

# POLITECNICO DI MILANO

Scuola di Ingegneria Industriale e dell'Informazione  
Corso di Laurea Specialistica in Ingegneria Elettronica  
Indirizzo Elettronica Fisica



## Studio dell'ampiezza dell'RTN nei dispositivi MOS di nanoscala mediante sIFM

Relatore: Prof. Christian Monzio Compagnoni

Correlatore: Ing. Andrea Ghetti

Tesi di Laurea di:  
Giulio Torrente  
Matr. 766457

Anno Accademico 2012-2013

*Alla mia famiglia*

# Indice

<b>Introduzione</b>	<b>4</b>
<b>1 Memorie Flash e Variability</b>	<b>7</b>
1.1 Memorie Flash . . . . .	7
1.1.1 Struttura e funzionamento di una cella Flash . . . . .	7
1.1.2 Flash Multilivello . . . . .	12
1.2 Variability . . . . .	16
1.2.1 Random Dopant Fluctuation (RDF) . . . . .	16
1.2.2 Random Telegraph Noise (RTN) . . . . .	23
1.3 Riduzione delle fluttuazioni . . . . .	36
1.3.1 Drogaggi disuniformi . . . . .	37
1.3.2 Celle con bordi arrotondati . . . . .	42
<b>2 Metodologie di simulazione</b>	<b>45</b>
2.1 Introduzione . . . . .	45
2.2 Metodo Monte Carlo . . . . .	46
2.2.1 Il Modello Drift-Diffusion . . . . .	46
2.2.2 Il Metodo . . . . .	47
2.3 Impedance Field Method . . . . .	51
2.3.1 Metodo di risoluzione della funzione di Green . . . . .	53
2.3.2 Funzione di Green nel metodo IFM . . . . .	59
2.3.3 “Noise-Like” IFM . . . . .	64
2.3.4 “Statistical” IFM . . . . .	66
2.3.5 sIFM per lo studio dell’RTN . . . . .	72
2.3.6 IFM in Sdevice . . . . .	78

<b>3</b>	<b>Risultati</b>	<b>81</b>
3.1	RDF . . . . .	81
3.1.1	Calcolo delle dispersioni . . . . .	81
3.1.2	Risultati . . . . .	83
3.2	RTN . . . . .	86
3.2.1	Drogaggio uniforme . . . . .	86
3.2.2	Drogaggi disuniformi lungo L e W . . . . .	99
3.2.3	Drogaggi disuniformi verticalmente . . . . .	106
3.2.4	Errori di linearizzazione dei metodi sIFM . . . . .	115
3.2.5	Verifiche sulla metodologia . . . . .	131
3.2.6	Osservazioni conclusive sui metodi . . . . .	137
<b>4</b>	<b>Applicazione di modelli analitici su un MOS</b>	<b>139</b>
4.1	Il dispositivo e le fluttuazioni “anomale” . . . . .	139
4.2	Modello analitico per il roll-off del MOS . . . . .	145
4.2.1	Il modello . . . . .	145
4.2.2	Tecniche di soluzione . . . . .	146
4.2.3	Tensione di soglia e dispersione . . . . .	149
4.3	Risultati del modello . . . . .	151
	<b>Conclusioni</b>	<b>153</b>

# Introduzione

I notevoli investimenti e sforzi nella ricerca sui dispositivi elettronici hanno portato un forte sviluppo nell'elettronica integrata, in particolare nell'ambito digitale e per quanto riguarda le tecnologie basate sul transistor MOSFET. Tale progresso scientifico è sospinto dallo scopo di aumentare la densità dei componenti in un singolo chip, tanto che la lunghezza di un device ha raggiunto ormai le dimensioni di poche decine di nanometri. L'aumento della densità di componenti per chip, che segue la nota legge di Moore (Fig.1), ossia raddoppia all'incirca ogni 24 mesi, permette di ottenere minori potenze dissipate e maggiore velocità di elaborazione e di conseguenza permette l'abbattimento dei costi per funzione (cost-per-function), che si riducono di poco meno del 30% all'anno. Tale progresso ha raggiunto un punto in cui risulta difficile immaginare il mantenimento dei trend storici di scaling negli anni a venire, dal momento che l'avanzamento nel range decananometrico e nanometrico porta con sé problemi sempre maggiori e tali per cui si inizia a pensare a dispositivi post-CMOS, siano essi evoluzioni della medesima tecnologia oppure soluzioni totalmente nuove.

