are possible. The same happens for indirect tunneling if we consider the number of phonons involved in the scattering process: the odds for BTBT assisted by one single phonon are much higher.

In indirect tunneling, the transmission probability is reduced due to the necessary scattering processes. The expression for $T(E)$ shown in Eq. 4.9 has to be reduced by a multiplier and E_g has to be replaced by $E_g + E_p$, where E_p is the phonon energy [46, 47, 160].

4.3.4 Inversion layer formation in TFETs

Similarly to MOSFETs, the capacitance of the semiconductor —usually known as quantum capacitance, C_q [161,162]— increases in TFETs when an inversion layer appears at the gate insulator interface as a result of the applied voltage at the gate. The formation of this inversion layer is of utmost importance in TFETs as it affects the relation between V_{gs} and the surface channel potential, ϕ_s (which for TFETs will be identified with the conduction band edge of the semiconductor at the insulator interface), which in turn controls the BTBT width, w. That relationship is governed by the capacitive voltage divider resulting of the series combination of the oxide capacitance, C_{ox}, and C_q

$$\Delta\phi_s = \frac{C_{ox}}{C_{ox} + C_q}\Delta V_{gs}. \tag{4.17}$$

When the inversion layer is not present, $C_q << C_{ox}$ and almost a one–to–one correspondence exists between V_{gs} and ϕ_s. This is known as "quantum capacitance limit" [161]. On the other hand, when inversion appears, C_q becomes larger and most of the voltage drops on C_{ox}. In this case, ϕ_s can hardly be moved by the external voltage, which in TFETs means that the gate loses control over band bending and, therefore, over w.

The key issue is that, even although w is indirectly affected when band bending dependence on V_{gs} is modified, $T(E)$ depends exponen-

tially on it (recall Eq. 4.10), which makes BTBT a very sensitive mechanism to the formation of the inversion layer. A detailed analysis of the role of inversion layer is then quite advisable in TFETs [163], and has been considered as one of the most important aspects for the further improvement of their performance [152].

It was originally believed that the inversion layer should be formed to trigger BTBT in the source/channel interface [164]. However, it was later on shown that tunneling may occur without inversion layer formation [163] as long as ϕ_s is different to the source potential [151].

The role that the inversion layer plays, shielding w from the supplied voltage at the gate, allows one of the special features of TFETs: the exponential onset of its output characteristic (unlike the linear onset of that of the MOSFET). This exponential regime only appears if $C_{ox} < C_q$. Let us see why. For small V_{ds} in combination with large V_{gs}, the inversion layer is formed and $C_q >> C_{ox}$. This prevents the bands in the channel from further movement and makes the channel potential depend heavily on V_{ds} as can be seen in Fig. 4.8. In this situation, V_{ds} also controls the band bending at the source/channel junction and by extension w. In other words, when inversion layer is formed by carriers injected from the drain, they pin ϕ_s and make channel resistance negligible. Due to this, the full V_{ds} appears across the tunneling junction and this, in turn, directly affects w.

For larger V_{ds}, carriers are pulled back to the drain, C_q again becomes small compared to C_{ox}, and the band bending control returns to V_{gs}. In this situation, when inversion layer has disappeared, I_{ds} is no longer controlled by V_{ds} and the output characteristic saturates at a value $V_{ds,sat}$ which is given by

$$V_{ds,sat} = V_{gs} - V_{inv}|_{V_{ds}=0}\,, \tag{4.18}$$

for a given V_{gs}. V_{inv} is defined as the value of V_{gs} at which inversion layer is formed.

In Chapter 6, we will deal in depth with the inversion layer along

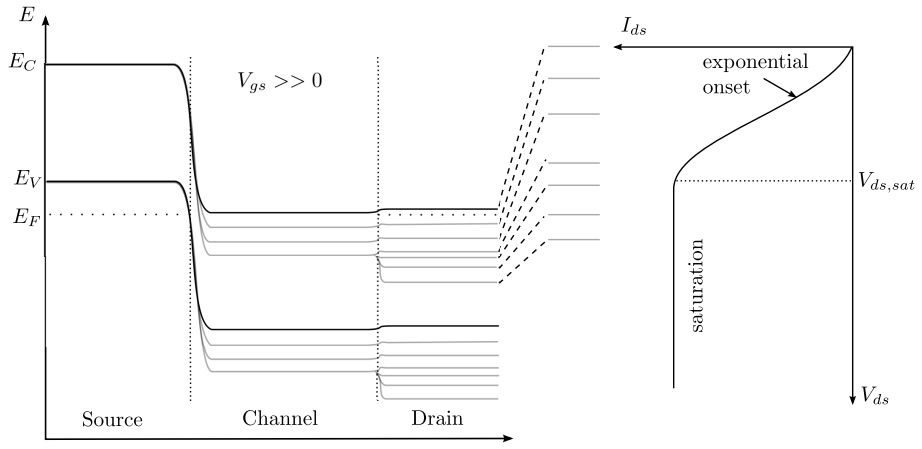

Figure 4.8: For large V_{gs} and small V_{ds}, ϕ_s is pinned. As a result, band bending at the tunnel junction is controlled by V_{ds} thus producing and exponential onset in the output characteristic. When V_{ds} increases and inversion layer disappears, the gate recovers the control over ϕ_s and V_{ds} has no longer influence on the current.

with quantum confinement and show that it turns to be determinant in the behaviour of electrical parameters —such as subthreshold swings or threshold voltages— for ultrathin body devices.

4.3.5 Subthreshold swing in TFETs

In TFETs, subthreshold swing, SS is not constant but rather a function of V_{gs}. This behaviour clearly differs from that of conventional MOSFETs, in which SS is a constant and does not vary with V_{gs}.

Let us illustrate this assertion for TFETs. SS is defined as the ΔV_{gs} required to change I_{ds} by one order of magnitude, i.e., one decade [3]

$$SS = \left[\frac{d\log(I_{ds})}{dV_{gs}}\right]^{-1} = \ln 10 \left[\frac{d\ln(I_{ds})}{dV_{gs}}\right]^{-1}. \quad (4.19)$$

To calculate it, we rewrite Eq. 4.11, which describes the BTBT current in tunnel diodes, and make explicit the terms that depend on V_{gs}

$$I_{btbt} = f(V_{gs})\, F(V_{gs}) \exp\left(-\frac{C}{F(V_{gs})}\right), \tag{4.20}$$

with

$$f(V_{gs}) = \frac{Aq^2}{36\pi\hbar^2}\sqrt{\frac{2m^*}{E_g}}D(V_{gs}) \qquad C = \frac{4\sqrt{2m^*}E_g^{3/2}}{3q\hbar}. \tag{4.21}$$

SS then results

$$\begin{aligned} SS &= \ln 10\left[\frac{d\ln f(V_{gs})}{dV_{gs}} + \frac{d\ln F(V_{gs})}{dV_{gs}} + \frac{d}{dV_{gs}}\left(-\frac{C}{F(V_{gs})}\right)\right]^{-1} = \\ &= \ln 10\left[\frac{1}{f(V_{gs})}\frac{df(V_{gs})}{dV_{gs}} + \left(\frac{F(V_{gs})+C}{F^2(V_{gs})}\right)\frac{dF(V_{gs})}{dV_{gs}}\right]^{-1}, \end{aligned} \tag{4.22}$$

which is very similar to what was obtained, for example, in [48]. In light of Eq. 4.22, it becomes obvious that SS strongly varies with V_{gs}. Swing is smallest at the lowest V_{gs} for which BTBT occurs, and increases as V_{gs} does likewise.

As a consequence of this variation, two different SS are defined in TFETs: the point swing, S_{pt}, and the average swing, S_{av}. S_{pt} is the smallest swing anywhere in the $I_{ds} - V_{gs}$ curve, and in most cases coincides with the point where BTBT starts. On the other hand, S_{av} is the swing taken from the point where BTBT begins, up to the threshold voltage[2]. These two swings are qualitatively shown in Fig. 4.9 along with the conventional MOSFET characteristic. Note that typically the TFET has a lower ON–state and OFF–state current. S_{av} is the most important swing for switch performance.

We can see that, unlike conventional MOSFETs where SS is a func-

[2]There is not an unified definition of the threshold voltage in TFETs. Some authors use the constant current technique (usually 10^{-7}A/μm). In our case, we choose a more physically–based definition and regard the threshold voltage as the voltage at which the control that the corresponding electrode exerts over the current changes from quasi–exponential to linear. Further details will be given in Chapter 6.

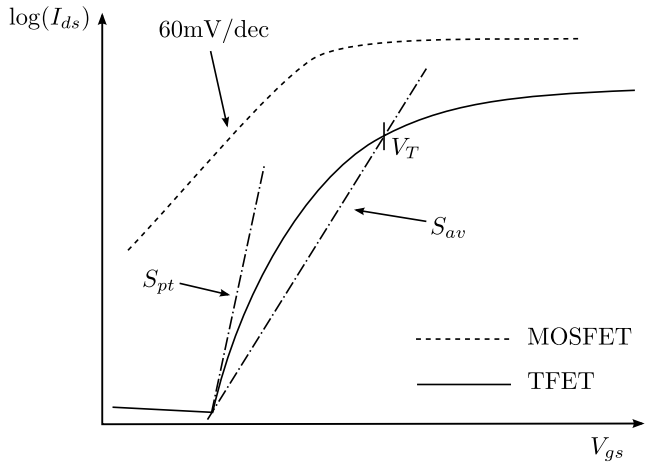

Figure 4.9: Qualitative comparison between conventional MOSFET and TFET characteristics. S_{pt} and S_{av} are depicted for TFETs showing the non-uniformity of SS.

tion of the thermal factor kT/q, SS does not depend on T to a first approximation in TFETs. This fact is not surprising since tunneling currents are weakly dependent on temperature. However, this does not imply that there are no degradation in S_{pt} or S_{av}. Degradation with T indeed exists given that rising temperatures clearly affect leakage current by increasing it, and making the steepest region of the curves disappear. What Eq. 4.22 indicates is that beyond the leakage level, current characteristics remain almost unaltered with temperature variations.

Chapter 5

Semiclassical and quantum simulations in TFETs

The quantum mechanical description implemented hereinafter in this chapter aims to go beyond the conventional semiclassical approaches that are commonly used in the literature. An example of how semiclassical models account for the effects of quantum mechanical confinement might be the inclusion of a quantum correction potential in the electron and hole current densities (as it happens in the density gradient model). Unfortunately, such models do not reproduce subband quantization and consequently regard conduction and valence bands as continuous instead of as a discrete spectrum of energy levels.

The actual dicrete nature of the bands alters the bandgap, increasing its effective value. This affects the BTBT probabilities of carriers through the energy barrier width, w, between the source and the channel. Semiclassical models will thus yield erroneous results as they predict nonzero current before tunneling conditions for the first subbands are established.

This chapter is devoted to present simulation results with and without quantum corrections and analyze their importance depending on the chosen device parameters. But prior to the introduction of the aforementioned quantum mechanical approach arising from the energy bands quantization, we must outline a brief description of the available BTBT

models to account for tunneling injection in ATLAS. These models are commonly grouped into two categories: local models, in which the electric field is regarded as constant across the tunneling junction; and non–local models, that allow spatial dependence of the electric field inside the forbidden gap. Along this chapter and the following, we will deal almost exclusively with non–local models —as they provide a more accurate description of BTBT—, and occasionally (see Sec. 6.1.2) with local models when analytical expressions need to be considered.

5.1 Band–to–band tunneling models

5.1.1 Local BTBT models

We have already established that the carrier injection mechanism that governs the TFET performance is BTBT. Consequently, an adequate choice of model accounting for it proves to be crucial in order to get an accurate description of these novel devices. Traditionally, BTBT models have been grouped in what has been labelled as "local models". Local models are based on a semiclassical conception of the tunneling process of carriers from the source into the channel. They reduce BTBT from its full quantum mechanical description with incoming wave functions, transmission probabilities and outgoing wave functions to a point–to–point process described by a generation rate per unit volume, G. In this framework, carriers disappear at the valence band edge in the source (this is the case of electrons, for example, in the n–channel ON–state) and reappear at the conduction band edge inside the channel. Total current must therefore be calculated through the integration of G over the entire device.

$$I_{semiclass} = q \iiint G \, dx \, dy \, dz \tag{5.1}$$

As the tunneling process is a point–to–point process, this results in a major impact of the exact choice of tunnel paths on the calculated current. Several analysis of how current levels depend on the selected

integration tunnel path when following the formalism of the generation rate functions may be found in [150, 151, 165]. In [150], it is shown, for example, that the shortest straight tunnel paths result in the largest current for the single gate device.

An alternative formulation, but also implying the concept of pair generation rate, may be found in [166] where the amplitude for an electron in the valence band μ at position $\mathbf{r}_{in}$ to tunnel into a conduction band state in band ν at the location $\mathbf{r}_{out}$ is calculated as a Feynman integral over all possible trajectories in the gap

$$\langle c, \nu, \mathbf{r}_{out}|\ v, \mu, \mathbf{r}_{in}\rangle \quad = \quad \sum_{\text{paths}(\mathbf{r}(s),\mathbf{k}(s))} \mathcal{B} \exp\left[i \int \mathbf{k}(s) \cdot d\mathbf{r}(s) \right] \quad (5.2)$$

with $\mathcal{B}$ a prefactor depending on the considered BTBT model (Kane, Klaasen, Schenk, Tanaka...). The last integral is performed selecting the path that determines an extreme point of the action integral.

A way to calculate the generation rate function, G, based on the works done by Keldysh and Kane [46, 160, 167] was shown [168] to be

$$G = -\frac{1}{q} \nabla \cdot \mathbf{J}_{btbt} = -\frac{1}{q} \frac{d\mathbf{J}_{btbt}}{dV} \cdot \nabla V = -\frac{d\mathbf{J}_{btbt}}{dE} \cdot \mathbf{F}, \quad (5.3)$$

with $\mathbf{F}$ the value of the electric field, which in local models is taken to be uniform along the tunneling junction. Current density may be calculated from Eq. 4.4.

The different local models existing in the literature differ in the expression that they consider for G. We briefly present hereinafter some of them currently implemented in the Silvaco ATLAS simulator as they are described in [9].

5.1.1.1 Kane model

This has been traditionally one of the most used models to account for BTBT processes. In it, G is written as

$$G_{Kane} = \frac{DA_{Kane}}{\sqrt{E_g}} F^{\gamma} \exp\left(-B_{Kane}\frac{E_g^{3/2}}{F}\right), \tag{5.4}$$

where, as indicated above, F is the magnitude of the electric field, E_g is the position dependent bandgap, A_{Kane}, B_{Kane} and γ are model constants and D is a statistical factor. The value of γ may be adjusted to account for direct or indirect transitions depending on the considered material. ATLAS uses as default values

$$\begin{aligned} D &= 1, \\ \gamma &= 2, \\ A_{Kane} &= 3.5 \times 10^{21} \mathrm{eV}^{1/2}\mathrm{cm}^{-1}\mathrm{s}^{-1}\mathrm{V}^{-2}, \\ B_{Kane} &= 2.25 \times 10^{7} \mathrm{Vcm}^{-1}\mathrm{eV}^{-3/2}. \end{aligned} \tag{5.5}$$

The Kane model can be modified by changing the definition of the statistical factor D as

$$D = \frac{n_{ie}^2 - np}{(n + n_{ie})\,(p + n_{ie})}, \tag{5.6}$$

with n, p the electron and hole concentrations; and n_{ie} the effective intrinsic concentration. When this modification is introduced, we refer to the model as the local Hurkx model [168].

5.1.1.2 Standard and Klaassen models

This two models share the following expression for the generation rate function

$$G_{Sta/Kla} = DA_{Sta/Kla} F^{\gamma_{Sta/Kla}} \exp\left(-\frac{B_{Sta/Kla}}{F}\right). \tag{5.7}$$

In which the explicit E_g dependence has disappeared and is now implicitly accounted for through the electric field. Depending on the constants, we will deal with one or the other. The default values for each model are

$$\begin{aligned} \gamma_{Sta} &= 2, \\ A_{Sta} &= 9.6615 \times 10^{18}\mathrm{cm}^{-1}\mathrm{V}^{-2}\mathrm{s}^{-1}, \\ B_{Sta} &= 3.0 \times 10^{7}\mathrm{Vcm}^{-1}. \end{aligned} \tag{5.8}$$

$$\begin{aligned} \gamma_{Kla} &= 2.5, \\ A_{Kla} &= 4.00 \times 10^{14}\mathrm{cm}^{-1/2}\mathrm{V}^{-5/2}\mathrm{s}^{-1}, \\ B_{Kla} &= 1.9 \times 10^{7}\mathrm{Vcm}^{-1}. \end{aligned} \tag{5.9}$$

The standard model should be used for direct transitions, whereas the Klaassen model is more suitable for indirect transitions. Both of them allow a modification that makes the simulator calculate these constants from first principles as

$$\begin{aligned} A &= \frac{q^2}{h^2}\sqrt{\frac{2m^*}{E_g}}, \\ B &= \frac{4\sqrt{2m^*}E_g^{3/2}}{3q\hbar}, \\ \gamma &= 2. \end{aligned} \tag{5.10}$$

In this case, both models become the same and B exactly matches the factor inside the exponential in Eq. 4.11.

5.1.1.3 Schenk model

The last local model that we present is the Schenk model [158]. This model is especially designed to be used with indirect materials because it assumes that the participation of a phonon in the BTBT process

is essential. This model consequently neglects direct tunneling. The expression for G then reads as

$$G_{Schenk} = A_{Schenk} F^{7/2} S \left[\frac{(A^{\mp})^{-3/2} \exp\left(\frac{A^{\mp}}{F}\right)}{\exp\left(\frac{C_{Schenk}}{kT}\right) - 1} + \frac{(A^{\pm})^{-3/2} \exp\left(\frac{A^{\pm}}{F}\right)}{1 - \exp\left(\frac{-C_{Schenk}}{kT}\right)} \right], \quad (5.11)$$

where S is a statistical factor dependent on carrier concentrations (similar to what happened in the Hurkx model), $\hbar\omega$ is the energy of the transverse acoustic phonon and $A^{\pm}$ is given by

$$A^{\pm} = B_{Schenk} \left(\hbar\omega \pm C_{Schenk}\right)^{3/2}. \quad (5.12)$$

The plus sign stands for tunneling generation of electron–hole pairs (reverse–biased junction), and the minus sign for recombination (forward–biased junction). The default constants for this model are

$$\begin{aligned} A_{Schenk} &= 8.977 \times 10^{20} \mathrm{cm^2 s^{-1} V^{-2}}, \\ B_{Schenk} &= 2.14667 \times 10^{7} \mathrm{eV^{-3/2} V cm^{-1}}, \\ C_{Schenk} &= 0.0186\,\mathrm{eV}. \end{aligned} \quad (5.13)$$

5.1.2 Non–local BTBT model

Unlike conventional local models, in non–local models the magnitude of the electric field is allowed to vary across the tunneling junction. The approximation made in Eq. 4.8 is therefore no longer valid and the simulator carries out numerically the integration to get the tunneling probabilities. As a result, the values for $T(E)$ strongly depend on the band bending along the junction and on the energy bandgap.

The simulator extends the 1D formulation of Sec. 4.3.2 to 2D by dividing the tunneling region into multiple horizontal slices perpendicular to the y–direction (that of the source/channel junction in our proposed

devices) [9], and performing for each one of them —and for each longitudinal energy $E_{\parallel}$ for which BTBT is allowed— the following integration

$$T\left(E_{\parallel}, y\right) = \exp\left[-2\int_{x_{start}}^{x_{end}} k(x,y)dx\right]. \tag{5.14}$$

x_{start} and x_{end} represent the limits of the horizontal integration path following the x–direction. Observe that in this integration we are clearly treating in a separate way the x and y coordinates. In that sense, this integration approach differs from that of Eq. 5.2, in which all possible trajectories had to be considered prior to selecting that one minimizing the action.

$k(x,y)$ is the wave vector of carriers inside the barrier and is given by

$$k(x,y) = \frac{k_e k_h}{\sqrt{k_e^2 + k_h^2}}, \tag{5.15}$$

with

$$k_e = \frac{1}{i\hbar}\sqrt{2m_e^*(x,y)\left(E - E_C(x,y)\right)}, \tag{5.16}$$

$$k_h = \frac{1}{i\hbar}\sqrt{2m_h^*(x,y)\left(E_V(x,y) - E\right)}. \tag{5.17}$$

E_C and E_V are the position dependent conduction and valence band energies, respectively. m_e^* and m_h^* are the electron and hole tunneling effective masses. Eqs. 5.15–5.17 ensure that the energy dispersion relationship is electron–like near the conduction band, hole–like near the valence band, and approximately mixed in between.

If quantum confinement is included, as will be later discussed in Sec. 5.2.2, the formerly continuous conduction and valence bands become a discrete spectrum of energy subbands and, subsequently, E_C and E_V in Eqs. 5.16 and 5.17 should be replaced by E_i and E_ν, being the first bound

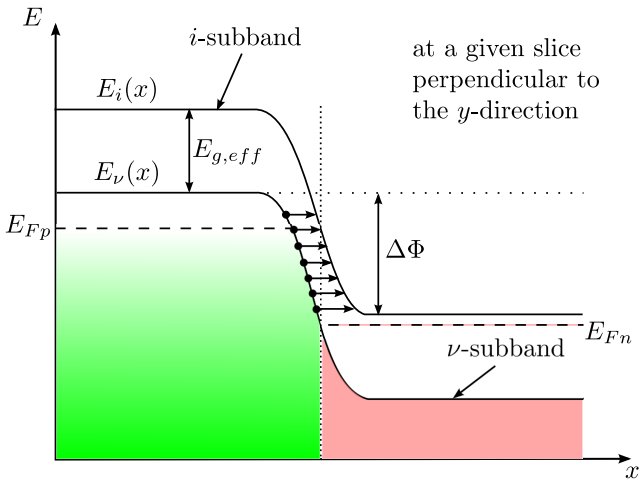

Figure 5.1: Schematic representation of several horizontal integration paths for the BTBT allowed range of energies, $\Delta\Phi$. Quantum confinement has been assumed and, therefore, tunneling takes place between the first bound state of the valence band (ν–subband) and the first bound state of the conduction band (i–subband).

states of the quantized conduction and valence bands, respectively[1]. In that situation, BTBT will occur between this two bound states. This will make the effective bandgap, E_g, and the tunneling barrier width, w, increase. As a result, BTBT current will be significantly reduced. Such a situation at a given horizontal slice perpendicular to the y–direction, and including quantum confinement, is schematically shown in Fig. 5.1 along with some integration paths.

Current density is calculated as

$$J(y) = \frac{q}{\pi\hbar}\iint T\left(E_{\parallel}, y\right)\left[f_l\left(E_{\parallel}+E_T\right)-f_r\left(E_{\parallel}+E_T\right)\right] \rho\left(E_T\right) dE_{\parallel}\, dE_T, \tag{5.18}$$

with E_T the transverse energy, $f_{l(r)}$ the Fermi–Dirac distribution on the left(right) hand side of the junction corresponding to the majority carrier

[1]In fact, the spatial dependence of E_i and E_ν is only on the x–direction, i.e. $E_i(x)$ and $E_\nu(x)$. See Sec. 5.2.2.

$$f_{l(r)}(E_{\parallel}+E_T) = \frac{1}{1+\exp\left(\frac{E_{\parallel}+E_T-E_{Fl(r)}}{kT}\right)}, \tag{5.19}$$

and $\rho(E_T)$ the 2D density of states corresponding to the two transverse components of the wave vector, which therefore depends on the electron and hole transverse effective masses

$$\rho(E_T) = \frac{\sqrt{m_e^T m_h^T}}{2\pi\hbar^2}. \tag{5.20}$$

By integrating over transverse energies in Eq. 5.18, we can get the contribution to the BTBT current corresponding to the longitudinal energy range from $E_{\parallel}-\Delta E_{\parallel}/2$ to $E_{\parallel}+\Delta E_{\parallel}/2$

$$J\left(E_{\parallel},y\right)\Delta E_{\parallel} = \frac{qkT\sqrt{m_e^T m_h^T}}{2\pi^2\hbar^3} T\left(E_{\parallel},y\right) \left| \ln\left[\frac{1+\exp\left(\frac{E_{\parallel}+E_T-E_{Fr}}{kT}\right)}{1+\exp\left(\frac{E_{\parallel}+E_T-E_{Fl}}{kT}\right)}\right]\right|_0^{E_{max}} \Delta E_{\parallel}, \tag{5.21}$$

where the upper limit of evaluation, E_{max}, is

$$E_{max} = \min\left(E_{\parallel}-E_{lower}, E_{upper}-E_{\parallel}\right), \tag{5.22}$$

provided that E_{upper} and E_{lower} are the limits of the $\Delta\Phi$ energy interval; and the total carrier energy satisfies $E_{lower} \leq E_{\parallel}+E_T \leq E_{upper}$. Substituting the evaluation limits we obtain

$$J\left(E_{\parallel},y\right)\Delta E_{\parallel} = \frac{qkT\sqrt{m_e^T m_h^T}}{2\pi^2\hbar^3}T\left(E_{\parallel},y\right) \ln\left\{\frac{\left[1+\exp\left(\frac{E_{\parallel}+E_{max}-E_{Fr}}{kT}\right)\right]\left[1+\exp\left(\frac{E_{\parallel}-E_{Fl}}{kT}\right)\right]}{\left[1+\exp\left(\frac{E_{\parallel}+E_{max}-E_{Fl}}{kT}\right)\right]\left[1+\exp\left(\frac{E_{\parallel}-E_{Fr}}{kT}\right)\right]}\right\}\Delta E_{\parallel}. \tag{5.23}$$

The calibration of the non–local model is carried out by adjusting the electron and hole effective masses. ATLAS assumes non–local BTBT as being an elastic process [9] and does not incorporate neither phonon–assisted nor trap–assisted processes.

5.2 Quantum effects modeling

In ATLAS, there are several models that aim to account for different effects arising from quantum mechanical confinement [9]. In what follows, we outline two of them that in principle are compatible with the aforesaid, more accurate, non–local BTBT model. They are the Density Gradient model and the Self–Consistent Coupled Schrödinger–Poisson model. Even when the second one is clearly more complete, we have also included the other since, traditionally, it has been widely used as an initial approach to quantum confinement.

5.2.1 Density Gradient model

This model is based on the moments of the Wigner function equations of motion. It can successfully reproduce the quantum carrier concentrations and transport properties, but, however, it fails to reproduce the bound state energies or wave functions given by the Schrödinger–Poisson model.

In this model, a quantum correction potential Λ has to be included in the electron and hole current densities [169, 170]

$$\mathbf{J}_n = qD_n\nabla n - qn\mu_n\nabla(V-\Lambda) - \mu_n n\left[kT\nabla\left(\ln n_{ie}\right)\right], \quad (5.24)$$

$$\mathbf{J}_p = -qD_p\nabla p - qp\mu_p\nabla(V-\Lambda) + \mu_p p\left[kT\nabla\left(\ln n_{ie}\right)\right], \quad (5.25)$$

which is calculated using either of the two following expressions

$$\begin{aligned}\Lambda &= -\frac{\gamma h^2}{12m^*}\left[\nabla^2\log n + \frac{1}{2}\left(\nabla\log n\right)^2\right],\\ \Lambda &= -\frac{\gamma h^2}{6m^*}\frac{\nabla^2\sqrt{n}}{\sqrt{n}}.\end{aligned} \quad (5.26)$$

with γ a fitting parameter, m^* the carrier effective mass, and n stands for the carrier concentration (electrons or holes as necessary).