I dispositivi di ultima generazione possono essere considerati pertanto "atomici", nel senso che contengono qualche manciata di atomi. Per questo nella loro progettazione è necessario tenere conto della natura discreta e granulare della materia. Infatti le fluttuazioni statistiche delle caratteristiche elettriche di dispositivi nominalmente identici riscontrate sperimentalmente sono riconducibili, per la maggior parte, proprio alla natura atomistica del drogaggio, in particolare per quanto riguarda i fenomeni di RTN (*Random Telegraph Noise*) ed RDF (*Random Dopant Fluctuation*). Data la complessità dei fenomeni fisici in questione, difficilmente possono essere sviluppati modelli analitici affidabili in grado di predire la loro distribuzione statistica, e dunque è reso necessario uno studio basato su simulazioni numeriche 3D e, dato il larghissimo numero di componenti per chip fabbricati, è fondamentale andare ad esplorare le bassissime probabilità.

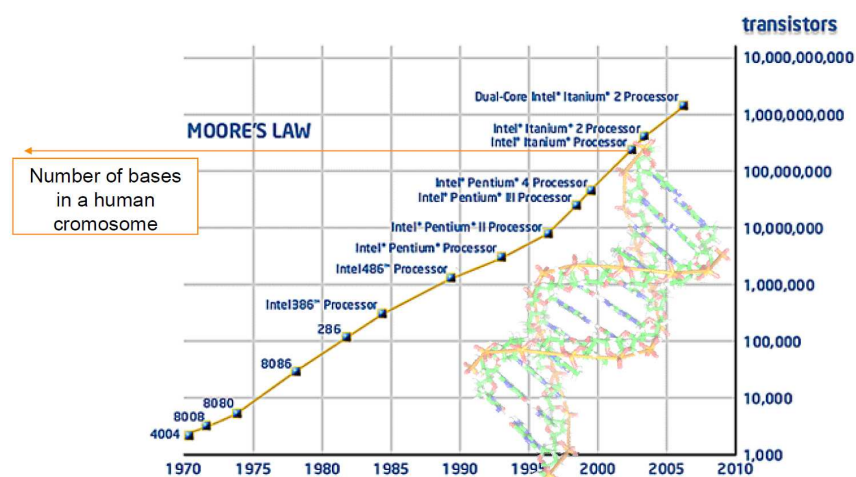


Figura 1: Legge di Moore. I componenti (transistor) per singolo chip raddoppiano ogni due anni.

Da un punto di vista computazionale, il metodo *Monte Carlo* (MC), che cerca di riprodurre le differenze microscopiche nei dispositivi simulati, è diffusamente utilizzato per l'analisi della variabilità statistica in relazione ai fenomeni di RTN ed RDF. Esso risulta molto preciso ma estremamente lento; ciò ha spinto la ricerca verso metodologie alternative con meno sforzo computazionale. Il presente lavoro di tesi, svolto presso la sede di Agrate Brianza di Micron nell'ambito del gruppo TCAD (*Technology Computer Aided Design*), si propone di analizzare questi fenomeni di *variability* con un approccio alternativo al MC: l'*Impedance Field Method* (IFM). Nel corso di questo elaborato spiegherò come utilizzare opportunamente l'IFM, e i pregi e i difetti che mostra rispetto al metodo di riferimento MC. L'attenzione sarà focalizzata principalmente sulle memorie Flash a floating-gate, dato che sono più soggette alle fluttuazioni statistiche ed hanno subito uno scaling delle dimensioni particolarmente aggressivo negli ultimi anni, e sulle dispersioni statistiche riguardanti il fenomeno di RTN, maggior ostacolo per lo scaling tecnologico delle Flash.

Nel capitolo 1 vengono sommariamente descritte le memorie Flash a floating-gate, mostrando le basi fisiche su cui si costruisce il funzionamento delle singole celle e le difficoltà di scaling dovute alle fluttuazioni statistiche. In tal capitolo descrivo inoltre le maggiori cause di dispersione dei parametri elettrici delle celle, in particolare l'RTN e l'RDF.

Nel capitolo 2 descrivo le metodologie di simulazione per l'RDF e l'RTN: il metodo *Monte Carlo* (MC) e l' *Impedance Field Method* (IFM). Il primo, utilizzato come riferimento in questa tesi, prevede la simulazione attraverso il modello *Drift Diffusion* di ogni singolo campione della statistica; il secondo, invece, è un metodo approssimato per lo studio delle fluttuazioni dovute al fenomeno dell'RDF. Lo *statistical-IFM* (sIFM), in particolare, è una variante dell'IFM implementata per la ricostruzione in maniera veloce delle transcaratteristiche  $I_D - V_{CG}$  dei singoli dispositivi della popolazione statistica. In questo elaborato di tesi ho valutato la possibilità di estendere l'uso di tale metodo per descrivere la distribuzione statistica dell'RTN. Di tal metodo ho proposto pertanto due implementazioni per lo studio dell'RTN: lo *sIFM-doping* e lo *sIFM-trap*.