5.2.2 Self–Consistent Coupled Schrödinger–Poisson model

This model allows to self–consistently solve the Schrödinger and the Poisson equations for electrons, for holes, or for both of them. In the 2D simulator that we employ, the Schrödinger equation can be solved in 1D slices or 2D plane. When cylindrical coordinates are employed, the Schrödinger equation is solved in radial direction for different orbital quantum numbers and for all slices perpendicular to the axis.

In our simulations, quantum confinement will occur in one dimension —which we choose to be the y–direction—. Therefore, the calculation of the quantum electron density relies upon a solution of a 1D Schrödinger equation, which has to be solved for eigen state energies E_{il} and wave functions $\Psi_{il}(x,y)$ at each slice perpendicular to the x–axis and for each electron valley l

$$-\frac{\hbar^2}{2}\frac{\partial}{\partial y}\left[\frac{1}{m_y^l(x,y)}\frac{\partial\Psi_{il}}{\partial y}\right] + E_C(x,y)\Psi_{il} = E_{il}\Psi_{il}. \quad (5.27)$$

$m_y^l(x,y)$ is a spatially dependent effective mass in the y–direction for the l valley and $E_C(x,y)$ is a conduction band edge. The equivalent equa-

tion for holes would require to replace E_{il} by $E_{\nu\alpha}$ (with α representing each hole band), $\Psi_{il}(x,y)$ by $\Psi_{\nu\alpha}(x,y)$, the electron effective masses by those corresponding to holes —$m_y^{\alpha}(x,y)$—, and $E_C(x,y)$ by $E_V(x,y)$ (the valence band edge).

The electron concentration (analogous treatment applies for holes) in 1D is then obtained using Fermi statistics along with the eigen energies and wave functions and reads as

$$\begin{aligned} n(x,y) &= 2\frac{kT}{\pi\hbar^2}\sum_l \sqrt{m_x^l(x,y)m_z^l(x,y)}\sum_{i=0}^{\infty}|\Psi_{il}(x,y)|^2 \\ &\quad \ln\left[1+\exp\left(-\frac{E_{il}-E_F}{kT}\right)\right]. \end{aligned} \tag{5.28}$$

Once the electron(hole) concentration is calculated, it is substituted into the Poisson equation and the corresponding potential is derived. The potential is brought back into the Schrödinger equation and a new modified carrier concentration is produced. This leads to a new recalculated potential, which in turn produces a new carrier concentration. This alternating process between the Schrödinger and Poisson equations is repeated until convergence is achieved and, therefore, a self–consistent solution is obtained[2]. A default predictor–corrector scheme is used to avoid instability and oscillations of Poisson convergence. The employed version of ATLAS also allows to self–consistently solve Schrödinger–Poisson equations under non–equilibrium conditions.

Given that ATLAS non–local BTBT model uses isotropic tunneling effective masses ($m_e^* = 0.322$ and $m_h^* = 0.549$), we choose to simplify the calculations regarding quantization and also assume isotropy of effective masses. The next logical step would be to differentiate tunneling and conventional effective masses making them anisotropic and dependent on the crystallographic orientation.

Note that, as long as the energies of the subbands come from a 1D quantization in the y–direction, they should only include an explicit

[2]The convergence criterion for potential in the Schrödinger–Poisson model is user–definable with a default value of 0.001V.

dependence on x. Hence, $E_C(x, y)$ and $E_V(x, y)$ in Eqs. 5.16 and 5.17 have to be replaced by $E_i(x)$ and $E_\nu(x)$.

5.3 Limitations, incompatibilities and solutions

In most of our simulations —unless otherwise specified— when quantum confinement has to be included, we will combine the non–local BTBT model with the self–consistent coupled Schrödinger Poisson model as their combination provides a more accurate description of the device operation. However, there exist two main limitations concerning the implementation and combination of these two models in ATLAS.

The first of them refers to the issue of combination. Due to incompatibilities of the numerical solvers that are used for each model, they cannot be incorporated at the same time for a given solving step. This incompatibility does not appear when local BTBT models are applied, so, we expect that this problem will be somehow fixed in future versions of the simulator. In our case, we have circumvented this hitch by dividing each calculation step into two separate iterations. In the first iteration, for a given bias condition, we self–consistently solve the Schrödinger and Poisson equations and obtain the potential, the carrier concentrations and the discrete subband spectrum of both conduction and valence bands. In the second iteration, we calculate non–local tunneling injection using the potential and charge distributions obtained in the previous iteration, and assume that BTBT of carriers does not significantly modify them. Some preliminary simulations showed us that this is a reasonable approximation.

The second limitation refers to the concept of BTBT that the simulator invariably assumes when accounting for it. It considers that the tunneling process always takes place between the conduction and valence bands and, consequently, it does not regard the possibility of those states being forbidden as it happens in the case of subband quantization. As a result, the simulator itself is not able to inject carriers (let us say, for example, electrons) from the first bound state of the valence band to the first bound state of the conduction band. However, what was a hitch in

the previous paragraph now proves to be a substantial advantage: as the Poisson equation is not solved in the second iteration of our calculation procedure —and therefore neither the potential nor the quantum charge distribution are modified—, we take profit of this and modify the bands profile in the tunnel junction for each bias point. This process represents a sort of calibration in which we increase the band gap and make it coincident with the effective gap between the first bound states (arising from the subband quantization of the first iteration). This way, when carriers are injected by means of the non–local BTBT mechanism, the modified conduction and valence bands play the role of their first bound states thus closely describing the actual situation, and giving sense to the $E_i(x)$ and $E_\nu(x)$ inclusion in Eqs. 5.16 and 5.17. To illustrate this calibration process of the effective bandgap, we will include some figures showing its behaviour when simulation results are shown.

Transport is implemented in our simulations by solving drift–diffusion equations as a postprocessing step when BTBT current is injected and considering the quantum charge distribution obtained from applying the Schrödinger–Poisson model. In addition, Shockley–Read–Hall recombination, bandgap narrowing and mobility models dependent on concentration and lateral electric field are used.

All the above constitutes, of course, not an exact solution but an approximate and interesting starting point that allows us to tackle the issue of confinement in TFETs in a way that otherwise would not be possible using ATLAS. We are aware that recently much more complex treatments based on 3-D atomistic full–band quantum transport simulators have been developed [171], but unfortunately only available to very few. On the contrary, the potential interest of the approach that we propose is not only to be an intermediate stage in between semiclassical G–based models and these 3-D atomistic models, but also the availability of ATLAS that makes it accessible to many researchers.

5.4 Device structure and operation

In what follows, we focus our attention on the single gate (SG) and double gate (DG) TFETs. In Fig. 5.2, the two considered structures are shown. The simulations are performed using the ATLAS version 5.18.3.R.

The doping profile at the source–to–channel junction is not abrupt but has a gaussian drop off profile: the uniform p^+ doping of the source decreases when approaching the junction with a gaussian characteristic length, x_{char}, of 1nm according to the following expression

$$p(x) = p_{peak} \exp\left[-\left(\frac{x}{x_{char}}\right)^2\right], \tag{5.29}$$

with p_{peak} the uniform doping at the source.

Unless otherwise explicitly specified, the default dopings that we use are: 10^{20}atoms/cm^3 for the source (p^+–region), 10^{20}atoms/cm^3 for the drain (n^+–region), and 10^{17}atoms/cm^3 for the lightly doped channel (n–region). The default dielectric is SiO_2 with t_{ox} =1nm. The BOX is also SiO_2. And the longitudinal dimensions are $L_{source} = L_{drain} = 100$nm, $L_g = 20$nm.

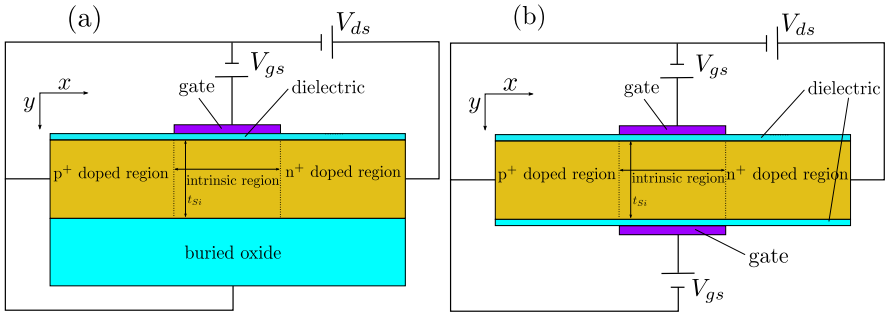

Figure 5.2: Schematic cross–section (not to scale) of the single gate (a) and double gate (b) n–channel TFETs considered throughout this chapter.

5.5 Simulation results

Let us compare the semiclassical and quantum characteristics of SG and DG–TFETs at low and high V_{ds} for different device configurations.

5.5.1 Effect of body thickness variation

5.5.1.1 Semiclassical simulations

Before including quantum confinement, Figs. 5.3, 5.4 and 5.5 show the transfer characteristics for $V_{ds} = 0.1$ and 1.1V corresponding to $t_{Si} = 3$, 5 and 10nm, respectively. As indicated at the beginning of the chapter, unless otherwise specified, we will use the non–local BTBT model.

It can be observed that for the different thicknesses, the n–branch of the DG–TFET lies over that of the SG–TFET. Furthermore, the thinness of the simulated structures allows the DG–TFET to have a greater electrostatic control of the tunnel junction, and this results in the higher values of the I_{ds}^{DG}/I_{ds}^{SG} ratio that are shown in the insets. In all cases and for the whole range of V_{gs}, this ratio remains above 2 indicating that semiclassically, for the considered thicknesses, the current is more than doubled when we design a DG configuration (in the subthreshold region, this ratio is even of several orders of magnitude).

The current levels for high V_{gs} ($V_{gs} = 2$V) when $V_{ds} = 0.1$V remain almost the same when the body thickness is reduced from 10 to 3nm, thus suggesting a saturating behaviour of I_{ds} around 0.6–0.7μA/μm for the DG and around 0.25μA/μm for the SG. On the other hand, when we increase V_{ds} to 1.1V, the saturating behaviour of I_{ds} is not so clear for the SG (current increases from 20μA/μm for 10nm to 40μA/μm approx. for 3nm), whereas for the DG it appears at 5nm ($\approx$150μA/μm).

5.5.1.2 Simulations including quantum confinement

As we previously commented, when quantum confinement is taken into account, we have to carry out a calibration process to adjust the effective bandgap in the source–to–channel junction to that existing between the first bound states arising from the conduction and valence band

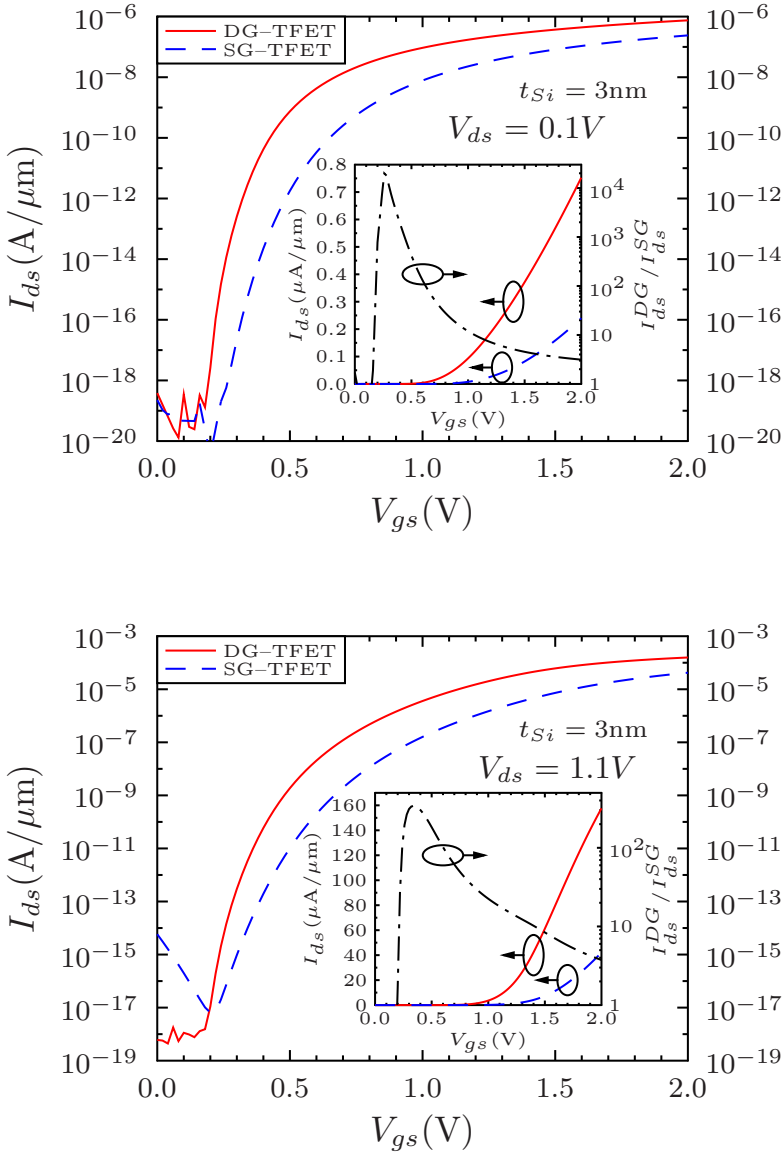

Figure 5.3: Transfer characteristics for Single Gate and Double Gate TFETs with $t_{Si} = 3$nm for $V_{ds} = 0.1$ and 1.1V. The inset shows the current on linear scale (left axis), and the ratio I_{ds}^{DG}/I_{ds}^{SG} (right axis).

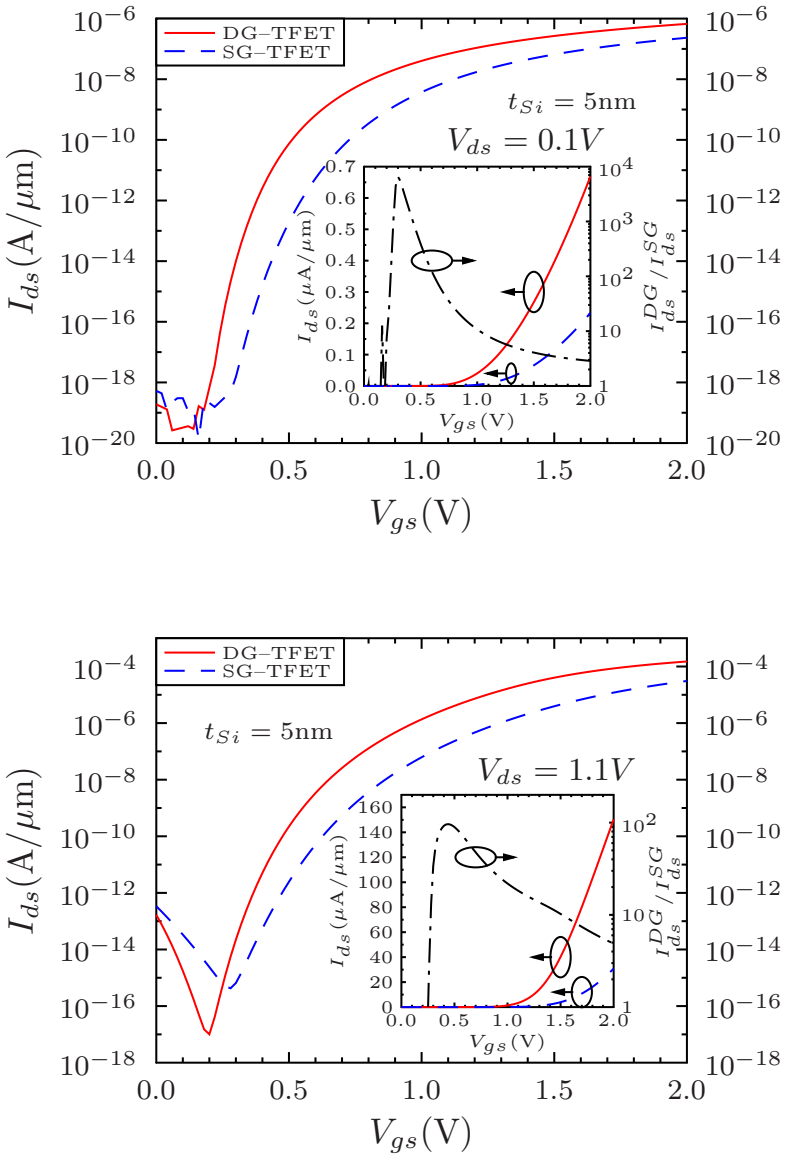

Figure 5.4: Transfer characteristics for Single Gate and Double Gate TFETs with $t_{Si} = 5$nm for $V_{ds} = 0.1$ and 1.1V. The inset shows the current on linear scale (left axis), and the ratio I_{ds}^{DG}/I_{ds}^{SG} (right axis).

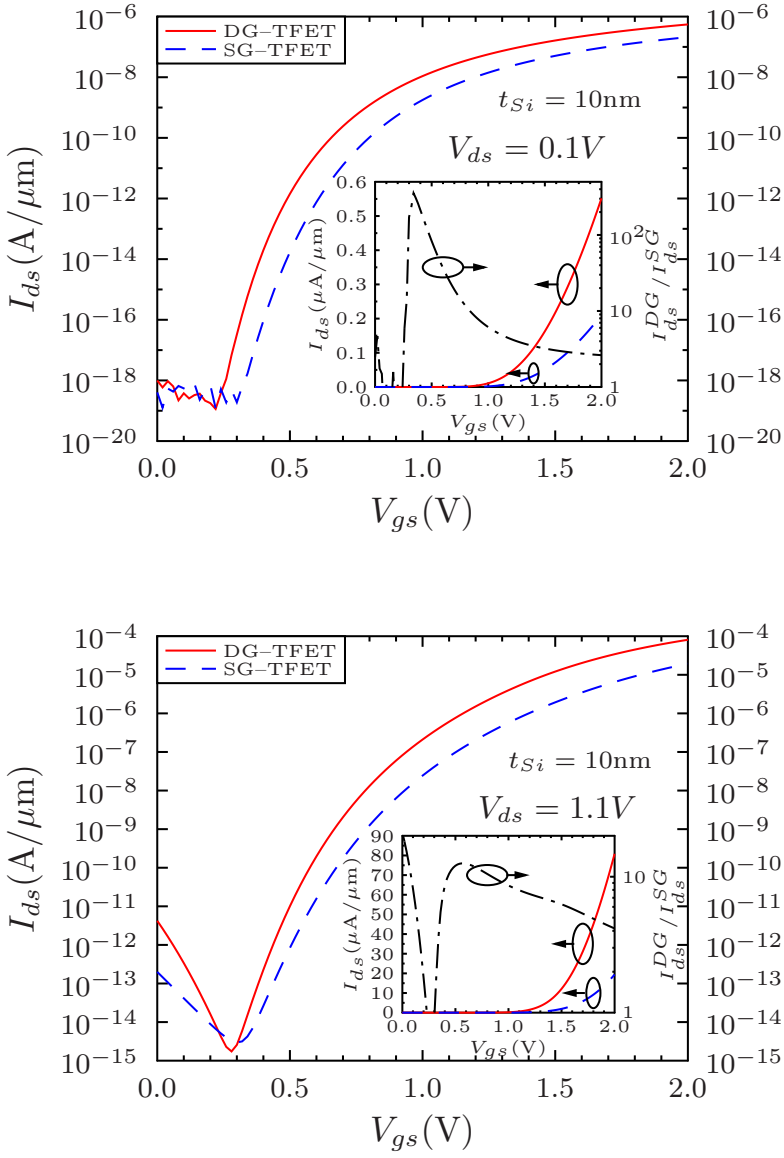

Figure 5.5: Transfer characteristics for Single Gate and Double Gate TFETs with $t_{Si} = 10$nm for $V_{ds} = 0.1$ and 1.1V. The inset shows the current on linear scale (left axis), and the ratio I_{ds}^{DG}/I_{ds}^{SG} (right axis).

quantization. As a result of this subband quantization, the effective bandgap turns out to be higher and, consequently, the current levels are significantly reduced. In Fig. 5.6, we show the transfer characteristics at $V_{ds} = 0.1$ including the effect of quantum confinement for $t_{Si} = 3$, 5 and 10nm.

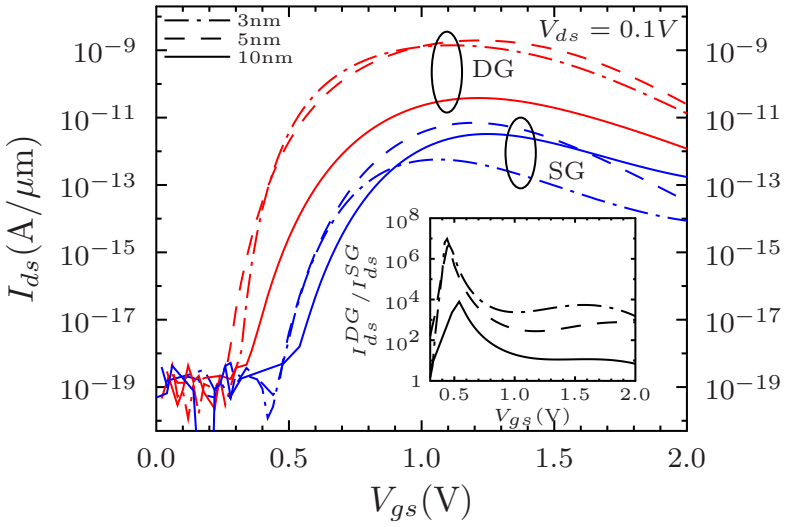

Figure 5.6: Transfer characteristics including quantum confinement for Single Gate and Double Gate TFETs with $t_{Si} = 3$, 5 and 10nm for $V_{ds} = 0.1$V. The inset shows the ratio I_{ds}^{DG}/I_{ds}^{SG}.

We clearly observe a global decreasing pattern of I_{ds} curves with V_{gs} once the subthreshold region is left behind. This was also reported in [148] for low V_{ds} when studying field–induced quantum confinement. As this behaviour is clearly connected with the inclusion of quantum confinement, it deserves some additional explanation. The key issue is the different effect that the formation of the inversion layer has over the surface channel potential, ϕ_s, depending on whether confinement is considered or not. It is known that classically the formation of the inversion layer modifies the ϕ_s dependence on the gate voltage, V_{gs}, making it stiffer to bending but not completely pinning it [163]. However, when the conduction and valence bands turn into a discrete spectrum of

subband energy levels (as a result of confinement), the surface channel potential proves to be more strongly pinned. These behaviours will be more extensively shown in Sec. 6.1.1 (Figs. 6.1 and 6.6).

The consequence of the above is that, once the gate voltage is increased beyond the value of inversion layer formation, the effective bandgap in the presence of confinement turns out to be gradually increased due to the aforementioned different pinning strength of semiclassical and quantum surface channel potentials (ϕ_s and $\tilde{\phi}_s$, respectively). To illustrate this, let us present in Fig. 5.7 a foretaste of the results shown in Sec. 6.1.1. As the bandgap plays the role of the tunneling barrier height in the region where BTBT occurs, the existence of an effectively increasing bandgap would result in a reduction of I_{ds} as it is indeed observed for $V_{ds} = 0.1$V. In Fig. 5.8, we superimpose both semiclassical and quantum corrected transfer characteristics at $V_{ds} = 0.1$V for each thickness.

Let us now study what happens when $V_{ds} = 1.1$V. In Fig. 5.9, we show the transfer characteristics at $V_{ds} = 1.1$V including the effect of quantum confinement for $t_{Si} = 3$, 5 and 10nm. The reason for not observing such a current reduction in quantum corrected $I_{ds} - V_{gs}$ curves for $V_{ds} = 1.1$V at high values of V_{gs} simply lies in the fact that in the

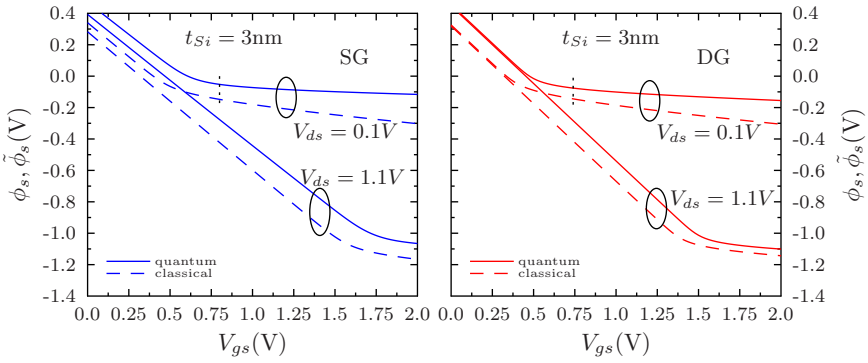

Figure 5.7: Semiclassical and quantum surface channel potentials (ϕ_s and $\tilde{\phi}_s$) vs. V_{gs} at $V_{ds} = 0.1$ and 1.1V for $t_{Si} = 3$nm. The formation of the inversion layer pins both potentials, but more tightly in the quantum case. Vertical dotted lines mark the gate voltages at which the difference $\tilde{\phi}_s - \phi_s$ begins to increase.

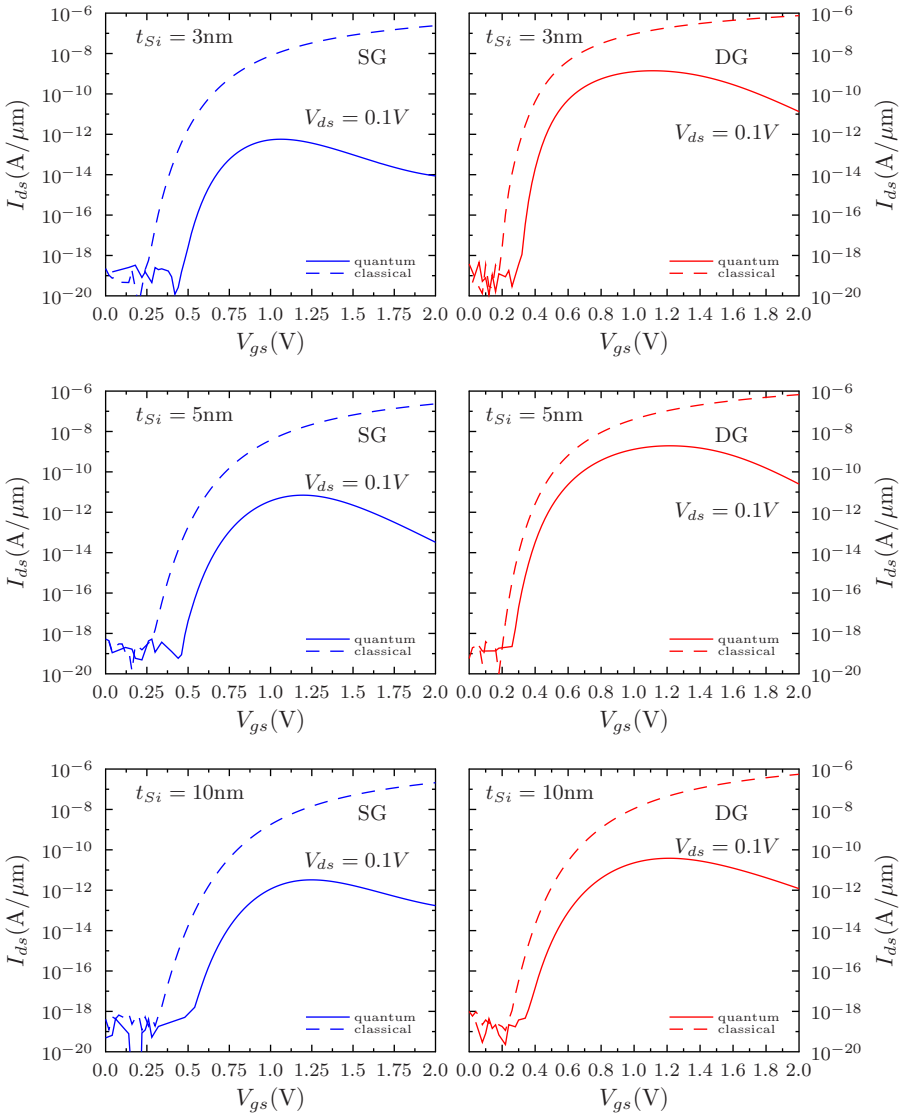

Figure 5.8: Comparison between semiclassical and quantum corrected $I_{ds} - V_{gs}$ curves for Single Gate and Double Gate TFETs with $t_{Si} = 3$, 5 and 10nm for $V_{ds} = 0.1$V.

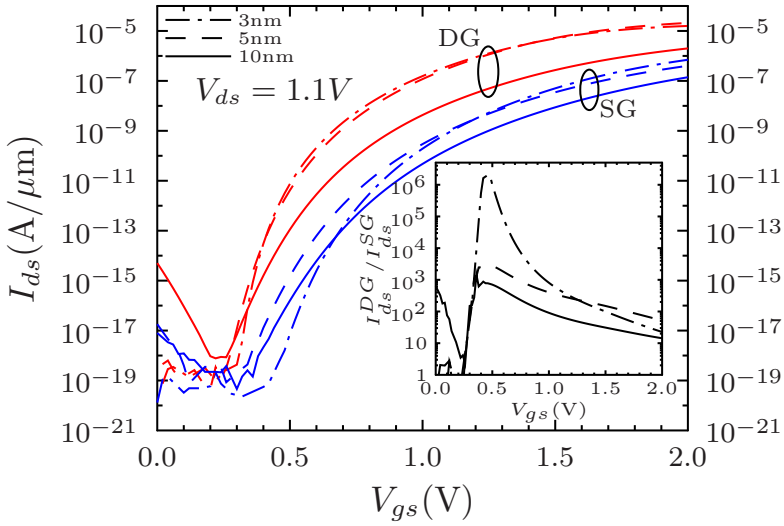

Figure 5.9: Transfer characteristics including quantum confinement for Single Gate and Double Gate TFETs with $t_{Si} = 3$, 5 and 10nm for $V_{ds} = 1.1$V. The inset shows the ratio I_{ds}^{DG}/I_{ds}^{SG}.

considered V_{gs} range, the diverting behaviour between ϕ_s and $\tilde{\phi}_s$ has not yet begun (see Fig. 5.7). This suggests that if the gate voltage were extended beyond the 2V limit for $V_{ds} = 1.1$V, one would also observe the beginning of a decreasing behaviour of I_{ds}. In fact, note that in Fig. 5.9 for the DG device at $V_{ds} = 1.1$V and high V_{gs}, the $I_{ds} - V_{gs}$ curves for $t_{Si} = 3$ and 5nm hint at the beginning of a decreasing trend. As well as we did for $V_{ds} = 0.1$V, in Fig. 5.10, we superimpose both semiclassical and quantum corrected curves at $V_{ds} = 1.1$V for the three considered thicknesses.

Focusing now on the peculiarities of both devices, we notice relevant differences between the DG and the SG structures at $V_{ds} = 0.1$V when observing in the $I_{ds} - V_{gs}$ curves of Fig. 5.6. For the DG, there is a common pattern in the whole range of V_{gs} consisting of a current increase when reducing the body thickness (from $t_{Si} = 10$ to 5nm), and a certain saturation —or even a decrease for high V_{gs}— of the current level when we go down from $t_{Si} = 5$nm to 3nm. There is also an apparent

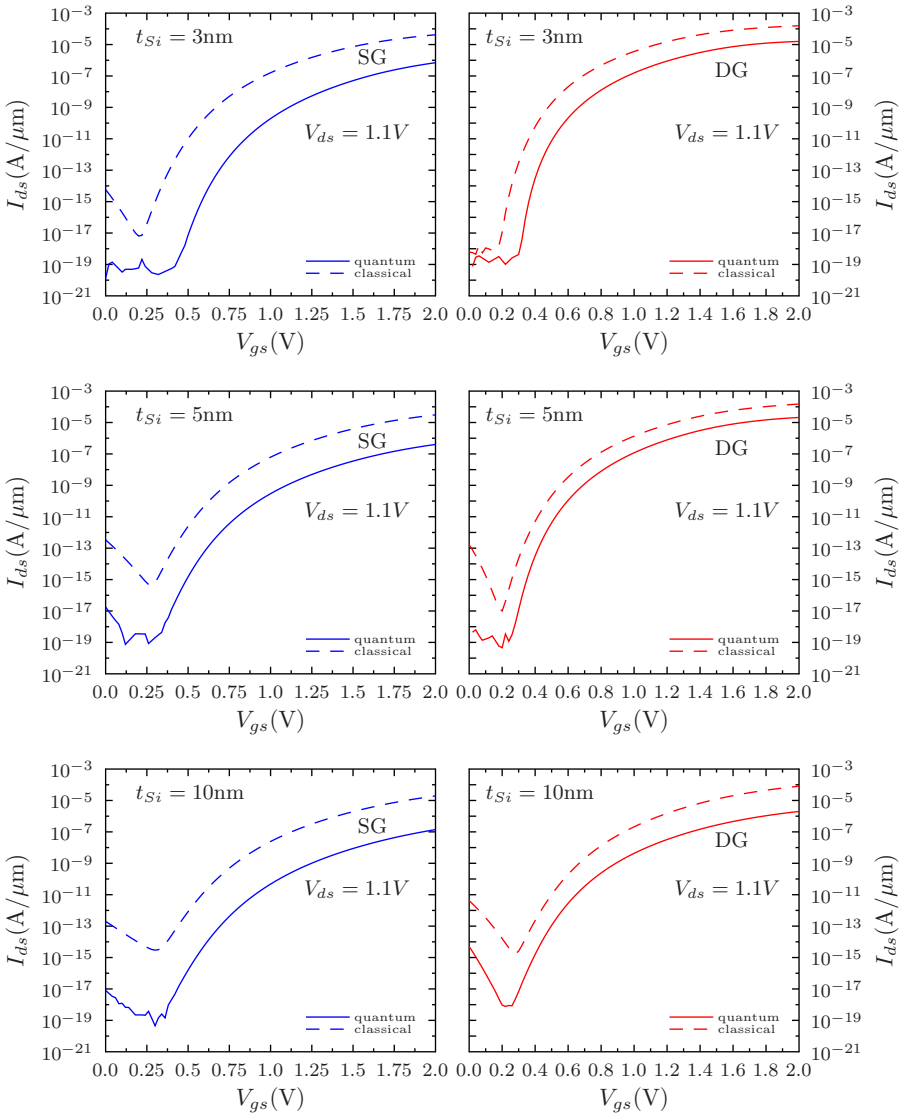

Figure 5.10: Comparison between semiclassical and quantum corrected $I_{ds} - V_{gs}$ curves for Single Gate and Double Gate TFETs with $t_{Si} = 3$, 5 and 10nm for $V_{ds} = 1.1$V.

improvement of subthrehold swings for smaller thicknesses[3]. However, in the case of the SG device, we find that when we go to high values of V_{gs}, the thinner the device gets, the lower its current becomes. As a consequence of this, for high V_{gs}, the ratio I_{ds}^{DG}/I_{ds}^{SG} approaches 10 when t_{Si} =10nm, whereas it turns to be two orders of magnitude higher when $t_{Si} = 3$ and 5nm.

When we go to $V_{ds} = 1.1$V (Fig. 5.9), both structures show a similar behaviour. We start with $t_{Si} = 10$nm and get a current level which increases as we reduce the body thickness to 5nm. If one then goes down to 3nm, both devices show a clear saturation of the current and no further improvement is obtained. In this case, the ratios I_{ds}^{DG}/I_{ds}^{SG} for the three silicon thicknesses lie between 10 and 100 when we consider higher values of V_{gs}.

Let us now show the effective bandgap calibration in the tunnel junction. In Fig. 5.11, we present different curves indicating how the bandgap has to be modified (increased in the case of quantum confinement) depending on the considered structure (SG/DG), the body thickness (3, 5 and 10nm) and the bias conditions (V_{gs} and V_{ds}).

In general, we can state that the effective bandgap tends to increase when V_{gs} does likewise, and that this growth becomes steeper for low V_{ds}. Comparing both structures, this increase is globally more noticeable for the SG device.

Recall (see Eq. 4.10) that the BTBT probabilities of carriers are exponentially dependent on the tunneling barrier width, w. As a result of this exponential dependence, we can state that the single most important factor that mainly determines the amount of tunnel current in a TFET is the minimum tunneling width across the device, w_{min} [139,150]. It becomes then interesting to analyze how w_{min} varies with V_{gs} in the presence of quantum confinement for the different thicknesses in both SG and DG structures. In Fig. 5.12, we present such variation for $V_{ds} = 0.1$ and 1.1V. w_{min} is extracted as the narrowest distance between the simulated band diagrams at a distance of 0.1nm from the gate insulator

[3]The issue of subthreshold swings and threshold voltages will be analyzed in detail in Chapter 6.

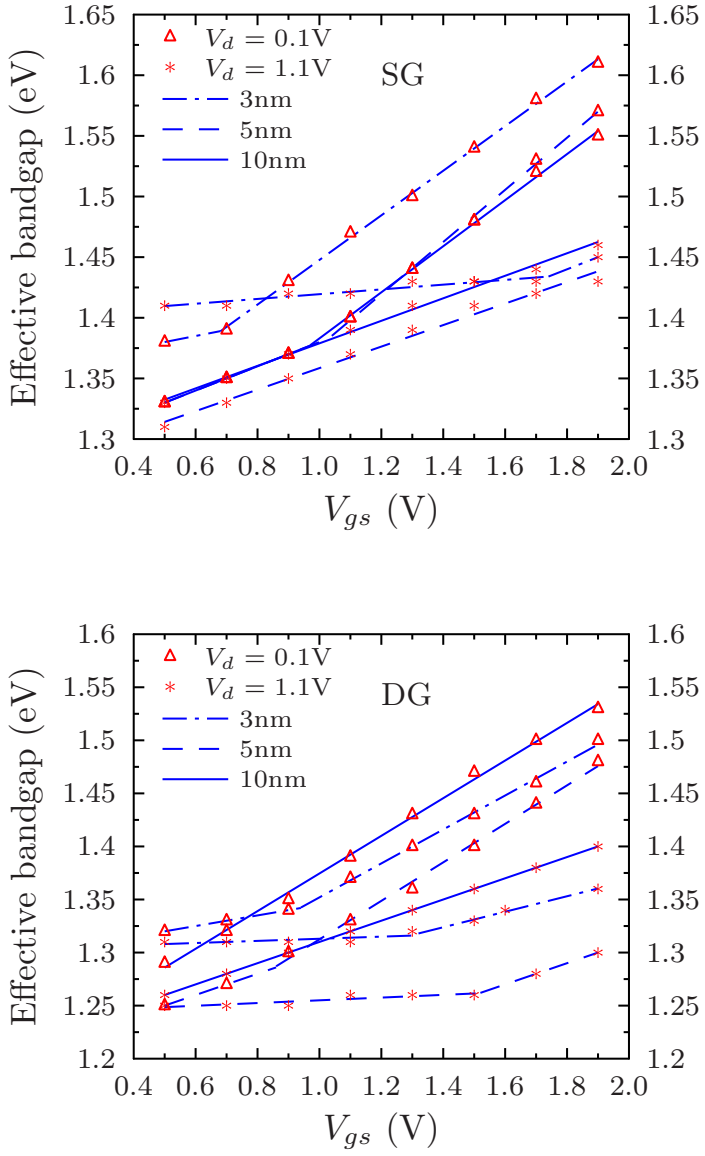

Figure 5.11: Effective bandgap calibration for the SG and DG devices when quantum confinement is included. Dependences on body thickness and bias conditions are shown.

surface. Note that the particular behaviour of the SG–TFET characteristics for low V_{ds} is directly correlated to the w_{min} variation with V_{gs}. Similarly, the saturation of w_{min} at $V_{ds} = 1.1$V when going down from

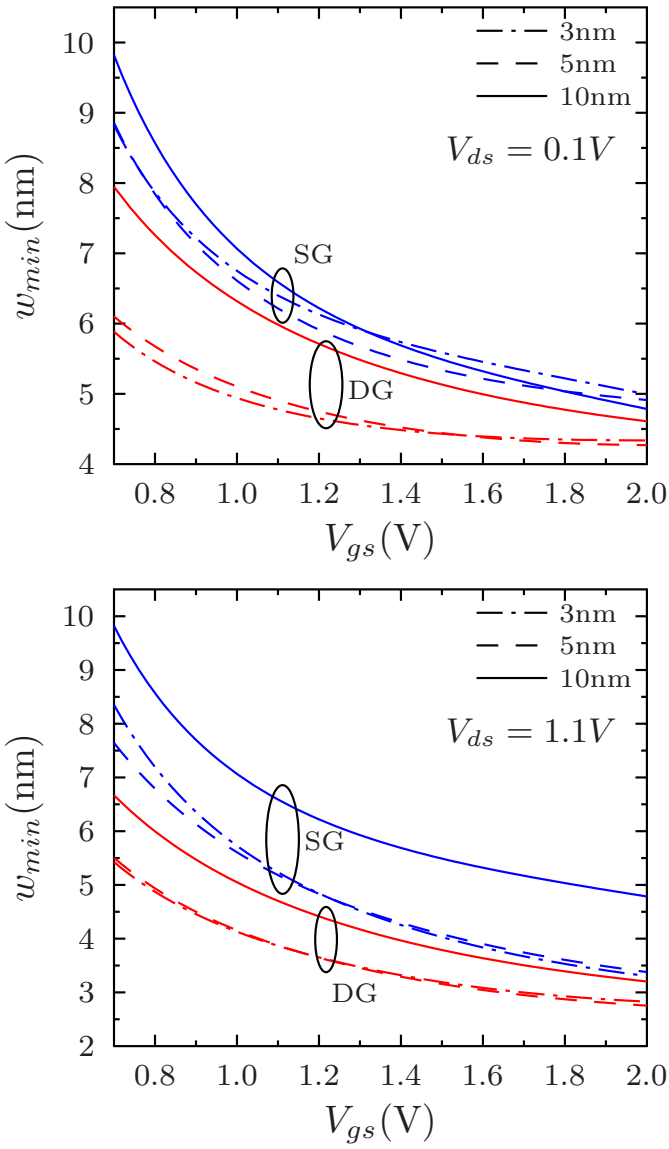

Figure 5.12: w_{min} dependence on V_{gs} when quantum confinement is considered for the SG and DG devices at $V_{ds} = 0.1$ and 1.1V, and for $t_{Si} = 3$, 5 and 10nm.

5nm to 3nm explains the apparent convergence of the $I_{ds} - V_{gs}$ curves (especially in the DG structure).

5.5.2 Effect of high–κ dielectric

Considering the results obtained in [111], it becomes greatly interesting to study the effects of a high–κ gate dielectric and see how the currents are affected by the inclusion of quantum confinement. The selected dielectric constants correspond to SiO_2 ($\varepsilon = 3.9$), Si_3N_4 ($\varepsilon = 7.5$), HfO_2 ($\varepsilon = 21$) and ZrO_2 ($\varepsilon = 29$). In all cases, we take $t_{ox} = 1$nm unless otherwise specified.

5.5.2.1 Semiclassical simulations

In Fig. 5.13, we present the obtained curves for $t_{Si} = 10$nm when varying the gate dielectric constants in the SG and DG configurations. We see that an increase of ε represents a higher current level except for the case $\varepsilon = 29$, where both devices show a slight decrease of I_{ds} for high V_{gs}.

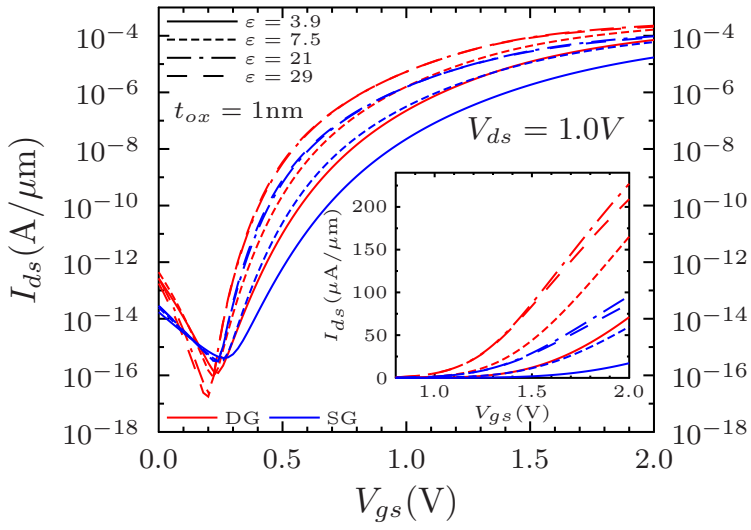

Figure 5.13: Semiclassical transfer characteristics for $t_{Si} = 10$nm with different gate dielectrics for SG and DG structures. V_{ds} =1.0V. The inset shows the current levels in a linear scale.

Similarly, note that the subthreshold swing is reduced when we raise ε.

If we go to a body thickness of $t_{Si} = 5$nm, a quick look to Fig. 5.14 makes us notice that the bunch of curves of SG and DG–TFETs globally tend to split off from each other, yielding a higher current in the case of the DG structure. The previously observed current reduction when increasing ε from 21 to 29 is reinforced in this case. Leakage current is also suppressed as ε gets higher.

When we further reduce the body thickness to 3nm, we obtain Fig. 5.15. In this situation, the two groups of curves are now clearly differentiated in the whole range of V_{gs}. Subthreshold swings are again improved as ε increases, but the total current does not follow this pattern. Fig. 5.15 not only deepens the previously observed current reduction for $\varepsilon = 29$, but also shows that this behaviour is extended to $\varepsilon = 21$ for $t_{Si} = 3$nm.

From these semiclassical simulations, we can therefore infer that for high-κ gate dielectrics and high V_{gs} when the body thickness is reduced from 10 to 3nm, so does the total current. Particularly, for $\varepsilon = 29$ and $V_{gs} = 1.8$V, the current decreases approximately to one third of its

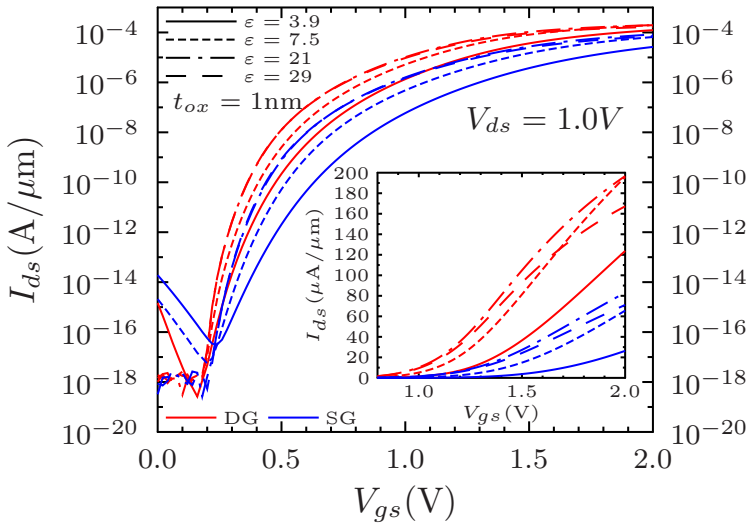

Figure 5.14: Semiclassical transfer characteristics for $t_{Si} = 5$nm and different gate dielectrics for SG and DG–TFETs. V_{ds} =1.0V. The inset shows the current levels in a linear scale.

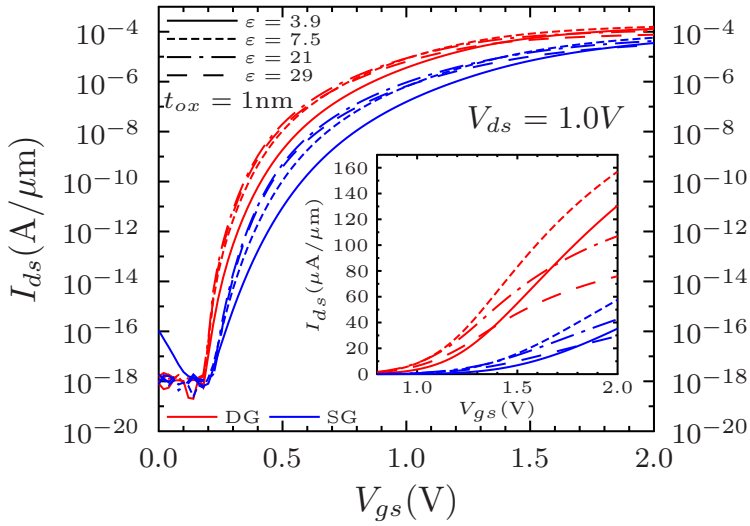

Figure 5.15: Semiclassical transfer characteristics for t_{Si} = 3nm with different gate dielectrics for SG and DG structures. V_{ds} =1.0V. The inset shows the curves in a linear scale.

original value in both SG and DG devices. On the contrary, if a SiO_2 gate dielectric is considered, the current increases when going down to 3nm. All this is depicted in Fig. 5.16, where we show the behaviour of I_{ds} for two fixed values of V_{gs} (V_{gs} = 1.0 and 1.8V) when varying t_{Si} and ε and considering both the SG and DG configurations.

The noteworthy trend of getting lower current levels for high–κ dielectrics at high V_{gs} when we reduce t_{Si} discards the possibility of attributing the difference between the $I_{ds} - V_{gs}$ curves of Figs. 5.13–5.15 to a purely capacitive effect (something similar was already noticed in [111] for DG–TFETs). Fig. 5.17 illustrates, for the DG structure and t_{Si} = 3nm, the transfer characteristics resulting from replacing t_{ox} under the gate by the corresponding equivalent oxide thickness (EOT) for each ε.

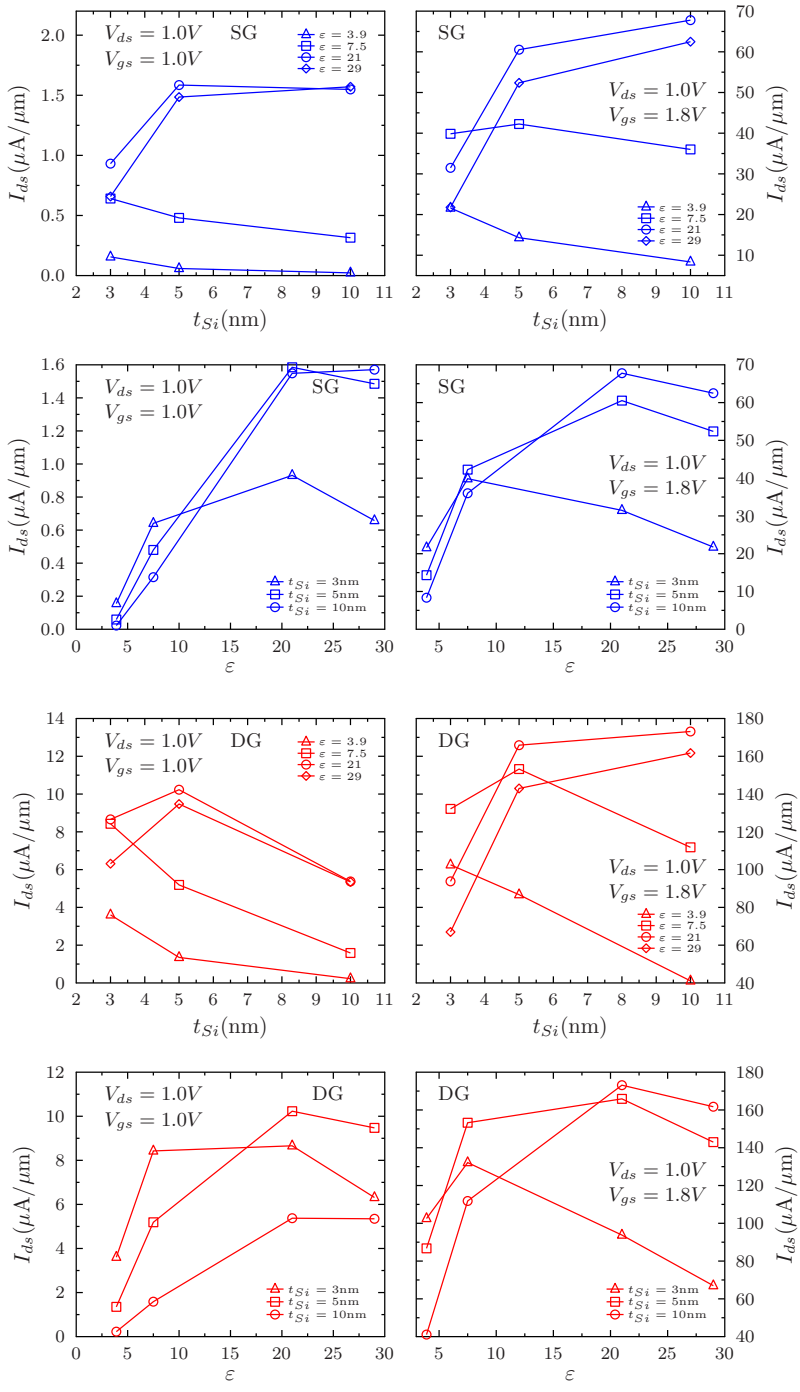

Figure 5.16: Semiclassical I_{ds} dependence on t_{Si} and ε for the SG and DG devices at $V_{ds} = 1.0$V, and for $V_{gs} = 1.0$ and 1.8V.

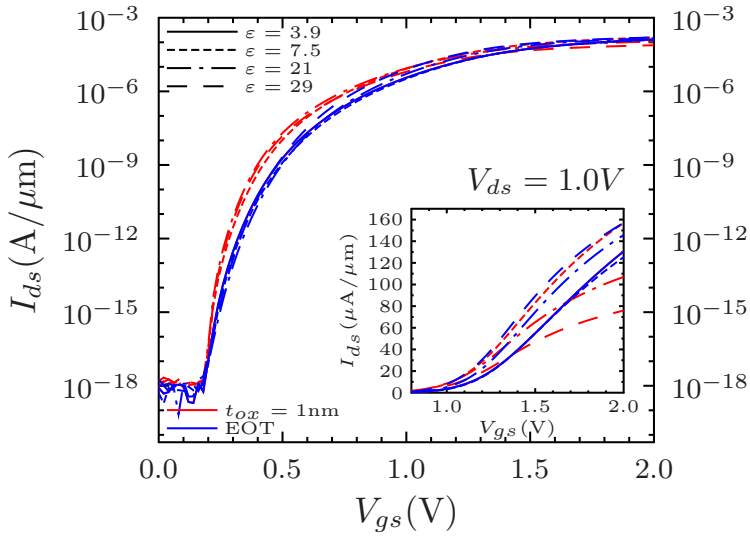

Figure 5.17: Comparison between semiclassical $I_{ds} - V_{gs}$ curves with $t_{ox} =$ 1nm and with equivalent oxide thickness under the gate for the different permittivities at $V_{ds} = 1.0$V. $t_{ox} = 1$nm for $\varepsilon = 3.9$ was taken as reference, and threfore solid lines logically overlap. The results correspond to the DG device with $t_{Si} = 3$nm.

5.5.2.2 Simulations including quantum confinement

Again, when quantum cofinement is accounted for, we expect a BTBT reduction as a consequence of the higher effective bandgap between the first bound states coming from the conduction and valence band quantization. In Fig. 5.18, we present the resulting transfer characteristics for $t_{Si} = 10$nm. Comparing the obtained curves with those of Fig. 5.13, we first confirm this announced current reduction. However, other interesting remarks arise in view of the comparison of these two figures:

- In the semiclassical framework, the characteristics corresponding to SG and DG configurations were to some extent entangled and there was no clear separation between them. As an example, the SG curves with $\varepsilon = 21$ and 29 lay above the DG curve with $\varepsilon = 3.9$. This entanglement tended to disappear when reducing t_{Si}. On the other hand, when conduction and valence band quantization is incorporated, the two bunch of curves are clearly differentiated

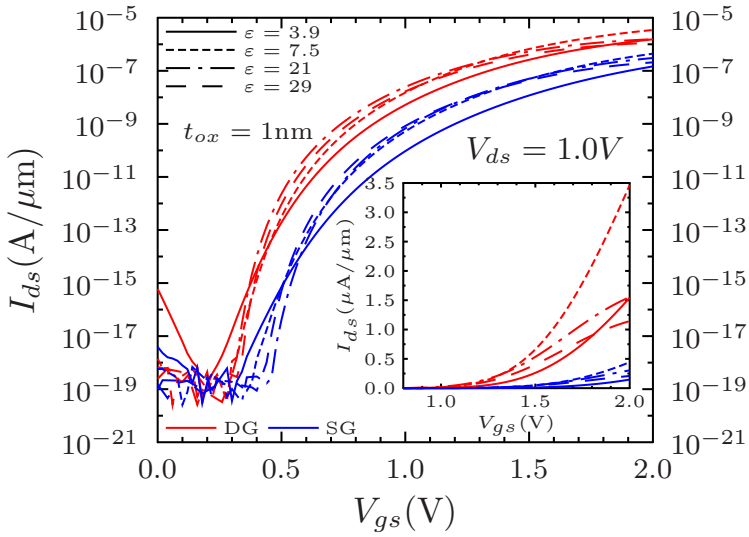

Figure 5.18: Transfer characteristics including subband quantization for $t_{Si} =$ 10nm with different gate dielectrics for SG and DG structures. $V_{ds} = 1.0$V. The inset displays I_{ds} in a linear scale.

even for $t_{Si} = 10$nm.

- Semiclassically, the highest current level was achieved for $\varepsilon = 21$ when $t_{Si} = 10$nm, whereas some reduction was observed if the dielectric constant was further increased to $\varepsilon = 29$. This trend is intensified in the presence of quantum confinement and, in that case, the value of the dielectric constant producing the highest current is shifted to $\varepsilon = 7.5$.

It becomes then greatly interesting to see what happens if we reduce the body thickness. If we proceed in such a way, the results corresponding to $t_{Si} = 3$nm including confinement are shown in Fig. 5.19. Comparing with Fig. 5.15, the separation of the two groups of curves is strengthened almost in the whole range of voltages except for the region of high V_{gs} and due to the behaviour of the DG device. In that region, for the DG structure, the presence of quantization makes the total current decrease for $V_{gs} \gtrsim 1.5$V and $\varepsilon \geq 7.5$. Apparently, a DG–TFET with SiO_2 gate dielectric is unaffected by this current reduction at high V_{gs}.

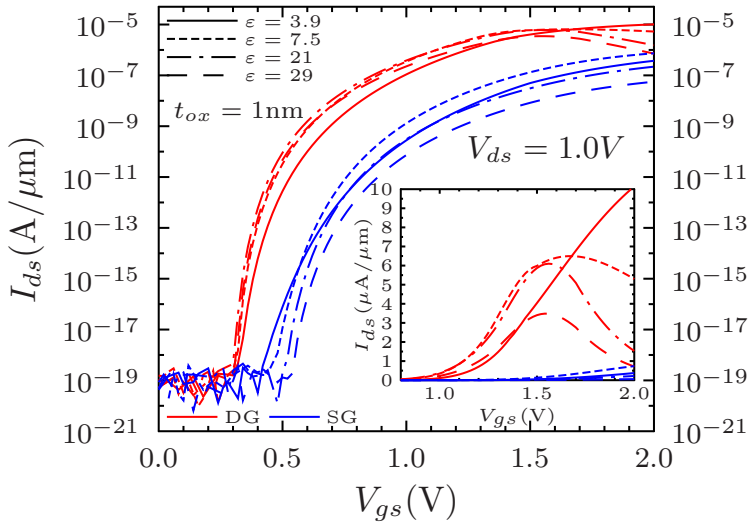

Figure 5.19: Transfer characteristics with quantum confinement for t_{Si} = 3nm with different gate dielectrics for SG and DG structures. V_{ds} = 1.0V. The inset displays I_{ds} in a linear scale.

Nonetheless, the apparent displacement of the current maxima observed in the inset of Fig. 5.19 suggests that for $\varepsilon = 3.9$ this current reduction would appear at gate voltages over 2V.

Once that we have studied how the current varies depending on the considered gate dielectric, it may result interesting —similarly to what was done in [111]— to check the validity of the approximation made in Eq. 4.16 where λ was taken as the uniform tunneling width over the range of energies for which BTBT was possible. In Figs. 5.20 and 5.21, we represent I_{ds} versus $(1/\varepsilon_{ox})^{1/2}$ for t_{Si} = 10nm and study the deviations of our simulations from the linear behaviour represented by the solid (semiclassical) and dashed (quantum) lines. Semiclassical simulations provide a reasonable fit to the suggested approximation, but this agreement worsens as we incorporate quantum confinement.

If we further reduce the body thickness to 3nm, therefore making the effects of quantization stronger, we can see how the results depicted in Fig. 5.22 and especially in Fig. 5.23 clearly bring up the limitations of Eq. 4.16.

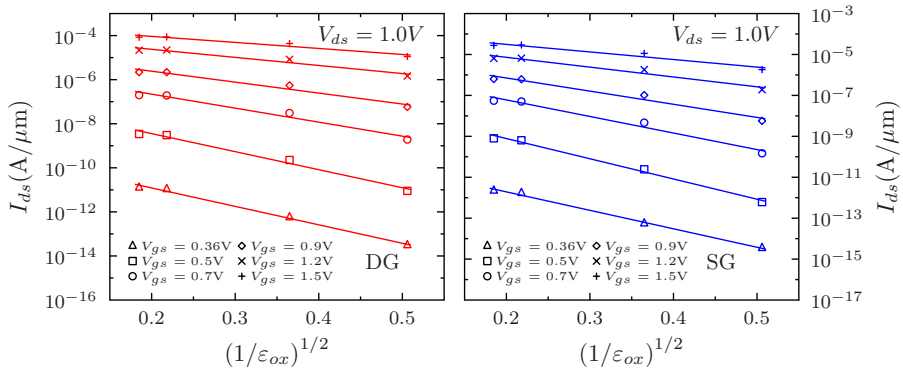

Figure 5.20: Semiclassical I_{ds} vs. $(1/\varepsilon_{ox})^{1/2}$ plots for $t_{Si} = 10$nm corresponding to SG and DG structures. $V_{ds} = 1.0$V.

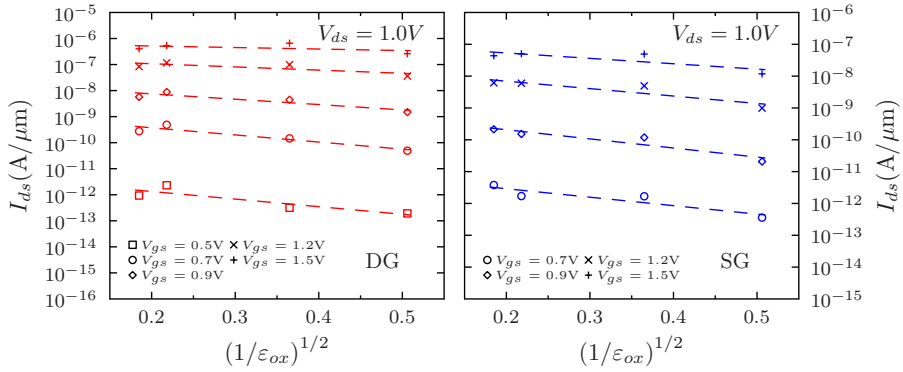

Figure 5.21: I_{ds} vs. $(1/\varepsilon_{ox})^{1/2}$ plots for $t_{Si} = 10$nm corresponding to SG and DG structures when we incorporate the effect of quantum confinement. $V_{ds} = 1.0$V.

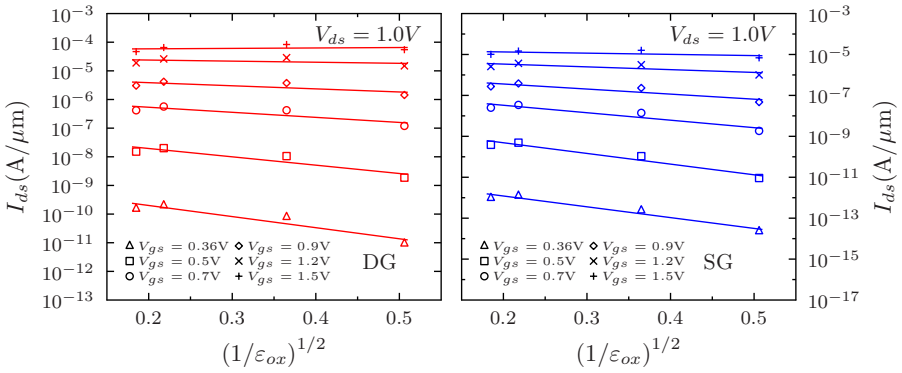

Figure 5.22: Semiclassical I_{ds} vs. $(1/\varepsilon_{ox})^{1/2}$ plots for $t_{Si} = 3$nm corresponding to SG and DG structures. $V_{ds} = 1.0$V.

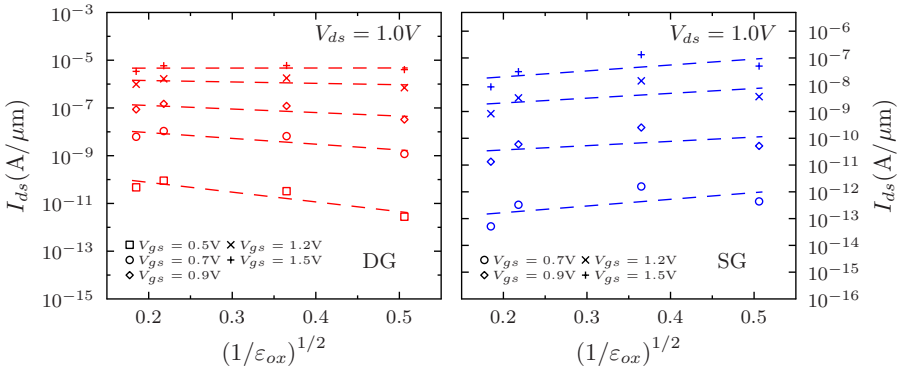

Figure 5.23: I_{ds} vs. $(1/\varepsilon_{ox})^{1/2}$ plots for $t_{Si} = 3$nm corresponding to SG and DG structures when we incorporate the effect of quantum confinement. $V_{ds} = 1.0$V.

5.5.3 Effect of doping variation

We now investigate the effect that a doping variation has over the DG and SG structures with $t_{Si} = 3$nm. In our n–channel TFET devices, the BTBT injection of electrons takes place between the p^+–doped region (which acts as the source) and the lightly n–doped region (the channel). Recall that the doping profiles employed in our simulations are not abrupt at the junction, but rather possess a gaussian drop off profile as stated in Sec. 5.4. This gaussian profile is not varied in this section, the doping modifications only refer to the maximum dopings.

5.5.3.1 Semiclassical simulations

In Fig. 5.24, we show the different transfer characteristics that one obtains when varying the doping level of the source (p–type). It becomes apparent the extraordinary degradation of the curves when we reduce it: when we go down from a 10^{20}cm^{-3} doping to 10^{19}cm^{-3}, the currents are reduced four orders of magnitude for the DG configuration and a

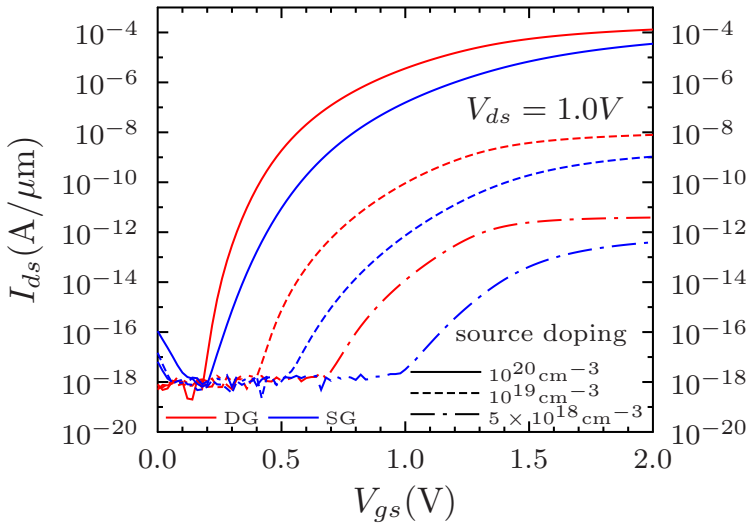

Figure 5.24: Semiclassical characteristics corresponding to SG and DG devices for $t_{Si} = 3$nm with different p–type doping levels at the source. $V_{ds} = 1.0$V.

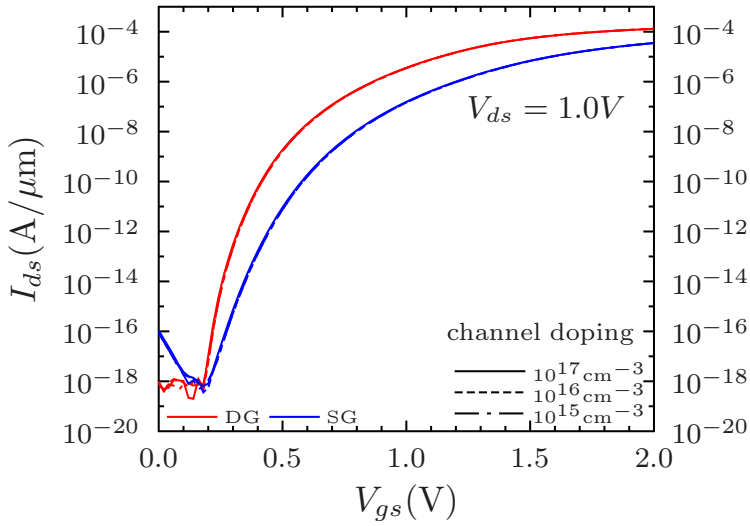

Figure 5.25: SG and DG semiclassical characteristics for $t_{Si} = 3$nm with different n–type doping levels at the channel. $V_{ds} = 1.0$V.

little bit more for the SG one. This enlightens the utmost importance of the source doping and how it is crucial to achieve the highest possible value for it.

On the other hand, it can be seen in Fig. 5.25 that variations in the channel n–type doping (keeping it in the range of intrinsic/lightly doped) do not affect neither the SG–TFET nor the DG–TFET performance.

5.5.3.2 Simulations including quantum confinement

As a result of the inclusion of subband quantization in the conduction and valence bands, we see in Fig. 5.26 that the I_{ds} curves are significantly reduced, which is in good agreement with what one would expect. Notice that in the case of a doping of $5 \times 10^{18}\mathrm{cm}^{-3}$, this reduction is critical for the SG device and no current is obtained beyond the leakage level.

By contrast, if the channel n–type doping is modified when confinement is present no changes arise in the characteristics as depicted in Fig. 5.27.

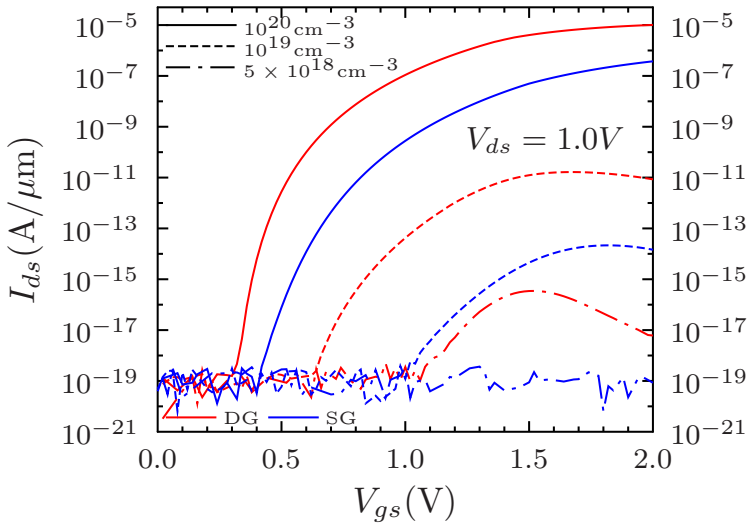

Figure 5.26: I_{ds} curves including quantum confinement for $t_{Si} = 3$nm with different p–type doping levels at the source. $V_{ds} = 1.0$V.

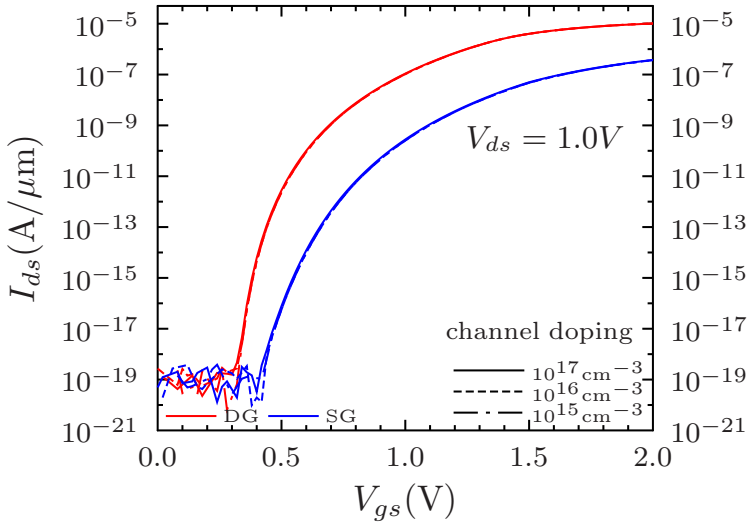

Figure 5.27: Characteristics incorporating confinement for $t_{Si} = 3$nm with different n–type doping levels at the channel. $V_{ds} = 1.0$V.

Chapter 6

Quantum effects on threshold voltages and SS in TFETs

This chapter tackles the study of quantum effects over two electrical parameters of paramount importance in the characterization and performance of TFETs. We make use of the quantum mechanical treatment described in Chapter 5 and shed light on novel trends of these parameters.

The actual configuration of conduction and valence bands as a discrete set of energy levels, necessarily modifies the control that the electrodes exert over total current thus affecting the behaviour of threshold voltages and subthreshold swings. These effects have not been analized in the literature and lead to interesting results for ultrathin body devices.

6.1 Threshold voltages

Unlike the usual constant current technique —which is somehow arbitrary and not based on physical processes taking place in the device—, we choose to deal with this parameter connecting it to nanoscale physics, and use to estimate it the point where the control that the corresponding

electrode has over the total current turns to be linear instead of quasi-exponential. This choice is very similar to that employed in conventional MOS transistors [172] where it refers to the transition between weak and strong inversion. In our case, we will quantify the threshold voltage by inspecting the maximum of the transconductance derivative or, in other words, examining the second derivative of the $I_{ds} - V$ curves. Analogously to MOS transistors, this point marks the separation between different variation regimes of total current when varying the supplied voltage. In these devices, as the carrier injection mechanism is dominated by BTBT, which depends on the tunneling width, w (recall Eq. 4.10), the threshold voltages will refer to the change of the control that the electrodes exert over it. This dependence on w being exponential allows to simplify the analysis, and focus on the study of w_{min} variations (as already discussed in Chapter 5).

We have avoided so far to openly talk about the gate when mentioning the electrodes because TFETs possess a unique feature that distinguishes them from conventional MOSFETs: they show two different threshold voltages, the gate threshold voltage, V_{tg}, and the drain threshold voltage, V_{td} [139]. The existence of these two voltages is related to how the gate and the drain affect w_{min}, which is a complex function of both V_{gs} and V_{ds}. In general —as will be shown in what follows—, when we increase one of those voltages, the other retains for a wider voltage range a stronger control over w_{min} (and subsequently over total current), which raises the corresponding threshold voltage of the latter.

6.1.1 The role of the inversion layer

6.1.1.1 Semiclassical approach

The appearance of an inversion layer in TFETs when we raise V_{gs} is an issue of major interest as previously discussed in Sec. 4.3.4. It is formed by carriers thermally injected from the drain [163] and causes the pinning of the surface channel potential, ϕ_s. This means that band bending, and indirectly w_{min}, is less sensitive to V_{gs} variations. Fig. 6.1 shows the variation of ϕ_s with V_{gs} for different values of V_{ds} corresponding

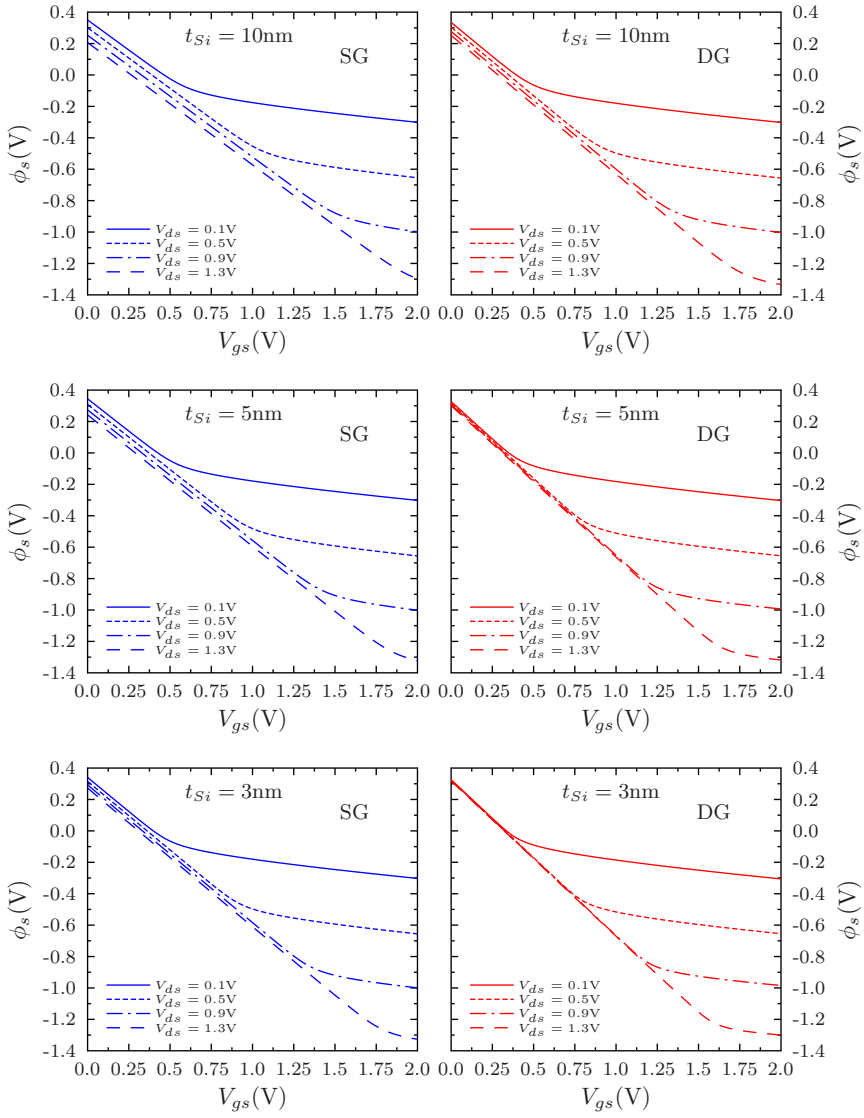

Figure 6.1: ϕ_s variation with V_{gs} for the SG and DG structures at different values of V_{ds}. $t_{Si} = 3$, 5 and 10nm.

to semiclassical simulations of the SG and DG structures depicted in Fig. 5.2. Notice how the appearance of the inversion layer modifies the ϕ_s linear dependence on V_{gs}.

According to Eq. 4.17, before the inversion layer is formed (i.e. $V_{gs} < V_{inv}$), the correspondence between V_{gs} and ϕ_s should be nearly one–to–one provided that $C_q << C_{ox}$. As shown in Fig. 6.2, this assumption turns out to be more accurate for the DG configuration rather than for the SG one; and —regarding each structure separately— when we reduce t_{Si}. Similarly, one would expect that the pinning of ϕ_s should result in a nearly zero slope of the $\phi_s - V_{gs}$ curves when $V_{gs} > V_{inv}$. Such a behaviour is again better achieved as we reduce the body thickness and for the DG device.

With the inversion layer formed (i.e. $V_{gs} > V_{inv}$ and, subsequently, $C_q >> C_{ox}$) and low V_{ds}, the accumulation of carriers reduces the channel resistance thus making the voltage applied at the drain drop at the tunneling junction. In this situation, it is the drain and not the gate which exerts the main control over ϕ_s and, by extension, over w_{min}. Fig. 6.3 illustrates for the DG structure and $t_{Si} = 3$nm how V_{ds} makes ϕ_s strongly vary provided that (recall Eq. 4.18)

$$V_{ds} < V_{gs} - V_{inv}|_{V_{ds}=0}\,. \tag{6.1}$$

V_{inv} values have been calculated through the maximum of the second derivative of the $\phi_s - V_{gs}$ curves. Consider now the point where the V_{gs} and V_{inv} lines intersect, which corresponds to $V_{ds,sat} = 0.82$V. On the left side, $V_{gs} > V_{inv}$ and therefore ϕ_s is controlled by V_{ds} (see how, in Fig. 6.4, the value of the slope is close to -1 in that region). Conversely, on the right side, as $V_{gs} < V_{inv}$, inversion layer has disappeared, the carriers have been pulled out of the channel, and ϕ_s is no longer affected by V_{ds} (the slope almost reaches zero).

Let us now illustrate with a particular example the influence of the inversion layer over w_{min}. Fig. 6.5 shows for the DG structure with $t_{Si} = 3$nm the variation of w_{min} with V_{gs} at $V_{ds} = 0.7$ and 1.4V, along with the corresponding $\phi_s - V_{gs}$ curve for $V_{ds} = 0.7$V. Note how the

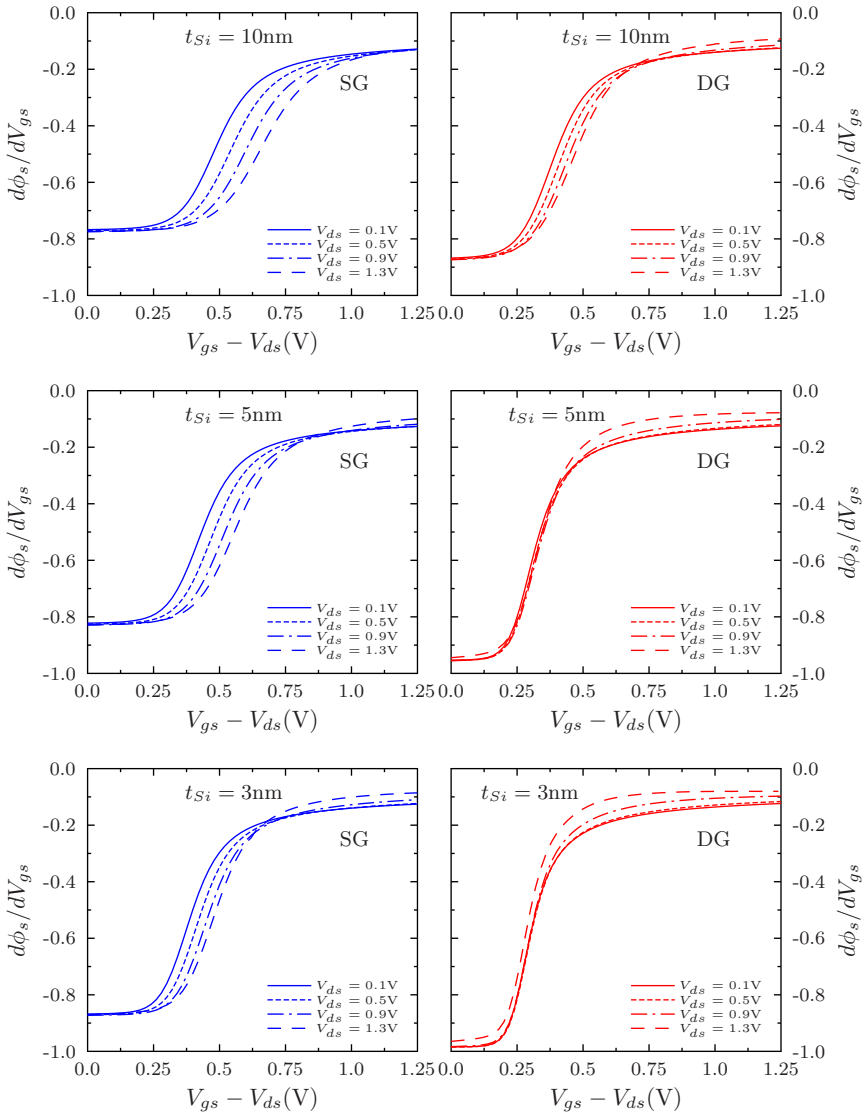

Figure 6.2: $d\phi_s/dV_{gs}$ vs. $V_{gs} - V_{ds}$ for the SG and DG structures at different values of V_{ds}. $t_{Si} = 3$, 5 and 10nm.

appearance of the inversion layer at $V_{gs} = 0.98$V for $V_{ds} = 0.7$V makes w_{min} decrease more slowly from that point onwards compared to the case where $V_{ds} = 1.4$V.

From the above, it is then reasonable to assume that changes in the

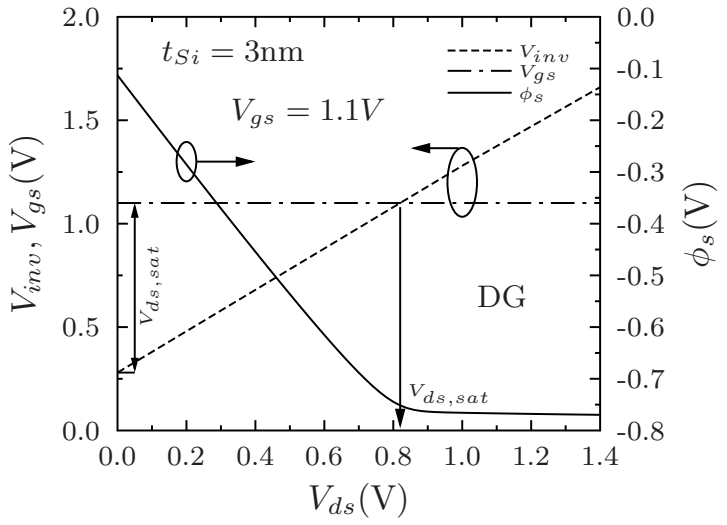

Figure 6.3: ϕ_s variation with V_{ds} for the DG device with $t_{Si} = 3$nm at $V_{gs} = 1.1$V, along with the gate voltages at which inversion is formed, V_{inv}.

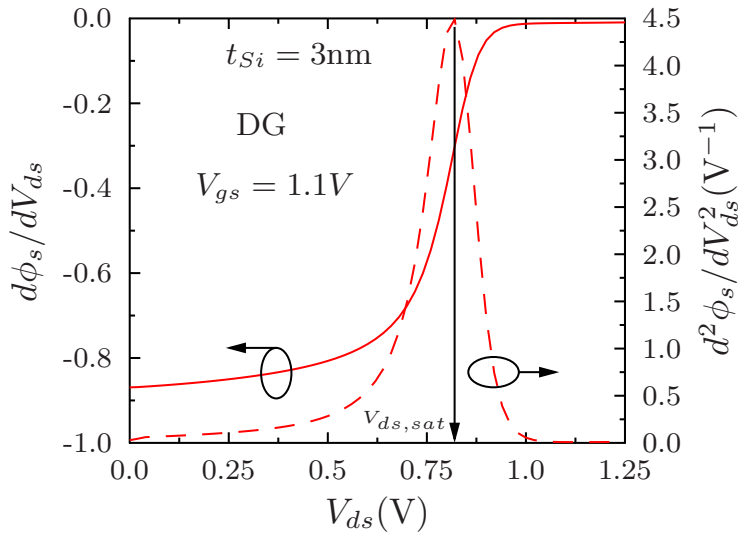

Figure 6.4: $d\phi_s/dV_{ds}$ and $d^2\phi_s/dV_{ds}^2$ for the DG device with $t_{Si} = 3$nm at $V_{gs} = 1.1$V. The maximum of the second derivative coincides with $V_{ds,sat}$.

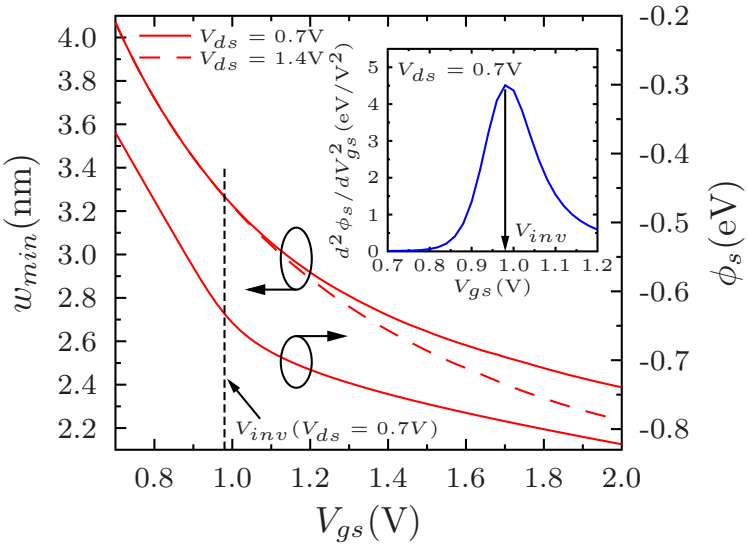

Figure 6.5: w_{min} variation with V_{gs} for the DG device with t_{Si} = 3nm at V_{ds} = 0.7 and 1.4V. The curve $\phi_s - V_{gs}$ for V_{ds} = 0.7V and its modified reduction rate enlightens the effect of the inversion layer formation at V_{inv} = 0.98V. The inset shows the second derivative of ϕ_s which provides the value of V_{inv}.

behaviour of BTBT current are correlated with the formation of the inversion layer and with the interplay between both the gate and the drain.

6.1.1.2 Quantum mechanical approach including confinement

The addition of quantum confinement to this scenario makes it even more interesting: as long as the conduction and valence bands become a discrete set of energy levels, the effective bandgap increases and this directly alters the value of w_{min} which turns to be bigger than that estimated from semiclassical simulations. In Fig. 6.6, we present the analogous curves to those of Fig. 6.1 but now assuming that the role of the surface channel potential is played by the first bound state of the conduction band, which we represent as $\tilde{\phi}_s$.

Compared to the evolution of ϕ_s in Fig. 6.1, we notice that once the inversion layer is formed, $\tilde{\phi}_s$ becomes more strongly pinned, as illustrated

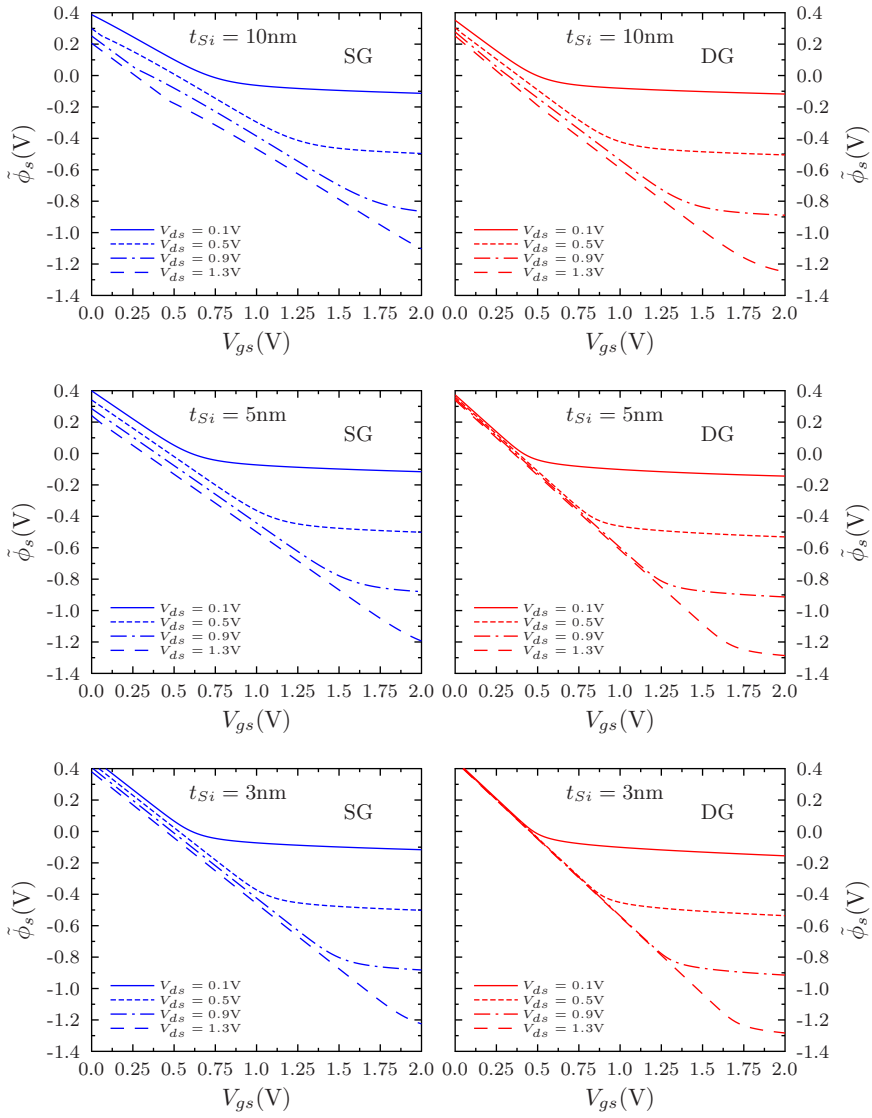

Figure 6.6: $\tilde{\phi}_s$ variation with V_{gs} for the SG and DG structures at different values of V_{ds}. t_{Si} = 3, 5 and 10nm.

in Fig. 6.7 by the fact that the slope of the $\tilde{\phi}_s - V_{gs}$ curves globally saturates to a value closer to zero in both SG and DG devices. Two additional remarks may be done in light of Figs. 6.2 and 6.7:

(i) When $V_{gs} < V_{inv}$, we observe that for the SG structure $\tilde{\phi}_s$ is not

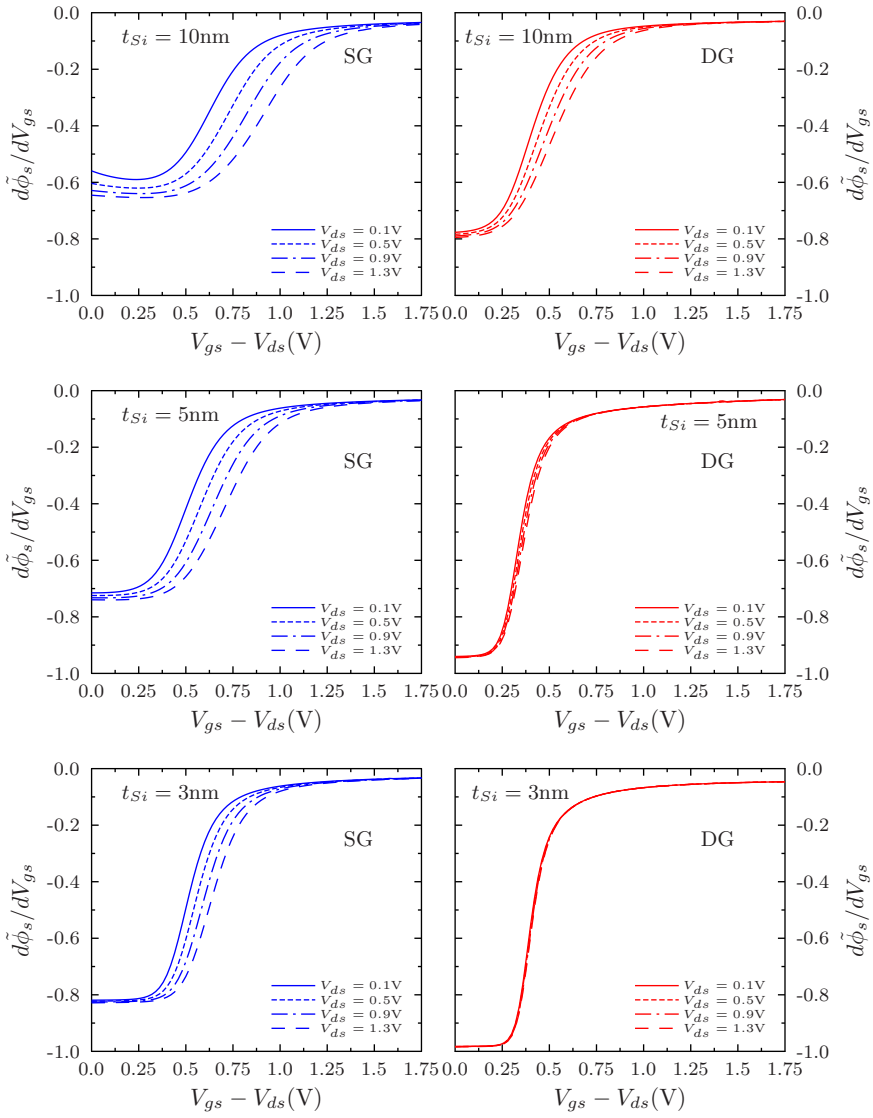

Figure 6.7: $d\tilde{\phi}_s/dV_{gs}$ vs. $V_{gs} - V_{ds}$ for the SG and DG structures at different values of V_{ds}. $t_{Si} = 3$, 5 and 10nm.

as tightly controlled by V_{gs} as ϕ_s was. This divergence between semiclassical and quantum approaches is especially noticeable for $t_{Si} = 10$nm and, particularly, if one regards the SG structure. Furthermore, the case of the SG structure with $t_{Si} = 10$nm shows

an unexpected behaviour of the $\tilde{\phi}_s - V_{gs}$ curve, which proves to have a non–uniform decreasing slope for low V_{gs} values.

(ii) Concerning the behaviour of the ϕ_s and $\tilde{\phi}_s$ pinning when we vary V_{ds}, we see that —whereas in semiclassical simulations ϕ_s is more tightly pinned as V_{ds} increases—, the pinning of $\tilde{\phi}_s$ turns out to be somehow independent of the drain bias when we incorporate quantum confinement.

6.1.2 Gate threshold voltage, V_{tg}

As previously mentioned, the gate threshold voltage of TFETs keeps a certain resemblance with that of conventional MOSFETs in the sense that it is also related to the gate and to how it modifies the device performance by switching from one regime of current increase to another. Yet, the physical origin of this switching is not the same in both devices as long as they are based in different carrier injection mechanisms.

6.1.2.1 DG–TFET

We first deal with the behaviour of V_{tg} for the DG structure. In Figs. 6.8, 6.9 and 6.10, we show for both semiclassical and quantum approaches the variation of V_{tg} with V_{ds} corresponding to $t_{Si} = 3$, 5 and 10nm, respectively. In those figures, we also depict V_{inv} and V_{on} (the gate voltage at which BTBT starts). V_{inv} has been calculated as the maximum of the second derivative of ϕ_s (see, for example, the inset of Fig. 6.5). V_{on} is extracted by measuring V_{gs} when I_{ds} becomes higher than the leakage current [163].

In all cases, we observe that $V_{on} < V_{inv}$ in the whole range of V_{ds}. This means that in the DG device BTBT always starts before the formation of the inversion layer. These results reinforce that the inversion layer is not necessary to trigger BTBT —as stated in [151]—, unlike that originally claimed in the early states of TFETs study [164].

Concerning the behaviour of V_{tg}, several conclusions may be extracted in light of Figs. 6.8, 6.9 and 6.10:

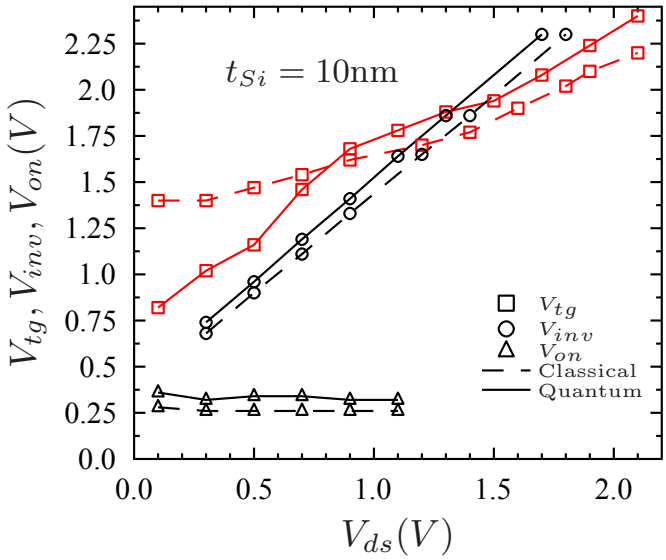

Figure 6.8: V_{tg} dependence on V_{ds} for the DG configuration with $t_{Si} = 10$nm. Gate voltages corresponding to V_{inv} and V_{on} have also been depicted.

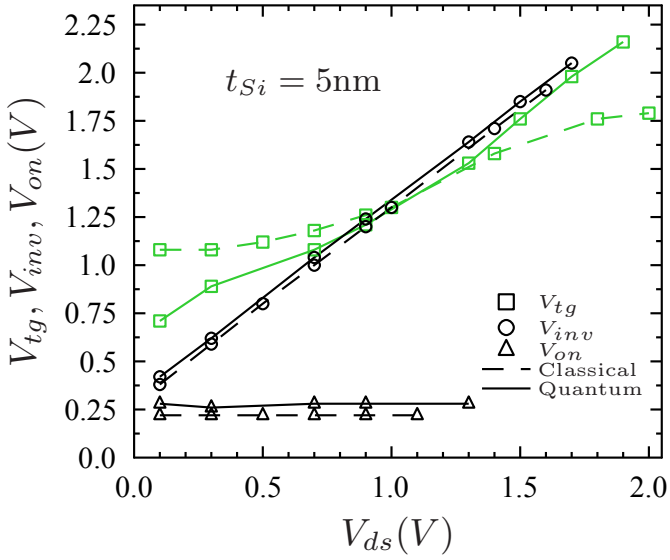

Figure 6.9: V_{tg} dependence on V_{ds} for the DG configuration with $t_{Si} = 5$nm along with gate voltages corresponding to V_{inv} and V_{on}.

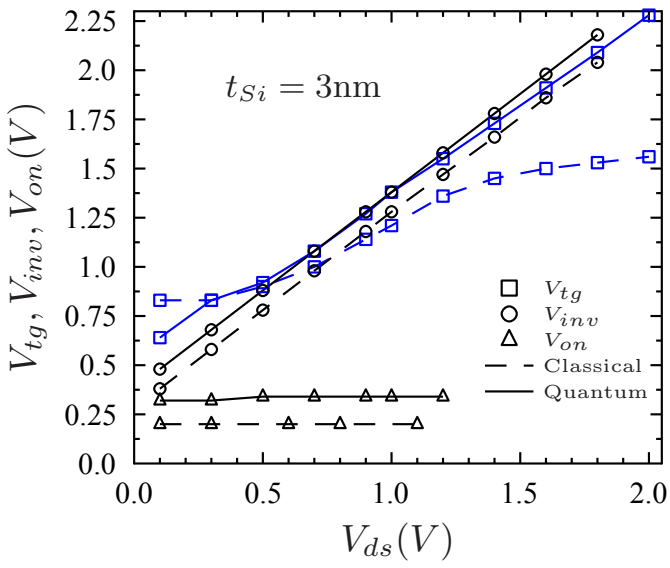

Figure 6.10: V_{tg} dependence on V_{ds} for the DG configuration with $t_{Si} = 3$nm. Gate voltages corresponding to V_{inv} and V_{on} are also included.

(i) For the different thicknesses considered in our simulations, semiclassical and quantum curves show an upward trend of V_{tg} with V_{ds}, hence indicating that an increase in V_{ds} allows the gate to keep a quasi-exponential control over the current for a wider range of voltages.

(ii) Semiclassical simulations show that for ultrathin Silicon bodies the increase of V_{tg} tends to saturate for high values of V_{ds} and that this behaviour is more acute as t_{Si} gets thinner. This saturation of V_{tg} in the absence of quantum corrections was originally suggested in [139]. However, inclusion of quantum confinement drastically changes this trend and the almost linear growth of V_{tg} with V_{ds} is preserved within the studied extent of voltages, even for ultrathin devices.

(iii) Regardless of the growing pattern of V_{tg} with V_{ds}, it can be observed that semiclassically V_{tg} clearly increases with t_{Si} in the whole range of V_{ds}; whereas in the quantum framework V_{tg} curves

appear to be somehow independent of the body thickness.

(iv) When quantum effects are considered, the gate voltage at which the inversion layer is formed, V_{inv}, proves to be a very good estimation of V_{tg} for ultrathin structures.

Let us now analyze the reason why V_{tg} curves do not saturate when quantum corrections are included. In this case, from the results presented in Figs. 6.8, 6.9 and 6.10, it is apparent —especially for $t_{Si} = 3$ and 5nm— that V_{inv} is responsible for the value of V_{tg}. The key to understand why this does not happen classically can be found studying the way in which the inversion layer affects the reduction of the tunneling width for carriers in both approaches. At this point, we take up again what we mentioned at the end of the introduction of Chapter 5 and make use of local BTBT models in order to deal with explicit analytical expressions that can be easily manipulated. Of course, the results shown herein using local models will not match those obtained in our simulations, as we employed non-local BTBT. However, they prove to be very useful if what we pursue is only an explanation to the aforementioned behaviour of V_{tg} including quantum confinement.

We consider the widely used Kane model and recall Eq. 5.4

$$G_{Kane} = \frac{DA_{Kane}}{\sqrt{E_g}} F^{\gamma} \exp\left(-B_{Kane}\frac{E_g^{3/2}}{F}\right), \tag{6.2}$$

with the letters standing for the magnitudes as described in Sec. 5.1.1.1. In local models, the electric field may be described [150] as

$$F = \frac{E_g}{qw}, \tag{6.3}$$

making G_{Kane} read as

$$G_{Kane} = \frac{DA_{Kane}}{\sqrt{E_g}} \left(\frac{E_g}{qw}\right)^{\gamma} \exp\left(-B_{Kane}E_g^{1/2}qw\right). \tag{6.4}$$

Again, due to the presence of the exponential factor, we can make the

approximation of dealing with w_{min} which leads to the following expression if one substitutes Eq. 6.4 into Eq. 5.1

$$\begin{aligned} I_{semiclass} &\approx qG_{Kane}\left(w_{min}\right)\left(\Delta x\Delta y\Delta z\right)_{min} \\ &\approx q\frac{DA_{Kane}}{\sqrt{E_g}}\left(\frac{E_g}{qw_{min}}\right)^{\gamma} \\ &\quad \exp\left(-B_{Kane}E_g^{1/2}qw_{min}\right)\left(\Delta x\Delta y\Delta z\right)_{min}. \end{aligned} \tag{6.5}$$

We now study the behaviour of the function $G_{Kane}(w_{min})$. To do so, we consider the device with $t_{Si} = 3$nm and extract the w_{min} dependence on V_{gs} for two values of V_{ds} that showed relevant differences between the semiclassical and quantum approaches in Fig. 6.10. In our case, we choose $V_{ds} = 1.4$V and 2.0V and represent the obtained curves in Fig. 6.11. w_{min} was extracted taking the narrowest distance from the simulated band diagrams at a distance of 0.1nm from the gate dielectric surface.

It can be seen how —when quantum effects are accounted for— the formation of the inversion layer affects the reduction rate of w_{min} with V_{gs}. However, this fact is not reproduced in semiclassical simulations where apparently the inversion layer has almost no effect over the variation of w_{min}. This absence of influence over w_{min} for the semiclassical simulations has clearly nothing to do with the nature of the semiclassical approach itself (recall that in Fig. 6.5 we indeed observed that influence), but rather seems to be motivated in this case by the considerably smaller value of w_{min} for the semiclassical simulations compared to that obtained including quantum confinement. A possible conclusion may be then that a very small value of w_{min} makes its variation less sensitive to the accumulation of carriers in the channel.

An appropriate fit of w_{min} curves in Fig. 6.11 leads to the following parametrization

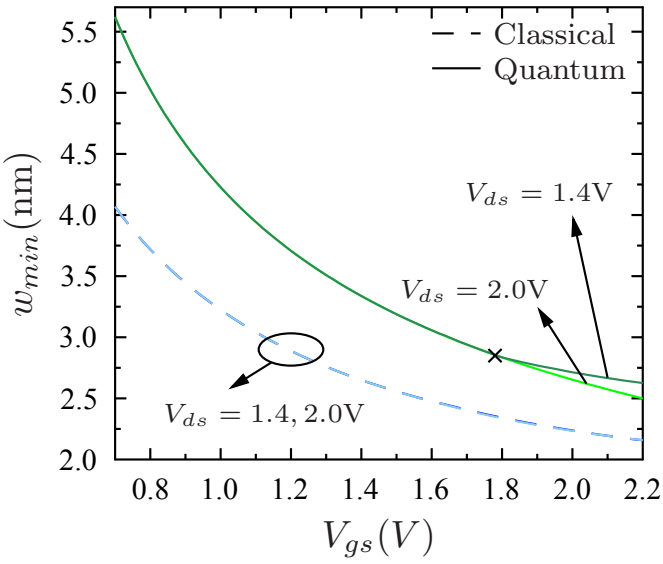

Figure 6.11: w_{min} dependence on V_{gs} for $V_{ds} = 1.4$ and 2.0V, t_{Si} =3nm. Dashed lines stand for semiclassical simulations, solid ones incorporate quantum confinement. Semiclassical simulations practically overlap in the whole range of V_{gs} whereas quantum curves divert at $V_{gs} \approx 1.78$V which is the value of V_{inv} at $V_{ds} = 1.4$V marked by the cross.

$$w_{min}^{quant}\,(V_{ds} = 1.4V) = 2.7188V_{gs}^{-1.1691} + 1.5005, \tag{6.6}$$

$$w_{min}^{quant}\,(V_{ds} = 2.0V) = 3.1377V_{gs}^{-1.0097} + 1.0978, \tag{6.7}$$

$$w_{min}^{class}\,(V_{ds} = 1.4V) = 1.925V_{gs}^{-1.0345} + 1.2943, \tag{6.8}$$

$$w_{min}^{class}\,(V_{ds} = 2.0V) = 1.9145V_{gs}^{-1.039} + 1.3046. \tag{6.9}$$

We apply these expressions to G_{Kane} in Eq. 6.4, and obtain the curves depicted in Fig. 6.12. The effect of the inversion layer becomes evident when one includes quantum corrections. However, as expected

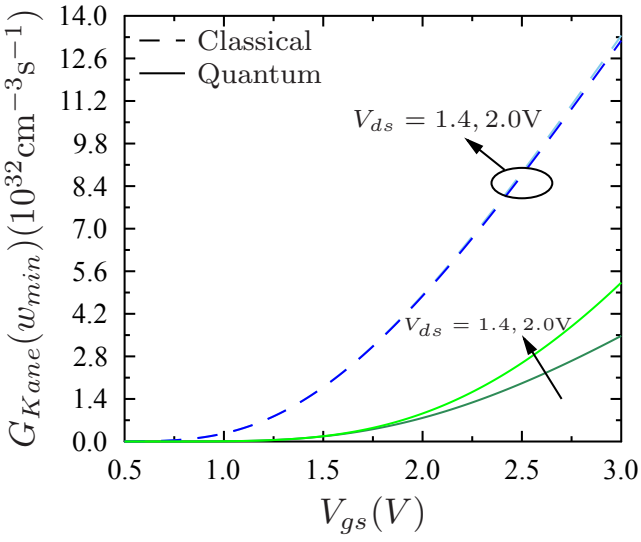

Figure 6.12: $G_{Kane}(w_{min})$ vs. V_{gs} for $V_{ds} = 1.4$ and 2.0V, t_{Si} =3nm. Dashed lines practically overlap and represent semiclassical simulations, solid ones include quantum effects. Quantum curves split up at $V_{gs} \approx 1.78$V which coincides with the value of V_{inv} at $V_{ds} = 1.4$V.

from Fig. 6.11, almost no differences are found semiclassically when we increase V_{ds} from 1.4 to 2.0V.

The next step is to calculate their second derivatives and use the value of the maxima to estimate V_{tg} for the different cases. This has been done in Fig. 6.13. Semiclassical curves show their maxima almost coincident at $V_{tg} \approx 1.3$V. On the other hand, the maxima of the quantum curves split showing values of $V_{tg} \approx 1.8$V and 2.3V for $V_{ds} = 1.4$V and 2.0V, respectively. These extracted values are in very good agreement with those presented in Fig. 6.10.

Hence, the employment of an analytical local BTBT model has allowed us to obtain a very accurate estimation of the V_{tg} values previously obtained from non–local simulations. This fact has helped us to provide a reasonable explanation of the observed differences between semiclassical and quantum simulations shown in Fig. 6.10. We have focused on the device with t_{Si} =3nm because the effects of quantum confinement were more apparent for that thickness.

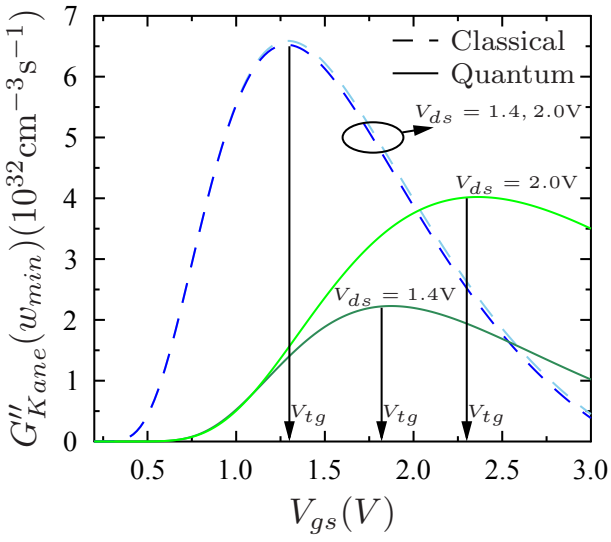

Figure 6.13: $d^2G_{Kane}(w_{min})/dV_{gs}^2$ vs. V_{gs} for V_{ds} =1.4 and 2.0V, t_{Si} = 3nm. Dashed lines lie very close and represent semiclassical simulations, solid ones incorporate quantum effects. Due to the overlap of semiclassical curves, V_{tg} remains almost constant for them, whereas the splitting of quantum curves result in different values of V_{tg}.

6.1.2.2 SG–TFET

An analogous treatment may be repeated to study the V_{tg} dependence on V_{ds} for the SG structure. In this case, the results have been depicted in Figs. 6.14, 6.15 and 6.16. We observe that, unlike what happens with the DG configuration, semiclassical simulations do not present a significant saturating pattern of V_{tg} for high V_{gs} (except for the curve corresponding to t_{Si} = 3nm where a first sign of saturation seems to appear). As a result of this apparent lack of saturation in the considered range of V_{gs}, both semiclassical and quantum corrected curves exhibit nearly linear growths. However, the $V_{tg} - V_{ds}$ curves including quantum confinement still show higher slopes.

Appart from the issue of saturation, some other remarks arise from the comparison between SG and DG devices:

(i) In both structures, the inclusion of confinement causes V_{on} to rise compared to the semiclassical $V_{on} - V_{ds}$ curves (which is logical

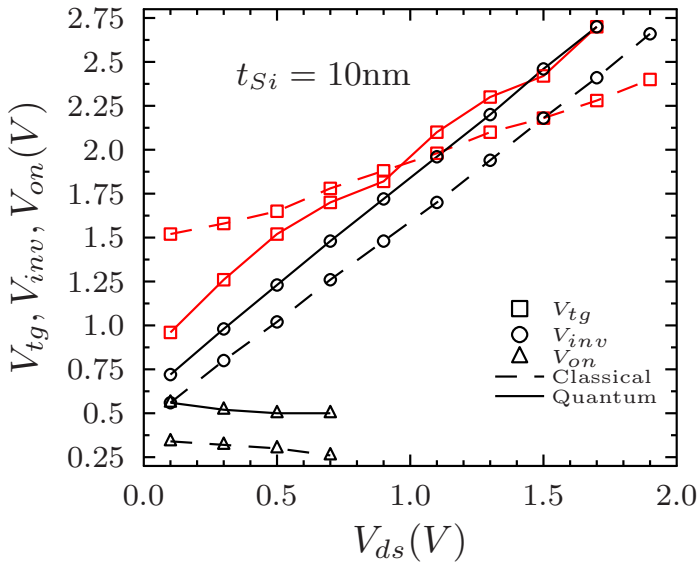

Figure 6.14: V_{tg} dependence on V_{ds} for the SG configuration with $t_{Si} = 10$nm. We also include V_{inv} and V_{on} values.

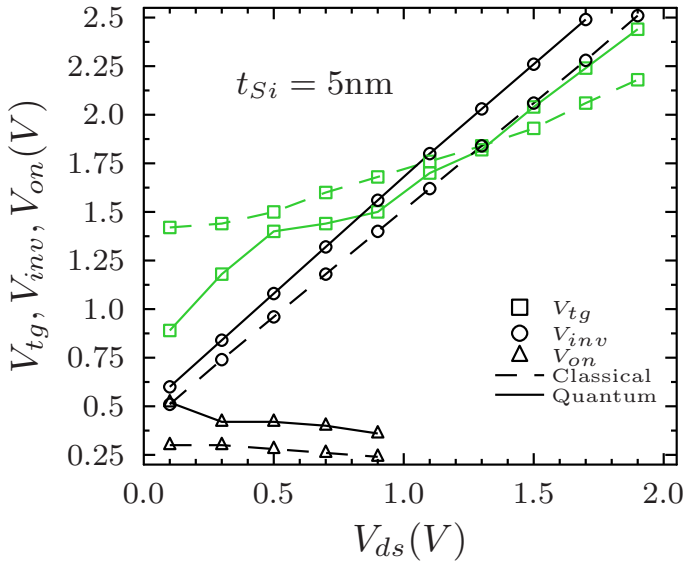

Figure 6.15: V_{tg} dependence on V_{ds} for the SG configuration with $t_{Si} = 5$nm along with gate voltages corresponding to V_{inv} and V_{on}.

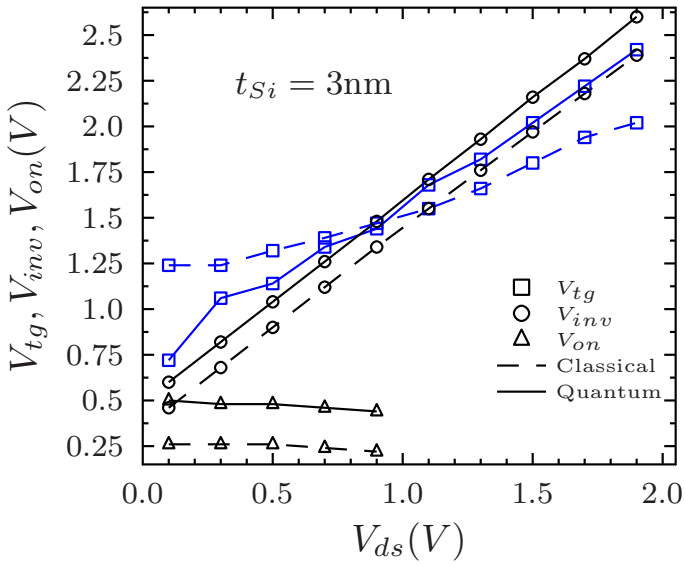

Figure 6.16: V_{tg} dependence on V_{ds} for the SG configuration with $t_{Si} = 3$nm. Gate voltages corresponding to V_{inv} and V_{on} are also included.

if we consider that the effective bandgap is bigger in that case). However, this increase turns out to be more apparent for the SG device than for the DG one. Furthermore, the employment of one gate instead of two seems to make V_{on} slightly decrease with V_{ds}.

(ii) When we move from a SG to a DG configuration, it can be stated that the $V_{tg} - V_{ds}$ curves globally decrease. This happens semiclassically as well as with quantum confinement included.

(iii) The tight correlation between V_{inv} and quantum corrected V_{tg} values so far observed for high V_{gs} in ultrathin DG–TFETs, now deteriorates to a certain extent if we consider a SG structure.

6.1.3 Drain threshold voltage, V_{td}

By changing the role of the electrodes, it becomes possible to extend the threshold voltage definition stated at the beginning of Sec. 6.1 to the drain. If we do so, we can define the so–called drain threshold voltage, V_{td}, and examine its dependence on V_{gs}. In this section, we will only

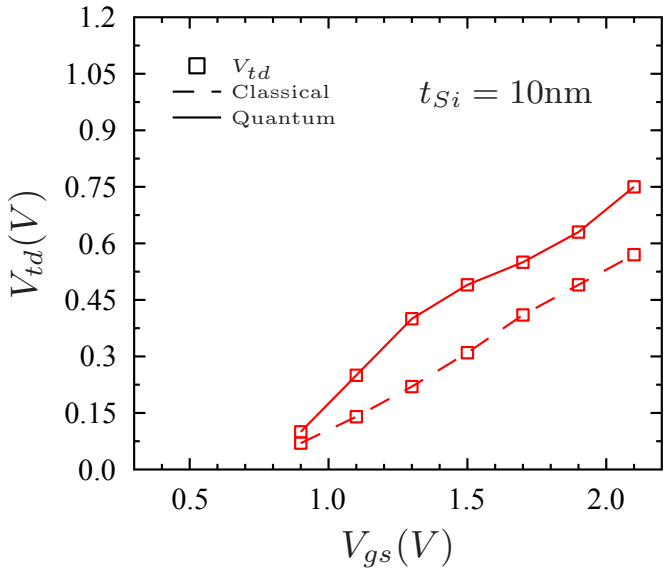

Figure 6.17: V_{td} dependence on V_{gs} for the DG structure and $t_{Si} = 10$nm. Semiclassical saturation of V_{td} is softened for this thickness.

present the results obtained for the DG–TFET and for the previously considered body thicknesses of 3, 5 and 10nm. Figs. 6.17, 6.18 and 6.19, show the $V_{td} - V_{gs}$ curves obtained from our simulations.

Similarly to what happened with V_{tg}, we observe how in the considered range of V_{gs} in Figs. 6.17, 6.18 and 6.19, when we include quantum effects, the lack of saturation of V_{td} becomes manifest —compared to the semiclassical framework— especially for $t_{Si} = 3$ and 5nm. This diverting behaviour between semiclassical and quantum curves is clearly less accentuated for a body thickness of 10nm. Observe that for $t_{Si} = 10$nm, semiclassical V_{td} values present a similar behaviour to those corresponding to semiclassical V_{tg} values (of course, assuming a change of role of the electrodes) in the sense that their curves do not show saturation in the range of values of interest corresponding to high V_{gs}. In addition, conversely to what happened with $V_{tg} - V_{ds}$ curves in the DG structure, semiclassical $V_{td} - V_{gs}$ curves do not decrease when reducing t_{Si}.

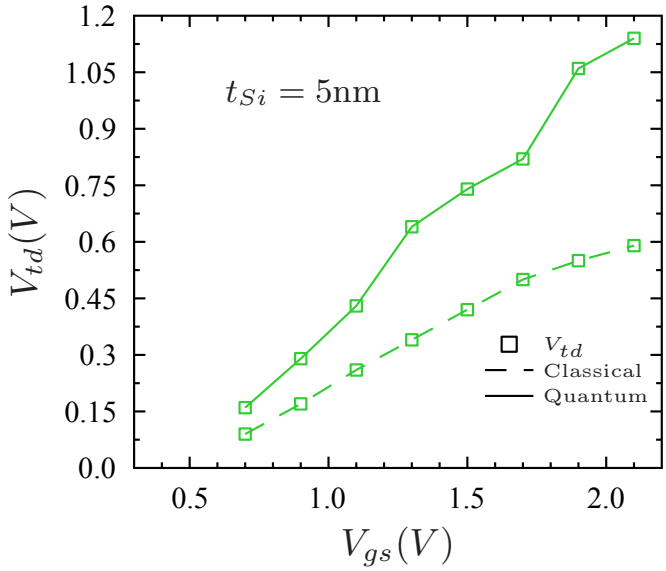

Figure 6.18: V_{td} dependence on V_{gs} for the DG structure and $t_{Si} = 5$nm. The quantum curve shows lack of saturation compared to the semiclassical one.

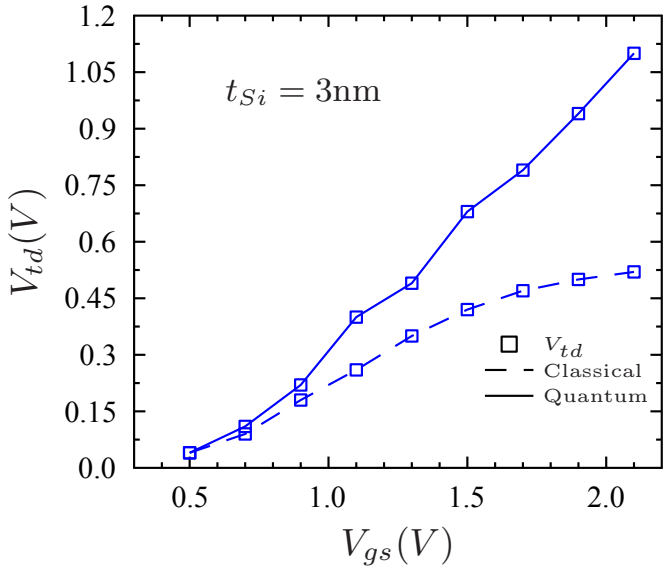

Figure 6.19: V_{td} dependence on V_{gs} for the DG device with $t_{Si} = 3$nm. Semiclassical curve clearly saturates compared to the behaviour including quantum confinement.

6.2 Subthreshold Swings (SS)

In Sec. 4.3.5, it was explained that two different swings need to be defined in TFETs given that in these devices the SS proves to strongly vary with V_{gs}. These two swings are the point swing, S_{pt}, corresponding to the smallest swing along the $I_{ds} - V_{gs}$ curve; and the average swing, S_{av}, which is defined as the swing between the point where BTBT starts (what we previously represented as V_{on}) and V_{tg}. S_{av} is the most interesting of these two swings from the point of view of switching performance.

6.2.1 DG–TFET

We first analyse the behaviour of S_{pt} and S_{av} for the DG structure. In Fig. 6.20, we show the variation of these parameters with the different body thicknesses considered in our simulations.

For $V_{ds} = 0.1$V, it can be noticed that semiclassical and quantum curves corresponding to S_{pt} lie very close to each other and manifestly below the reference level of 60mV/dec. These values for S_{pt} are consistent with the extraordinary sharpness of the transfer characteristics when BTBT condition is achieved and carriers begin to be injected into the channel. However, the most interesting remark concerns S_{av} that turns out to be smaller than 60mV/dec for $t_{Si} \leq 5$nm when quantum effects are taken into account. This represents an improvement on what we obtain in the semiclassical framework, where S_{av} only reaches that reference value for $t_{Si} = 3$nm. The reason for this considerable difference between semiclassical and quantum results regarding S_{av} —especially for $t_{Si} = 3$ and 5nm— can be found recalling the behaviour of V_{tg} reported in Figs. 6.8–6.10: the reduction of V_{tg} at $V_{ds} = 0.1$V due to confinement along with a slight increase of V_{on} causes S_{av} to be reduced.

Let us increase V_{ds} to 1.1V. In that case, the quantum corrected S_{av} curve now proves to be very close to the semiclassical one for $t_{Si} = 3$, 5 and 10nm, as the V_{tg} values arising from both approaches turn out to be very similar for that drain bias. From a global perspective, both S_{pt} and S_{av} tend to be raised when V_{ds} is increased from 0.1 to 1.1V.

If we further raise V_{ds} to 1.5V, another feature that may result in-

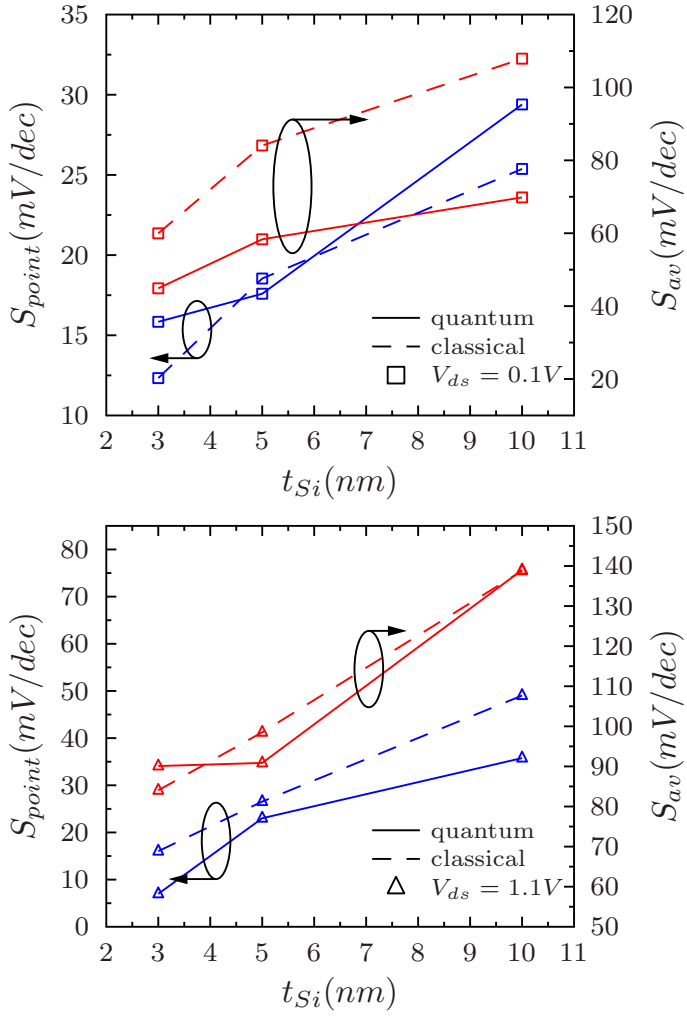

Figure 6.20: Subthreshold swings (point and average) as a function of t_{Si} at $V_{ds} = 0.1$ and 1.1V for the DG–TFET. Dashed lines stand for semiclassical simulations, solid lines include quantum confinement.

teresting to study is the degradation of the S_{pt} semiclassical curve as a consequence of the leakage current increase, which causes the steepest region of the n–branch to be hidden. Such behaviour becomes evident in light of Fig. 6.21. Observe how this effect is particularly noticeable when we increase t_{Si}.

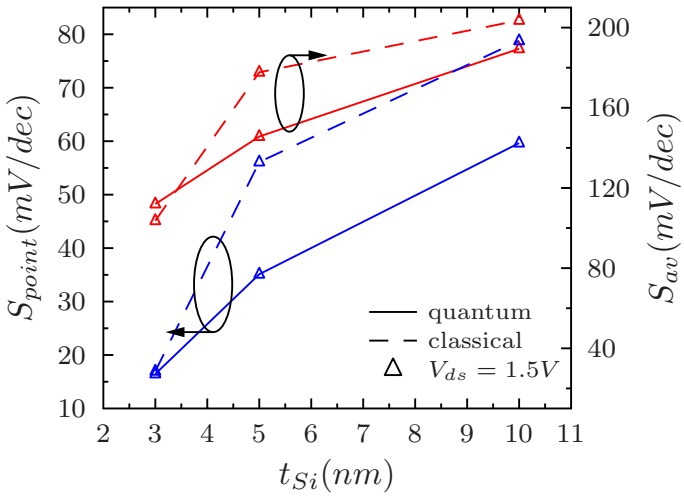

Figure 6.21: S_{pt} and S_{av} vs. t_{Si} at $V_{ds} = 1.5$V for the DG–TFET. Dashed lines, semiclassical simulations. Solid lines, including confinement.

6.2.2 SG–TFET

An equivalent analysis of how S_{pt} and S_{av} vary with t_{Si} for the SG–TFET is presented herein. In Fig. 6.22, the behaviour of both swings is shown for $V_{ds} = 0.1$ and 1.1V.

Similarly to what happened with the DG device, it can be noticed that for low V_{ds} the reduction of V_{tg} arising from the inclusion of quantum confinement again makes the semiclassical and quantum corrected curves corresponding to S_{av} greatly divert from each other. Concerning S_{pt}, the resulting curves for both approaches remain very close as leakage current is not significant when $V_{ds} = 0.1$V.

In the SG–TFET, leakage begins to be significant for semiclassical simulations at lower V_{ds} compared to the DG device. For that reason, at $V_{ds} = 1.1$V, the S_{av} curves do not remain so close as they do for the DG case. Additionally, as the n–branch of the semiclassical $I_{ds} - V_{gs}$ curves becomes progressively hidden due to leakage increase, we also observe a deviation between S_{pt} curves when confinement is accounted for.

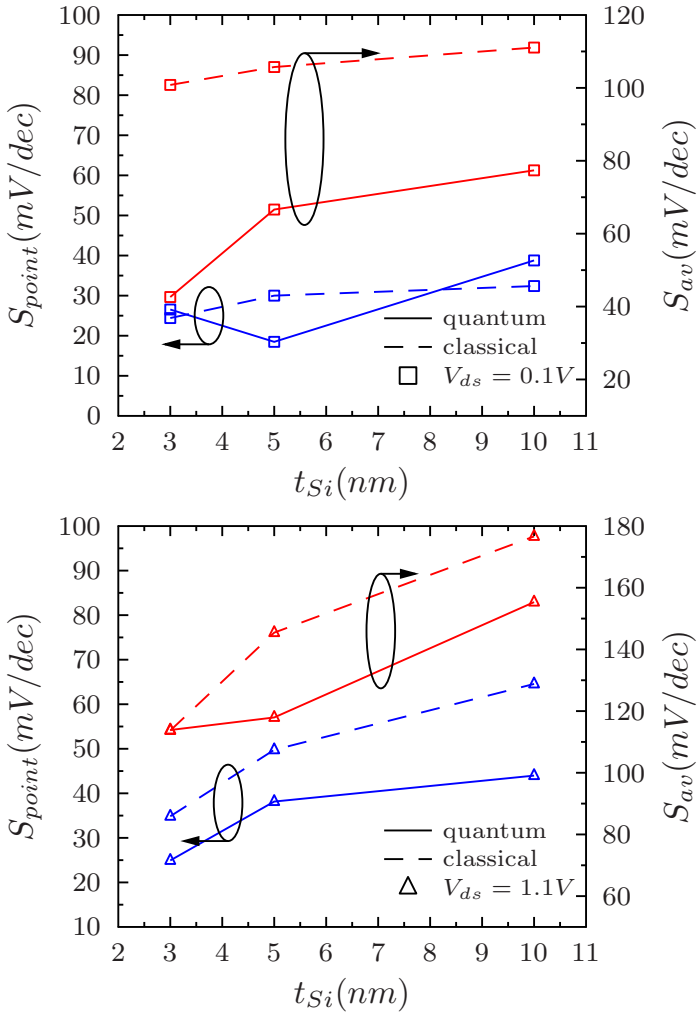

Figure 6.22: S_{pt} and S_{av} as a function of t_{Si} at $V_{ds} = 0.1$ and 1.1V for the SG configuration. Dashed lines stand for semiclassical simulations, solid lines include quantum confinement.

Chapter 7

Conclusions, future work and publications

7.1 Conclusions

7.1.1 Schottky Barrier MOSFETs

- We have confirmed that barrier lowering mechanisms are not satisfactorily implemented in Silvaco ATLAS when applied to tunneling injection of carriers. Simulations performed with the internal BL of ATLAS showed deviations from experimental results when tunneling contributes to total current.
- We have developed an iterative and self–consistent mechanism using existing ATLAS capabilities to account for barrier lowering mechanisms applied to thermionic emission, field emission and thermionic field emission. The proposed mechanism is a trade-off that allows to keep the widely used WKB approximation and corrects its deviations from the so–called "wide barrier" assumption through an adequate fitting of the carrier tunneling effective masses.
- We have combined this mechanism with a vertical discretization of the channel inside the SB–MOSFET that accounts for the Schottky barrier height dependence in the direction perpendicular to

the gate. This vertical discretization proves to be a fine–tuning that improves our iterative approach.

- To ensure convergence and avoid possible fluctuations that might destabilize the iterative calculations, the value of the electric field used to estimate barrier lowering processes must be extracted very close to the contacts but not at the interface. Typically, a distance of 0.5–1nm guarantees convergence without an excessive computational cost.

- We have applied the mechanism to estimate barrier lowering in NiSi and epitaxial $NiSi_2$ S/D SB–MOSFETs and compared the simulation results to experimental data. Very accurate fits are obtained for the regions dominated by tunnel current.

- We have reproduced in our simulations the experimental current reduction obtained when downsizing the gate length to 20nm as a consequence of the overlapping between source and drain Schottky barrier potential profiles.

- DIBL has been confirmed in devices with 20nm gate length and a lack of saturation in their output characteristics has been also accurately simulated. The effect of variations in the gate–to–silicide underlap has been analyzed.

7.1.2 Tunneling Field–Effect Transistors

- We have stated that the combination of the non–local band–to–band tunneling (BTBT) model for carrier injection (based on the non–uniformity of the electric field inside the tunneling barrier width), with the inclusion of quantum mechanical confinement through the self consistent coupled Schrödinger–Poisson model of ATLAS is not possible.

- We have determined that this was caused by (1) incompatibilities of the employed numerical solvers and (2) the impossibility of the

non–local BTBT model to inject carriers from (and to) bound states of the quantized conduction and valence bands.

- We have overcome the first obstacle by dividing each solving step into two separate iterations. In the first one, we self–consistently solve the Schrödinger and Poisson equations and obtain the potential, the quantum carrier concentrations and the discrete subband spectrum of both conduction and valence bands. In the second one, we calculate non–local tunneling injection using the potential and charge distributions of the previous iteration.

- We have circumvented the second hitch modifying the non–local BTBT model by increasing the bandgap at the tunneling junction making it coincident with the effective gap between the first bound states of conduction and valence bands. This way, the modified conduction and valence bands play the role of those first bound states, hence closely describing the actual situation.

- The simple approach outlined in the two previous points represents a quite interesting tool to account for quantum confinement using the widely accessible ATLAS simulator.

- We have applied the proposed approach to the simulation of semiclassical and quantum corrected currents in Double Gate TFETs and SOI Single Gate TFETs. The effects of body thickness variation, high–κ dielectrics and doping variation were elucidated.

- We have studied the effect of the inversion layer formation and the effect quantum confinement over threshold voltages and subthreshold swings, along with their impact on the surface channel potential.

- We have stated —using analytical expressions coming from local BTBT models— that the simulated semiclassical and quantum trends in threshold voltages can be understood by examining the influence of the minimum tunneling barrier width (w_{min}) variation on tunneling current.

7.2 Future work

We feel that the obtained results and conclusions of our work can be regarded as stages of a bigger roadmap which could significantly contribute to widen the existing knowledge of these novel devices.

Concerning the study of SB–MOSFETs, some ideas involve the application of our study of barrier lowering processes to asymmetric structures (with Schottky tunneling at the source and Ohmic junction at the drain created, for example, using highly doped pocket implants), or analyze its implementation in the presence of silicidation induced dopant segregation, which it is known to also affect the effective Schottky barrier height. Further work would be oriented to the inclusion of quantum confinement and to study how the replacement of conduction and valence bands by a discrete set of energy subbands causes the barrier heights to be increased.

In the case of TFETs, we generated simulation results comparing our developed procedure to include confinement with semiclassical existing approaches, and analyzed the impact on threshold voltages and subthreshold swings that the inclusion of confinement represents. This approach may be from now on combined with the implementation of engineering solutions leading, for example, to bandgap reductions at the source–to–channel junction either by straining the silicon or by using a material with a smaller bandgap at the source side. Analogously, the issue of scaling will need to be faced taking into account that in these devices the scaling conventions of MOSFETs do not necessarily apply. As far as the injection mechanisms in these devices are different, length scaling will have to be raised separately from the scaling of other parameters in order to better isolate and identify their effects.

7.3 Publications

The results presented in this thesis have given rise to the following publications:

7.3.1 Journals

- J.L. Padilla, L. Knoll, F. Gámiz, Q.T. Zhao, A. Godoy and S. Mantl, "Simulation of Fabricated 20nm Schottky Barrier MOSFETs on SOI: Impact of Barrier Lowering", *IEEE Transactions on Electron Devices*, vol. 59, no. 5, pp. 1320–1327, 2012.

- J.L. Padilla, F. Gámiz and A. Godoy, "A simple approach to quantum confinement in Tunneling Field–Effect Transistors", *IEEE Electron Device Letters*, vol. 33, no. 10, 2012.

- J.L. Padilla, F. Gámiz and A. Godoy , "Impact of quantum confinement on gate threshold voltage and subthreshold swings in Double Gate Tunnel FETs", *IEEE Transactions on Electron Devices*, In press.

7.3.2 International Conferences

- J.L. Padilla and F. Gámiz, "Barrier lowering implementation in SB–MOSFETs on SOI substrates", *in Proceedings ULIS 2010*, pp. 201–204, 2010.

- L. Donetti, F. Gámiz, F. Martínez–Carricondo, J.L. Padilla and N. Rodríguez, "Transport mass of holes in ultra–thin DGSOI devices", *in Proceedings EUROSOI 2011*, pp. 121–122, 2011.

Appendix A

ATLAS simulator

A.1 Description

The ATLAS simulator is one of the simulating tools developed by Silvaco (Silvaco, Inc.) [9], which provides general capabilities for physically–based 2D and 3D simulation of semiconductor devices. It essentially predicts the electrical characteristics that are associated with specified physical structures and bias conditions. To do so, the operation of the considered device is accounted for using a two or three dimensional grid divided in a series of points called nodes.

The main advantages of ATLAS as a physically–based simulator are:

- It is predictive.
- It provides insight.
- It conveniently captures and visualizes theoretical knowledge.

This physically–based simulation approach differs from the also existing concept of emprical modeling. The aim of the latter is to obtain analytic formulae that approximate existing data with good accuracy and minimum complexity. In that context, any hint of predictive capability disappears when dealing with this empirical modeling.

Physically–based simulators have become a paramount tool in the characterization of semiconductor devices, as using them is manifestly

quicker and cheaper than performing experiments. On the other hand, however, the employment of this type of simulators also presents some obvious drawbacks: prior to begin simulating, it is mandatory to incorporate all the relevant physics involved in the processes as well as the details of the numerical solvers that will be used. That is to say, before the simulating process starts, all the details concerning the problem to be solved (exact structure of the device under study, physical models, numerical solvers and bias conditions) need to be perfectly delimited. Or, in other words, before assembling the jigsaw, all its pieces need to be known.

Efficient numerical techniques are implemented in ATLAS involving:

- Accurate and robust discretization techniques.
- Gummel, Newton and block–Newton nonlinear iteration strategies.
- Efficient solvers, both direct and iterative, for linear subproblems.
- Powerful initial guess strategies.
- Small–signal calculation techniques that converge at all frequencies.
- Stable and accurate time integration

In the development of our work, other Silvaco software fully compatible with ATLAS was employed: (*i*) DeckBuild, as the interactive run–time environment; and (*ii*) TonyPlot, as the interactive graphics and analysis package.

A.2 Simulation process structure

Using Deckbuild as the run–time environment, a text file including all the commands to be executed by ATLAS needs to be created as input file. After running this input file, ATLAS produces three types of output files:

1 **The run–time outputs**. They give information about the progress of the simulation, as well as report possible error and warning messages during execution.

2 **The *log* files**. They store all terminal voltages and currents from the device analysis.

3 **The *solution* or structure files**. They store the 2D or 3D values of solution variables within the device at a given bias point.

Fig. A.1 illustrates how the information flows during the simulation process.

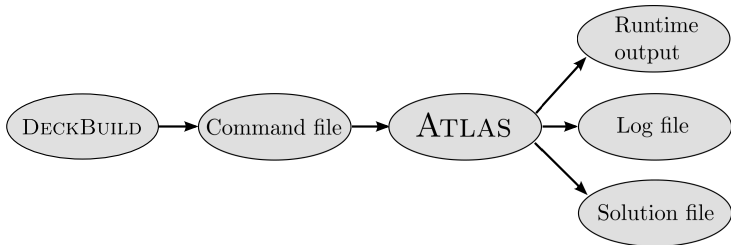

Figure A.1: Schematic representation of a simulation involving ATLAS.

A.3 ATLAS input files

Commands in the input file need to follow a definite order. They are commonly nested in five different statement groups depending on the function that they carry out. These groups must also occur in the correct order. The following table shows the appropriate sequence in which commands must be specified.

Group	*Statements*
1. Structure Specification	MESH
	REGION
	ELECTRODE
	DOPING
2. Material Models Specification	MATERIAL
	MODELS
	CONTACT
	INTERFACE
3. Numerical Method Selection	METHOD
4. Solution Specification	LOG
	SOLVE
	LOAD
	SAVE
5. Results Analysis	EXTRACT
	TONYPLOT

Part III

Conclusions in Spanish

Conclusiones

Schottky Barrier MOSFETs

- Hemos confirmado que los mecanismos de bajada de barrera no se implementan de manera satisfactoria en ATLAS en lo que se refiere a la inyección de portadores por efecto túnel. Las simulaciones realizadas con los mecanismos internos de ATLAS mostraron claras discrepancias con los resultados experimentales en aquellas regiones de las curvas características donde la contribución por túnel era dominante.

- Hemos desarrollado un mecanismo iterativo y autoconsistente tomando como base las herramientas de ATLAS para describir los procesos de bajada de barrera aplicados tanto emisión termoiónica, como a emisión de campo, como a emisión de campo termoiónica. El mecanismo propuesto permite seguir usando la aproximación WKB y corrige sus desviaciones mediante ajustes en las masas efectivas de túnel.

- Hemos combinado este mecanismo con una discretización vertical del canal en los transistores SB–MOSFET que permite tener en cuenta la dependencia de la altura de la barrera Schottky en la dirección perpendicular a la longitud de puerta. Esta discretización vertical demostró ser un ajuste fino que mejora el mecanismo de bajada de barrera.

- Para garantizar la convergencia, los valores del campo eléctrico usados en la estimación de la bajada de barrera deben extraerse muy próximos a la unión metal–semiconductor, pero no justamente en la interfase. Una distancia de 0.5–1nm garantiza convergencia sin que ello suponga un excesivo coste computacional.

- Hemos aplicado el mecanismo desarrollado para estimar las bajadas de barrera en transistores SB–MOSFET con contactos de fuente y drenador fabricados con NiSi y $NiSi_2$ epitaxial, y comparado los resultados experimentales disponibles con los proce-

dentes de nuestras simulaciones. El resultado fue una notable precisión en los ajustes.

- Hemos sido capaces de reproducir mediante simulación el efecto de canal corto consistente en la disminución de corriente cuando la longitud de puerta se reduce de 50nm a 20nm.

- Hemos reproducido el efecto DIBL para una longitud de puerta de 20nm, y hemos predicho mediante simulación el efecto que variaciones en la distancia de underlap tendrían sobre los niveles de corriente en el dispositivo fabricado con $NiSi_2$.

Tunnel FETs

- Hemos constatado que la descripción del fenómeno de corriente túnel banda a banda mediante el modelo no local de ATLAS es incompatible a priori con la inclusión de los efectos cuánticos de confinamiento mediante el modelo autoconsistente de Schrödinger–Poisson del simulador.

- Determinamos que este problema se debía a: (1) incompatibilidades de los métodos numéricos empleados en cada modelo, y (2) la imposibilidad del modelo no local de inyectar portadores entre los primeros estados excitados procedentes de la discretización cuántica en subniveles de energía de las bandas de conducción y valencia.

- El primer obstáculo lo solucionamos dividiendo cada etapa de cálculo en dos pasos separados. En el primero, se resuelven de manera autoconsistente las ecuaciones de Schrödinger–Poisson y se obtienen las distribuciones cuánticas de carga, el potencial y el espectro de niveles de energía de las bandas de conducción y valencia. En el segundo, se implementa la inyección de portadores tomando como marco el escenario obtenido en el paso anterior.

- El segundo inconveniente lo resolvimos modificando el modelo no local de túnel banda a banda mediante el incremento de la anchura

del gap para hacer coincidentes los extremos de las bandas de conducción y valencia con los primeros estados excitados obtenidos de la inclusión de confinamiento. De esta manera, al efectuar la inyección se consigue simular de manera bastante aproximada la situación real.

- El mecanismo descrito en los dos puntos anteriores constituye una interesante herramienta para abordar los efectos cuánticos de confinamiento en estos dispositivos mediante el uso de ATLAS.

- Hemos aplicado este procedimiento para la simulación de transistores TFET de puerta única en silicio sobre aislante y de doble puerta. Se han estudiado en ellos los efectos de la variación del espesor del canal, de la permitividad del aislante de puerta y de los niveles de dopado.

- Hemos analizado el papel desempeñado por la formación de la capa de inversión y por la inclusión del confinamiento sobre las tensiones umbrales y sobre las pendientes subumbrales, así como el efecto que dichos actores ejercen sobre el potencial de superficie en el canal.

- Hemos mostrado, con la ayuda de expresiones analíticas procedentes de los modelos locales de corriente túnel banda a banda, que los comportamientos semiclásicos y cuánticos de las tensiones umbrales están directamente asociados con cómo las modificaciones en la distancia mínima de túnel afectan a la corriente total en estos dispositivos.

Bibliography

[1] D. Connelly, C. Faulkner, D. Grupp, and J. Harris, "A new route to zero–barrier metal source/drain MOSFETs," *IEEE Trans. Nanotechnol.*, vol. 3, no. 1, pp. 98–104, 2004.

[2] L. Knoll, Q. Zhao, R. Luptak, S. Trellenkamp, K. Bourdelle, and S. Mantl, "20nm gate length Schottky MOSFETs with ultra thin NiSi/epitaxial $NiSi_2$ source/drain," *in Proc. ULIS*, 2011.

[3] S. Sze, *Physics of Semiconductor Devices*. Wiley, 2007.

[4] R. D. et al, "Design of micron MOS switching devices," in *Electron Devices Meeting, 1972 International*, 1972.

[5] J. Larson and J. Snyder, "Overview and status of metal S/D Schottky–barrier MOSFET technology," *IEEE Transactions on Electron Devices*, vol. 53, no. 5, pp. 1048–1058, 2006.

[6] D. Lenoble, "Advanced junction fabrication challenges at the 45nm node," in *Semiconductor Fabtech 30th edition*, vol. 114, 2006.

[7] T. Skotnicki, A. Pouydebasque, and D. Villanueva, in *USJ Conf. Proceedings*, 2003.

[8] D. H. et al., "FinFET a self–aligned double–gate MOSFET scalable to 20nm," *IEEE Trans. Electron Devices*, vol. 47, no. 12, pp. 2320–2325, 2000.

[9] Silvaco, *ATLAS users manual*, June 2012.

[10] E. Dubois and G. Larrieu, "Low Schottky barrier source/drain for advanced MOS architecture: device design and material considerations," *Solid–State Electronics*, vol. 46, 2002.

[11] G. Larrieu and E. Dubois, "Schottky-barrier source/drain MOSFETs on ultrathin SOI body with a tungsten metallic midgap gate," *IEEE Electron Device Lett.*, vol. 25, no. 12, pp. 801–803, 2004.

[12] J. Larson and J. Snyder, *Schottky barrier CMOS – technology overview.* Spinnaker Semiconductor, 2004.

[13] S. Thompson, P. Packan, and M. Bohr, *MOS scaling: Transistor challenges for the 21st century.* Intel Technology Journal, Q3, 1998.

[14] S. Swirhun, E. Sangiorgi, A. Weeks, R. Swanson, K. Saraswat, and R. Dutton, "A VLSI–suitable Schottky–barrier CMOS process," *IEEE Transactions on Electron Devices*, vol. 32, no. 2, pp. 194–202, 1985.

[15] M. Sugino, L. Akers, and M. Rebeschini, "CMOS latch–up elimination using Schottky barrier PMOS," in *IEDM Tech. Dig.*, 1982, pp. 462–465.

[16] ——, "Latch–up free Schottky–barrier CMOS," *IEEE Transactions on Electron Devices*, vol. 30, no. 2, pp. 110–118, 1983.

[17] S. Zhu, H. Yu, S. Whang, J. Chen, C. Shen, C. Zhu, S. Lee, M. Li, D. Chan, W. Yoo, A. Du, C. Tung, J. Singh, A. Chin, and D. Kwong, "Schottky-barrier S/D MOSFETs with high κ gate dielectrics and metal-gate electrode," *IEEE Electron Device Lett.*, vol. 25, no. 5, pp. 268–270, 2004.

[18] J. Knoch and J. Appenzeller, "Impact of the channel thickness on the performance of Schottky barrier metal-oxide-semiconductor field-effect transistors," *App. Phys. Lett.*, vol. 81, no. 16, pp. 3082–3084, 2002.

[19] E. Dubois and G. Larrieu, "Measurement of low Schottky barrier heights applied to metallic source/drain metal-oxide-semiconductor field effect," *J. Appl.Phys.*, vol. 96, pp. 729–737, 2004.

[20] V. Chin, J. Storey, and M. Green, "Characteristics of p–type PtSi Schottky diodes under reverse bias," *J. Appl. Phys.*, vol. 68, pp. 4127–4132, 1990.

[21] S. Zhu, J. Chen, M.-F. Li, S. Lee, J. Singh, C. Zhu, A. Du, C. Tung, A. Chin, and D. Kwong, "N–type Schottky barrier S/D MOSFET using Ytterbium silicide," *IEEE Electron Device Lett.*, vol. 25, no. 8, pp. 565–567, 2004.

[22] M. Jang, Y. Kim, J. Shin, and S. Lee, "Characterization of erbium–silicided Schottky diode junction," *IEEE Electron Device Lett.*, vol. 26, no. 6, pp. 354–356, 2005.

[23] D. Connelly, C. Faulkner, P. Clifton, and D. Grupp, "Fermi–level depinning for low–barrier Schottky source/drain transistors," *Appl. Phys. Lett.*, vol. 88, no. 1, pp. 012 105–012 108, 2006.

[24] Q. Zhao, E. Rije, U. Bruer, S. Lenk, and S. Mantl, "Tuning of silicide SBHs by segregation of sulfur atoms," in *Proc. IEEE*, 2004, pp. 456–459.

[25] M. Tao, D. Udeshi, N. Basit, E. Maldonado, and W. Kirk, "Removal of dangling bonds and surface states on silicon (001) with a monolayer of selenium," *Applied Physics Letters*, vol. 82, no. 10, pp. 1559–1561, 2003.

[26] M. Tao, S. Agarwal, D. Udeshi, N. Basit, E. Maldonado, and W. Kirk, "Low Schottky barriers on *n*–type silicon (001)," *Applied Physics Letters*, vol. 83, no. 13, pp. 2593–2595, 2003.

[27] G. Song, M. Ali, and M. Tao, "A high Schottky barrier of 1.1eV between Al and S–Passivated p–type Si (100) surface," *IEEE Electron Device Lett.*, vol. 28, no. 1, pp. 71–73, 2007.

[28] R. Saiz-Pardo, R. Perez, F. Garcia-Vidal, R. Whittle, and F. Flores, "Systematic theoretical studies of the Schottky barrier control by passivating atomic intralayers," *Surf. Sci.*, vol. 426, no. 1, pp. 26–37, 1999.

[29] T. Yamauchi, A. Kinoshita, Y. Tsuchiya, J. Koga, and K. Kato, "1nm NiSi/Si junction design based on first-principles calculation for ultimately low contact resistance," *IEDM Tech. Dig.*, pp. 385–388, 2006.

[30] R. A. Vega, "A Comparative Study of Dopant-Segregated Schottky and Raised Source/Drain Double-Gate MOSFETs," *IEEE Transactions on Electron Devices*, vol. 55, no. 10, pp. 2665–2677, 2008.

[31] R. A. Vega, V. C. Lee, and T.-J. K. Liu, "The Effect of Random Dopant Fluctuation on Specific Contact Resistivity," *IEEE Transactions on Electron Devices*, vol. 57, no. 1, pp. 273–281, 2010.

[32] H. Bethe, "Theory of the Boundary Layerof Crystal Rectifiers," *MIT Radiat. Lab. Rep.*, vol. 43, p. 12, 1942.

[33] W. Schottky, "Halbleittertheorie der Sperrschicht," *Naturwissenschaften*, vol. 26, p. 843, 1938.

[34] E. Rhoderick and R. Williams, *Metal–Semiconductor Contacts*. Oxford Science Publications, 1988.

[35] F. Padovani and R. Stratton, "Field and Thermionic–Field Emission in Schottky Barriers," *Solid–State Electron.*, vol. 9, p. 695, 1966.

[36] J. Snyder, C. Helms, and Y. Nishi, "Experimental investigation of a PtSi source and drain field emission transistor," *Applied Physics Letters*, vol. 67, no. 10, pp. 1420–1422, 1995.

[37] A. Cowley and S. Sze, "Surface states and barrier height of metal-semiconductor systems," *J. Appl. Phys.*, vol. 36, pp. 3212–3220, 1965.

[38] C. Crowell and G. Roberts, "Surface state and interface effects on the capacitance-voltage relationship in Schottky barriers," *J. Appl. Phys.*, vol. 40, pp. 3726–3730, 1969.

[39] J. Levine, "Schottky-barriers anomalies," *J. Appl. Phys.*, vol. 42, pp. 3991–3999, 1971.

[40] V. Heine, "Theory of Surface States," *Phys. Rev.*, vol. 138, p. 1689, 1965.

[41] K. Shenai, "Current Transport Mechanisms in Atomically Abrupt Metal-Semiconductor Interfaces," *IEEE Transactions on Electron Devices*, vol. 35, no. 4, pp. 468–482, 1988.

[42] W. Mönch, "Electronic properties of semiconductor interfaces," *Springer*, 2004.

[43] J. Andrews and M. Lepselter, "Reverse Current-Voltage Characteristics of Metal-Silicide Schottky Diodes," *Solid-State Electronics*, vol. 13, pp. 1011–1023, 1970.

[44] L. Knoll, Q. Zhao, S. Habicht, C. Urban, B. Ghyselen, and S. Mantl, "Ultrathin Ni Silicides with low contact resistance on strained and unstrained silicon," *IEEE Electron Device Letters*, vol. 31, no. 4, pp. 350–352, 2010.

[45] M. Shin, "Computational Study on the Performance of Multiple-Gate Nanowire Schottky Barrier MOSFETs," *IEEE Transactions on Electron Devices*, vol. 55, no. 3, pp. 737–742, 2008.

[46] E. O. Kane, "Theory of Tunneling," *Journal of Applied Physics*, vol. 32, no. 1, pp. 83 –91, 1961.

[47] E. Kane, "Tunneling in InSb," *J. Phys. Chem. Solids*, vol. 12, p. 181, 1960.

[48] Q. Zhang, W. Zhao, and A. Seabaugh, "Low-subthreshol-swing tunnel transistors," *IEEE Electron Device Letters*, vol. 27, p. 297, 2006.

[49] G. Baccarani, M. Wordeman, and R. Dennard, "Generalized scaling theory and its application to a 1/4 micrometer MOSFET design," *IEEE Trans. Electron Devices*, vol. 31, no. 4, pp. 452–462, 1984.

[50] ITRS, *The International Technology Roadmap for Semiconductors (ITRS), Semiconductor Industry Association.* http://public.ITRS.net, 2011.

[51] J. Kedzierski, P. Xuan, V. Subramanian, J. Bokor, T.-J. King, C. Hu, and E. Anderson, "A 20nm gate–length ultra–thin body p–mosfet with silicide S/D," *Superlattices Microstruct.*, vol. 28, no. 5/6, pp. 445–452, 2000.

[52] Y. Nishi, *Insulated gate field effect transistor and its manufacturing method.* Patent 587 527, 1970.

[53] M. Lepselter and S. Sze, "SB–IGFET: An insulated–gate field–effect transistor using Schottky barrier contacts for source and drain," *Proc. IEEE*, vol. 56, no. 8, pp. 1400–1402, 1968.

[54] C. Koeneke, S. Sze, R. Levin, and E. Kinsbron, "Schottky MOSFET for VLSI," in *IEDM Tech. Dig.*, 1981, pp. 367–370.

[55] C. Koeneke and W. Lynch, "Lightly doped Schottky MOSFET," in *IEDM Tech. Dig.*, 1982, pp. 466–469.

[56] T. Mochizuki and K. Wise, "An n–channel MOSFET with Schottky source and drain," *IEEE Electron Device Lett.*, vol. 5, no. 4, pp. 108–111, 1984.

[57] C. Oh, Y. Koh, and C. Kim, "A new P–Channel MOSFET structure with Schottky–Clamped source and drain," in *IEDM Tech. Dig.*, 1984, pp. 609–612.

[58] B.-Y. Tsui and M.-C. Chen, "A novel process for high–performance Schottky barrier PMOS," *J. Electrochem. Soc.*, vol. 136, no. 5, pp. 1456–1459, 1989.

[59] M. Kimura and T. Matsudate, "A new type of Schottky tunnel transistor," *IEEE Electron Device Lett.*, vol. 15, no. 10, pp. 412–414, 1994.

[60] J. Tucker, C. Wang, and P. Carney, "Silicon field–effect transistor based on quantum tunneling," *Appl. Phys. Lett.*, vol. 65, no. 5, pp. 618–620, 1994.

[61] J. Tucker, C. Wang, J. Lyding, T.-C. Shen, and G. Abeln, "Nanometer scale MOSFETs and STM paterning on Si," in *Proc. Int. Conf. Solid State Devices Mater.*, 1994, pp. 322–324.

[62] J. Snyder, *The physics and technology of platinum silicide source and drain field effect transistors.* Ph.D. dissertation, Stanford Univ., Stanford, CA, 1996.

[63] L. Calvet, H. Luebben, M. Reed, C. Wang, J. Snyder, and J. Tucker, "Subthreshold and scaling of PtSi Schottky barrier MOSFETs," *Superlattices and Microstructures*, vol. 28, no. 5/6, pp. 501–506, 2000.

[64] ——, "Suppression of leakage current in Schottky barrier metal–oxide–semiconductor field–effect–transistors," *Journal of Applied Physics*, vol. 91, no. 2, pp. 757–759, 2002.

[65] M. Fritze, C. Chen, S. Calawa, D. Yost, B. Wheeler, P. Wyatt, C. Keast, J. Snyder, and J. Larson, "High–speed Schottky–barrier pMOSFET with f_T = 280ghz," *IEEE Electron Device Letters*, vol. 25, no. 4, pp. 220–222, 2004.

[66] K. Ikeda, Y. Yamashita, A. Endoh, T. Fukano, K. Hikosaka, and T. Mimura, "50–nm gate Schottky source/drain p–MOSFETs with a SiGe channel," *IEEE Electron Device Letters*, vol. 23, no. 11, 2002.

[67] J. Kudzierski, P. Xuan, V. Subramanian, J. Bokor, T.-J. King, C. Hu, and E. Anderson, "A 20nm gate–length ultra–thin body

p–MOSFET with silicide source/drain," *Superlattices and Microstructures*, vol. 28, no. 5/6, pp. 445–452, 2000.

[68] G. Larrieu and E. Dubois, "Integration of PtSi–based Schottky–barrier p–MOSFETs with a midgap tungsten gate," *IEEE Transactions on Electron Devices*, vol. 52, no. 12, pp. 2720–2726, 2005.

[69] C. Wang, J. Snyder, and J. Tucker, "Sub–40nm PtSi Schottky source/drain metal–oxide–semiconductor field–effect transistor," *Applied Physics Letters*, vol. 74, no. 8, pp. 1174–1176, 1999.

[70] S.-Y. Zhu and M.-F. Li, "High performance p–channel Schottky barrier MOSFETs with self–aligned PtSi source/drain on thin film SOI substrate," *Chinese Physics Letters*, vol. 22, no. 8, pp. 2020–2022, 2005.

[71] M. Jang, S. Lee, and K. Park, "Erbium silicided n–type Schottky barrier tunnel transistors for nanometer regime applications," *IEEE Transactions on Nanotechnology*, vol. 2, no. 4, pp. 205–209, 2003.

[72] M. Jang, J. Oh, S. Maeng, W. Cho, K. Kang, and K. Park, "Characteristics of erbium–silicided n–type Schottky barrier tunnel transistors," *Applied Physics Letters*, vol. 83, no. 13, pp. 2611–2613, 2003.

[73] M. Jang, Y. Kim, J. Shin, and S. Lee, "A 50nm gate–length erbium–silicided n–type Schottky barrier metal–oxide–semiconductor field–effect transistor," *Applied Physics Letters*, vol. 84, no. 5, pp. 741–743, 2004.

[74] M. Jang, Y. Kim, J. Shin, S. Lee, and K. Park, "Novel properties of erbium–silicided n–type Schottky barrier metal–oxide–semiconductor field–effect transistors," *Journal of Semiconductor Technology and Science*, vol. 4, no. 2, pp. 94–98, 2004.

[75] T. Kinoshita, R. Hasumi, M. Hamaguchi, K. Miyashita, T. Komoda, A. Kinoshita, J. Koga, K. Adachi, Y. Toyoshima,

T. Nakayama, S. Yamada, and F. Matsuoka, "Ultralow voltage operations in bulk CMOS logic circuits with dopant segregated Schottky source/drain transistors," in *IEDM 2006*, 2006.

[76] J. Knoch, M. Zhang, Q. Zhao, S. Lenk, S. Mantl, and J. Appenzeller, "Effective Schottky barrier lowering in silicon–on–insulator Schottky–barrier metal–oxide–semiconductor field–effect transistors using dopant segregation," *Applied Physics Letters*, vol. 87, pp. 263 505–1–263 505–3, 2005.

[77] M. Zhang, J. Knoch, Q. Zhao, S. Lenk, J. Appenzeller, U. Breuer, and S. Mantl, "Dopant segregation in Schottky–barrier SOI–MOSFETs," in *ULIS*, 2005.

[78] Q. Zhao, M. Zhang, S. Lenk, and S. Mantl, "Tuning of NiSi/Si Schottky barrier heights by sulfur segregation during Ni silicidation and fabrication of Schottky MOSFETs," in *ULIS*, 2005, pp. 19–22.

[79] S. Feste, M. Zhang, J. Knoch, and S. Mantl, "Impact of variability on the performance of SOI Schottky barrier MOSFETs," *Solid–State Electronics*, vol. 53, no. 4, pp. 418–423, 2009.

[80] G. Patil, , and S. Qureshi, "A novel δ-doped partially insulated dopant-segregated Schottky barrier SOI MOSFET for analog/RF applications," *Semiconductor Science and Technology*, vol. 26, no. 8, p. 085002, 2011.

[81] J. Luo, Z.-J. Qiu, D. Zhang, P.-E. Hellstrom, M. Ostling, and S.-L. Zhang, "Effects of Carbon on Schottky Barrier Heights of NiSi Modified by Dopant Segregation," *IEEE Electron Device Letters*, vol. 30, no. 6, pp. 608–610, 2009.

[82] J. Guo and M. Lundstrom, "A computational study of thin–body, double–gate, Schottky barrier MOSFETs," *IEEE Transactions on Electron Devices*, vol. 49, no. 11, pp. 1897–1902, 2002.

[83] M. Jang, K. Kang, S. Lee, and K. Park, "Simulation of Schottky barrier tunnel transistor using simple boundary conditions," *Applied Physics Letter*, vol. 82, no. 16, pp. 2718–2720, 2003.

[84] R. A. Vega, "On the modeling and design of Schottky field–effect transistors," *IEEE Transactions on Electron Devices*, vol. 53, no. 4, pp. 866–874, 2006.

[85] ——, "Comparison Study of Tunneling Models for Schottky Field Effect Transistors and the Effect of Schottky Barrier Lowering," *IEEE Transactions on Electron Devices*, vol. 53, no. 7, pp. 1593–1600, 2006.

[86] M. Schwarz, M. Weidemann, A. Kloes, and B. Íñiguez, "2d analytical calculation of the electrostatic potential in lightly doped Schottky barrier Double-Gate MOSFET," *Solid–State Electronics*, vol. 54, no. 11, pp. 1372–1380, 2010.

[87] J. Padilla, L. Knoll, F. Gamiz, Q. Zhao, A. Godoy, and S. Mantl, "Simulation of fabricated 20nm Schottky Barrier MOSFETs on SOI: Impact of Barrier Lowering," *IEEE Transactions on Electron Devices*, vol. 59, no. 5, pp. 1320–1327, 2012.

[88] L. Zeng, Z. Xia, G. Du, J. Kang, R. Han, and X. Liu, "Gate-Induced Image Force Barrier Lowering in Schottky Barrier FETs," *IEEE Transactions on Nanotechnology*, vol. 8, no. 1, pp. 10–15, 2009.

[89] D. Neamen, *Semiconductor Physics and Devices*. McGraw Hill, 2003.

[90] J. Bardeen, "Surface States and Rectification at a Metal Semiconductor Contact," *Phys. Rev.*, vol. 71, p. 717, 1947.

[91] A. Cowley and S. Sze, "Carrier transport across Metal-Semiconductor barriers," *Mass. Inst. Technol. Radiat. Lab. Rep.*, vol. 43, p. 12, 1965.

[92] C. Mead and W. Spitzer, "Fermi level position at metal–semiconductor interfaces," *Physical Review*, vol. 134, no. 3A, pp. A713–A716, 1964.

[93] C. Crowell and S. Sze, "Current Transport in Metal–Semiconductor Barriers," *Solid–State Electronics*, vol. 9, p. 1035, 1966.

[94] C. Chang and S. Sze, "Carrier transport across metal–semiconductor barriers," *Solid–State Electron.*, vol. 13, p. 727, 1970.

[95] J.-P. Colinge, *Silicon–On–Insulator Technology: materials to VLSI*. 3rd Edition, Kluwer Academic Publishers, 2004.

[96] M. Zhang, J. Knoch, J. Appenzeller, and S. Mantl, "Improved Carrier Injection in Ultrathin–Body SOI Schottky–Barrier MOSFETs," *Electron Device Letters*, vol. 28, no. 3, pp. 223–225, 2007.

[97] J. Knoch and J. Appenzeller, "Impact of the channel thickness on the performance of Schottky barrier metal–oxide–semiconductor field–effect transistors," *Applied Physics Letters*, vol. 81, no. 16, pp. 3082–3084, 2002.

[98] J. Knoch, M. Zhang, S. Mantl, and J. Appenzeller, "On the performance of single–gated ultrathin–body SOI Schottky–barrier MOSFETs," *IEEE Transactions on Electron Devices*, vol. 53, no. 7, pp. 1669–1674, 2006.

[99] J. Knoch, M. Zhang, J. Appenzeller, and S. Mantl, "Physics of ultrathin–body silicon–on–insulator Schottky–barrier field–effect transistors," *Applied Physics A: Materials Science & Processing*, vol. 87, no. 3, pp. 351–357, 2007.

[100] C. Urban, M. Emam, C. Sandow, J. Knoch, Q.-T. Zhao, J.-P. Raskin, and S. Mantl, "Radio–Frequency study of Dopant–Segregated n–type SB–MOSFETs on Thin–Body SOI," *IEEE Electron Device Letters*, vol. 31, no. 6, pp. 537–539, 2010.

[101] Z. Qiu, Z. Zhang, M. Östling, and S.-L. Zhang, "A Comparative Study of Two Different Schemes to Dopant Segregation at NiSi/Si and PtSi/Si Interfaces for Schottky Barrier Height Lowering," *IEEE Transactions on Electron Devices*, vol. 55, no. 1, pp. 396–403, 2008.

[102] L. Hutin, M. Vinet, T. Poiroux, C. L. Royer, B. Previtali, C. Vizioz, D. Lafond, Y. Morand, M. Rivoire, F. Nemouchi, V. Carron, T. Billon, S. Deleonibus, and O. Faynot, "Dual metallic source and drain integration on planar Single and Double Gate SOI CMOS down to 20nm: Performance and scalability assessment," *IEDM Tech. Dig.*, pp. 1–4, 2009.

[103] L. Knoll, Q. Zhao, S. Habicht, C. Urban, K. Bourdelle, and S. Mantl, "Formation and characterization of ultra-thin ni silicides on strained and unstrained silicon," *in 2010 International Workshop on Junction Technology Extended Abstracts*, 2010.

[104] Synopsys, *Sentaurus TCAD*, ver. 10.0.

[105] K. Matsuzawa, K. Uchida, and A. Nishiyama, "A unified simulation of Schottky and ohmic contacts," *IEEE Transactions on Electron Devices*, vol. 47, no. 1, pp. 103–108, 2000.

[106] M. Ieong, P. Solomon, S. Laux, H. Wong, and D. Chidambarro, "Comparison of raised and Schottky source/drain MOSFETs using a novel tunneling contact model," *IEDM Tech. Dig.*, pp. 733–736, 1998.

[107] S. Xiong, T. King, and J. Bokor, "A comparison study of symmetric ultrathin-body double-gate devices with metal source/drain and doped sorce drain," *IEEE Trans. Electron Devices*, vol. 52, no. 8, pp. 1859–1867, 2005.

[108] C. Huang, W. Zang, and C. Yang, "Two-dimensional numerical simulation of Schottky barrier MOSFET with channel length to 10nm," *IEEE Trans. Electron Devices*, vol. 45, no. 4, pp. 842–848, 1998.

[109] K. Brennan and C. Summers, "Theory of resonant tunneling in a variably spaced multiquantum well structure," *J. Appl. Phys.*, vol. 61, no. 2, pp. 614–623, 1987.

[110] R. Rengel, E. Pascual, and M. Martin, "Injected current and quantum transmission coefficient in low Schottky barriers: WKB and Airy approaches," *IEEE Electron Device Lett.*, vol. 28, no. 2, pp. 171–173, 2007.

[111] K. Boucart and A. Ionescu, "Double Gate Tunnel FET with High-κ Gate Dielectric," *IEEE Transactions on Electron Devices*, vol. 54, no. 7, pp. 1725–1733, 2007.

[112] L. Esaki, "New Phenomenon in Narrow Germanium p-n junctions," *Physical Review*, vol. 109, pp. 603–604, 1958.

[113] J. Quinn, G. Kawamoto, and B. McCombe, "Subband spectroscopy by surface channel tunneling," *Surface Science*, vol. 73, pp. 190–196, 1978.

[114] K. Takeda, H. Matsuoka, Y. Igura, and S. Asai, "A band-to-band tunneling MOS device (B^2T-MOSFET)," in *IEEE IEDM Tech. Dig.*, 1988, pp. 402–405.

[115] S. Banerjee, W. Richardson, J. Coleman, and A. Chaterjee, "A new three-terminal tunnel device," *IEEE Electron Device Lett.*, vol. 8, no. 8, pp. 347–349, 1987.

[116] T. Baba, "Proposal for surface tunnel transistors," *Jpn. J. Appl. Phys.*, vol. 31, pp. L455–L457, 1992.

[117] T. Uemura and T. Baba, "Characterization of depletion-type surface tunnel transistors," *Jpn. J. Appl. Phys.*, vol. 31, pp. L1727–L1729, 1992.

[118] W. Reddick and G. Amaratunga, "Silicon surface tunnel transistor," *Appl. Phys. Lett.*, vol. 67, no. 4, pp. 494–497, 1995.

[119] J. Koga and A. Toriumi, "Three-terminal silicon surface junction tunneling device for room temperature operation," *IEEE Electron Device Lett.*, vol. 20, no. 10, pp. 529–531, 1999.

[120] W. Hansch, C. Fink, J. Schulze, and I. Eisele, "A vertical MOS-gated Esaki tunneling transistor in silicon," *Thin Solid Films*, vol. 369, pp. 387–389, 2000.

[121] J. Schulze, C. Fink, T. Sulima, I. Eisele, and W. Hansch, "MOS-gated tunneling transistor in Si(100) and Si(111)," *Thin Solid Films*, vol. 380, pp. 154–157, 2000.

[122] W. Hansch, P. Borthen, J. Schulze, C. Fink, T. Sulima, and I. Eisele, "Performance improvement in vertical surface tunneling transistors by a boron surface phase," *Jpn. J. Appl. Phys.*, vol. 40, no. 5A, pp. 3131–3136, 2001.

[123] S. Sedlmaier, J. Schulze, T. Sulima, C. Fink, C. Tolksdorf, A. Bayerstadler, I. Eisele, P.-F. Wang, K. Hilsenbeck, and W. Hansch, "Phonon assisted tunneling in gated p-i-n diodes," *Mat. Sci. & Eng.*, vol. B89, no. 1-3, pp. 116–119, 2002.

[124] C. Aydin, A. Zaslavsky, S. Luryi, S. Cristoloveanu, D. Mariolle, D. Fraboulet, and S. Deleonibus, "Lateral interband tunneling transistor in silicon-on-insulator," *Appl. Phys Lett.*, vol. 84, no. 10, pp. 1780–1782, 2004.

[125] J. Appenzeller, Y.-M. Lin, J. Knoch, and P. Avouris, "Band-to-Band Tunneling in Carbon Nanotube Field-Effect Transistors," *Physical Review Letters*, vol. 93, no. 19, pp. 1–4, 2004.

[126] J. Appenzeller, Y.-M. Lin, J. Knoch, Z. Chen, and P. Avouris, "Comparing Carbon Nanotube Transistors - The Ideal Choice: A Novel Tunneling Device Design," *IEEE Trans. Elec. Devices*, vol. 52, no. 12, pp. 2568–2576, 2005.

[127] K. Bhuwalka, M. Born, M. Schindler, M. Schmidt, T. Sulima, and I. Eisele, "P-Channel Tunnel Field-Effect Transistors down

to Sub-50nm Channel Lengths," *Jpn. J. Appl. Phys*, vol. 45, pp. 3106–3109, 2006.

[128] W. Choi, B. Park, J. Lee, and T. Liu, "Tunneling field-effect transistors (TFETs) with subthreshold swing (SS) less than 60mV/dec," *IEEE Electron Device Lett.*, vol. 28, no. 8, pp. 743–745, 2007.

[129] M. Björk, J. Knoch, H. Schmid, H. Riel, and W. Riess, "Silicon nanowire tunneling field-effect transistors," *Applied Physics Letters*, vol. 92, p. 193504, 2008.

[130] C. Sandow, J. Knoch, C. Urban, Q. Zhao, and S. Mantl, "Impact of electrostatics and doping concentration on the performance of silicon tunnel field-effect transistors," *Solid-State Electronics*, vol. 53, pp. 1126–1129, 2009.

[131] K. Moselund, H. Ghoneim, M. Björk, H. Schmid, S. Karg, E. Lörtscher, W. Riess, and H. Riel, "Comparison of VLS grown Si NW tunnel FETs with different gate stacks," in *Proc. ESSDERC*, 2009, pp. 448–451.

[132] F. Mayer, C. L. Royer, J. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B. Previtali, and S. Deleonibus, "Impact of SOI, $Si_{1-x}Ge_xOI$ and GeOI substrates on CMOS compatible Tunnel FET performance," in *2008 IEEE International Electron Devices Meeting*, vol. 4, 2008, pp. 1–5.

[133] D. Kazazis, P. Jannaty, A. Zaslavsky, C. L. Royer, C. Tabone, L. Clavelier, and S. Cristoloveanu, "Tunneling field-effect transistor with epitaxial junction in thin germanium-on-insulator," *Applied Physics Letters*, vol. 94, p. 263508, 2009.

[134] J. Knoch, S. Mantl, and J. Appenzeller, "Impact of the dimensionality on the performance of tunneling FETs: Bulk versus one-dimensional devices," *Solid-State Electronics*, vol. 51, pp. 572–578, 2007.

[135] S. Koswatta, D. Nikonov, and M. Lundstrom, "Computational study of carbon nanotube pin tunnel FETs," in *IEEE International Electron Devices Meeting*, vol. I, 2005, pp. 518–521.

[136] S. Koswatta, M. Lundstrom, and D. Nikonov, "Influence of phonon scattering on the performance of p-i-n band-to-band tunneling transistors," *Applied Physics Letters*, vol. 92, p. 43125, 2008.

[137] M. Born, K. Bhuwalka, M. Schindler, U. Aelein, M. Schmidt, T. Sulima, and I. Eisele, "Tunnel FET: A CMOS Device for high Temperature Applications," in *25th International Conference on Microelectronics*, vol. 1, 2006, pp. 124–127.

[138] A. Verhulst, W. Vandenberghe, K. Maex, and G. Groeseneken, "Tunnel field-effect transistor without gate-drain overlap," *Applied Physics Letters*, vol. 91, p. 053102, 2007.

[139] K. Boucart and A. Ionescu, "A new definition of threshold voltage in tunnel FETs," *Solid-State Electronics*, vol. 52, pp. 1318–1323, 2008.

[140] M. Schlosser, K. Buwalka, M. Sauter, T. Zilbauer, T. Sulima, and I. Eisele, "Fringing-Induced Drain Current Improvement in the Tunnel Field-Effect Transistor With High-κ Gate Dielectrics," *IEEE Transactions on Electron Devices*, vol. 56, no. 1, pp. 100–108, 2009.

[141] M. Vadizadeh and M. Fathipour, "Using Low-κ Oxide for Reduction of Leakage Current in Double Gate Tunnel FET," in *Proc. ULIS*, 2009, pp. 301–304.

[142] S. Koswatta, M. Lundstrom, and D. Nikonov, "Performance Comparison Between p-i-n Tunneling Transistors and Conventional MOSFETs," *IEEE Transactions on Electron Devices*, vol. 56, pp. 456–465, 2009.

[143] A. Verhulst, W. Vandenberghe, K. Maex, S. D. Gendt, M. Heyns, and G. Groeseneken, "Complementary Silicon-Based Heterostruc-

ture Tunnel-FETs With High Tunnel Rates," *IEEE Electron Device Letters*, vol. 29, pp. 1398–1401, 2008.

[144] E. Toh, G. Wang, L. Chan, G. Samudra, and Y. Yeo, "Device physics and guiding principles for the design of double-gate tunneling field-effect transistor with silicon-germanium source heterojunction," *Applied Physics Letters*, vol. 91, p. 243505, 2007.

[145] T. Krishnamohan, D. Kim, S. Raghunathan, and K. Saraswat, "Double-Gate Strained-Ge Heterostructure Tunneling FET (TFET) with record high drive currents and 60mV/dec subthreshold slope," in *2008 IEEE International Electron Devices Meeting*, vol. 67, 2008, pp. 1–3.

[146] A. Bowonder and P. Patel, "Low-voltage green transistor using ultra shallow junction and hetero-tunneling," in *Extended Abstracts - 2008 8th International Workshop on Junction Technology (IWJT'08)*, vol. 1, 2008, pp. 93–96.

[147] K. Boucart, W. Riess, and A. Ionescu, "Lateral Strain Profile as Key Technology Booster for All-Silicon Tunnel FETs," *IEEE Electron Device Letters*, vol. 30, pp. 656–658, 2009.

[148] W. Vandenberghe, B. Soree, W. Magnus, G. Groeseneken, and M. Fischetti, "Impact of field-induced quantum confinement in tunneling field-effect devices," *App. Phys. Lett.*, vol. 98, p. 143503, 2011.

[149] W. Vandenberghe, B. Soree, W. Magnus, G. Groeseneken, , A. Verhulst, and M. Fischetti, "Field induced quantum confinement in Indirect Semiconductors: Quantum mechanical and modified semiclassical model," in *SISPAD 2011*, 2011, pp. 271–274.

[150] A. Verhulst, B. Soree, D. Leonelli, W. Vandenberghe, and G. Groeseneken, "Modeling the single-gate, double-gate and gate-all-around tunnel field-effect transistor," *J. Appl. Phys.*, vol. 107, no. 2, pp. 024 518–1–024 518–8, 2010.

[151] W. Vandenberghe, A. Verhulst, G. Groeseneken, B. Soree, and W. Magnus, "Analytical model for a tunnel field-effect transistor," in *Proc. MELECON*, May 2008, pp. 923–928.

[152] C. Shen, S. Ong, C. Heng, G. Samudra, and Y. Yeo, "A variational approach to the two dimensional nonlinear Poisson's equation for the modeling of tunneling transistors," *IEEE Electron Device Lett.*, vol. 29, no. 11, pp. 1252–1255, 2008.

[153] C. Shen, L.-T. Yang, G. Samudra, and Y.-C. Yeo, "A new robust non-local algorithm for band-to-band tunneling simulation and its application to Tunnel-FET," *Solid-State Electronics*, vol. 57, pp. 23–30, 2011.

[154] Q. Zhang, W. Zhao, and A. Seabaugh, "Analytic Expression and Approach for Low Subthreshold-Swing Tunnel Transistors," in *Device Research Conference (DRC)*, June 2005.

[155] A. Chynoweth, W. Feldmann, and R. Logan, "Excess Tunnel Current in Silicon Esaki Junctions," *Phys. Rev.*, vol. 121, p. 684, 1961.

[156] S. Datta, *Electronic Transport in Mesoscopic Systems*. Cambridge University Press, Cambridge, 1995.

[157] J. Knoch and J. Appenzeller, "A novel concept for field-effect transistors - the tunneling carbon nanotube FET," in *Device Research Conference Digest, 2005. DRC '05. 63rd*, vol. 1, 20-22, 2005, pp. 153 – 156.

[158] A. Schenk, "Rigorous Theory and Simplified Model of the Band–to–Band Tunneling in Silicon," *Solid-State Electronics*, vol. 36, no. 1, pp. 19–34, 1993.

[159] J. Morgan and E. Kane, "Observation of Direct Tunneling in Germanium," *Phys. Rev. Lett.*, vol. 3, p. 466, 1959.

[160] L. Keldysh, "Behavior of non-metallic crystals in strong electric fields," *Sov. Phys.-JEPT*, vol. 33, no. 4, pp. 763–770, 1958.

[161] S. Luryi, "Quantum capacitance devices," *Appl. Phys. Lett.*, vol. 52, pp. 501–503, 1988.

[162] A. Rahman, J. Guo, S. Datta, and M. Lundstrom, "Theory of Ballistic Nanotransistors," *IEEE Transactions on Electron Devices*, vol. 50, no. 9, pp. 1853–1864, 2003.

[163] W. Lee and W. Y. Choi, "Influence of Inversion Layer on Tunneling Field-Effect transistors," *IEEE Electron Device Letters*, vol. 32, no. 9, pp. 1191–1193, 2011.

[164] K. Bhuwalka, J. Schulze, and I. Eisele, "A simulation approach to optmize the electrical parameters of a vertical tunnel FET," *IEEE Trans. Electron Devices*, vol. 52, no. 7, pp. 1541–1547, 2005.

[165] W. Vandenberghe, A. Verhulst, G. Groeseneken, B. Soree, and W. Magnus, "Analytical Model for Point and Line Tunneling in a Tunnel Field-Effect Transistor," in *SISPAD 2008*, 2008, pp. 137–140.

[166] M. Fischetti, T. O'Regan, S. Narayanan, C. Sachs, S. Jin, J. Kim, and Y. Zhang, "Theoretical Study of Some Physical Aspects of Electronic Transport in nMOSFETs at the 10-nm Gate-Length," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2116–2136, 2007.

[167] L. Keldysh, "Influence of the lattice vibrations of a crystal on the production of electron-hole pairs in strong electric fields," *Sov. Phys.-JEPT*, vol. 34, no. 4, pp. 665–668, 1958.

[168] G. Hurkx, D. Klaassen, and M. Knuvers, "A new recombination model for device simulation including tunneling," *Electron Devices, IEEE Transactions on*, vol. 39, no. 2, pp. 331 –338, Feb 1992.

[169] M. Ancona and H. Tiersten, "Macroscopic physics of the silicon inversion layer," *Phys. Rev. B*, vol. 35, pp. 7959–7965, 1987.

[170] M. Ancona and G. Iafrate, "Quantum correction to the equation of state of an electron gas in a semiconductor," *Phys. Rev. B*, vol. 39, no. 13, pp. 9536–9540, 1989.

[171] M. Luisier and G. Klimeck, "Atomistic Full–Band Design Study of InAs Band–to–Band Tunneling Field–Effect Transistors," *IEEE Electron Device Letters*, vol. 30, no. 6, pp. 602–604, 2009.

[172] R. Booth, M. White, H. Wong, and T. Krutsick, "The effect of channel implants on MOS transistor characterization," *IEEE Trans. Electron Devices*, vol. ED-34, pp. 2501–2509, 1987.