Nel capitolo 3, dopo aver mostrato brevemente alcuni risultati per quanto riguarda l'RDF, ho riportato una gran quantità di confronti per l'RTN tra i metodi sIFM e il MC. In tal capitolo ho cercato di capire quando e come i metodi approssimati risultano dare risultati i cui discostamenti dai rispettivi MC, presi come riferimenti, possono essere considerati trascurabili. I confronti vengono fatti soprattutto sulle distribuzioni delle ampiezze dell'RTN per la stessa statistica di campioni. Dopo tali analisi sono arrivato alla conclusione che entrambe le metodologie proposte per lo studio dell'RTN offrono un pratico e conveniente compromesso tra accuratezza e sforzo computazionale per la ricostruzione della distribuzione dell'ampiezza dell'RTN. Per quanto riguarda invece il singolo campione della statistica i metodi faticano ad approssimare con sufficiente precisione l'ampiezza dello shift di soglia e le transcaratteristiche rispetto al riferimento MC.

Nel capitolo 4, infine, mostro lo studio della dispersione statistica della tensione di soglia su transistori MOS, per spiegare le particolari caratteristiche riscontrate sperimentalmente.

Per quanto riguarda i risultati della tesi in merito all'analisi statistica dell'ampiezza dell'RTN, sono state scritte due pubblicazioni scientifiche: [25] e [26].

# Capitolo 1

## Memorie Flash e Variability

### 1.1 Memorie Flash

Le memorie Flash sono le memorie a stato solido non volatili che hanno segnato l'evoluzione degli ultimi anni grazie alla possibilità di essere programmate e cancellate elettricamente. Si rivelano quindi estremamente flessibili e versatili e per questo si sono affermate in tutti i tradizionali segmenti di mercato quali telecomunicazioni, Personal Computer, automotive e industriale. Il loro sviluppo però, si sta portando ai limiti della scalabilità, poiché più di ogni altro dispositivo richiedono una transcaratteristica ben definita. Per questo motivo lo studio delle fluttuazioni statistiche dei parametri elettrici dei dispositivi diventa fondamentale al fine di prevedere e possibilmente abbattere tali dispersioni con l'avanzare della tecnologia.

#### 1.1.1 Struttura e funzionamento di una cella Flash

La struttura di una cella Flash a Floating Gate (Fig 1.1a) è assolutamente analoga a quella di un transistor MOS planare classico, con la differenza che sono presenti due gate: il Control Gate (CG) e il Floating Gate (FG). Il primo è direttamente contatto con il circuito esterno, mentre il secondo risulta elettricamente isolato (flottante), in quanto completamente circondato da dielettrico (tipicamente biossido di silicio), ed è ubicato tra il canale ed il Control Gate.



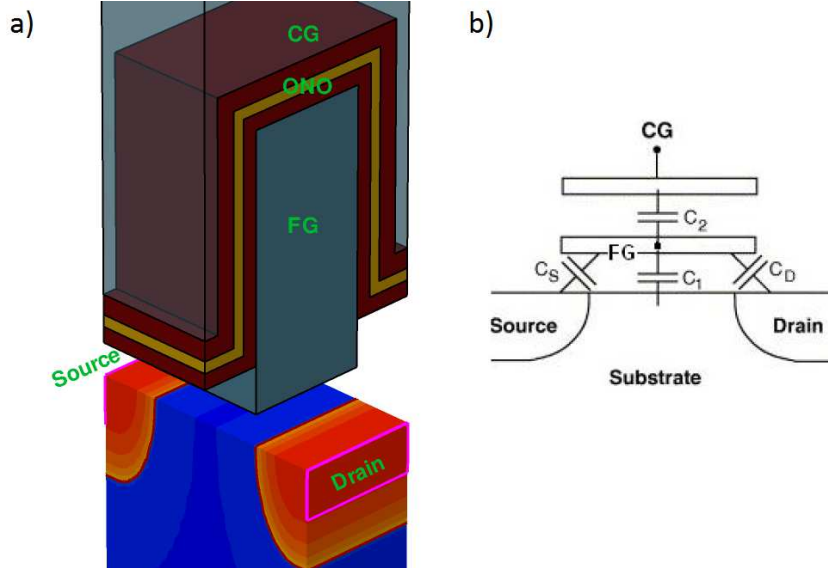


Figura 1.1: a) struttura di una cella a floating gate e b) accoppiamenti capacitivi fra control, floating gate, source, drain e substrato.

Control gate e floating gate controllano la conduttanza di canale e, come mostra lo schema in Figura 1.1b), sono legati capacitivamente fra loro e con gli altri nodi del circuito attraverso le capacità  $C_S$ ,  $C_D$ ,  $C_1$  e  $C_2$ . Se cortocircuitiamo bulk e source a massa possiamo quantitativamente esprimere il potenziale al floating gate come:

$$V_{FS} = \frac{C_2}{C_T} V_{GS} + \frac{C_D}{C_T} V_{DS} + \frac{Q}{C_T} \quad (1.1)$$

dove  $C_T = C_2 + C_S + C_1 + C_D$  è la capacità totale vista dal nodo di floating gate. Avendo definito il coefficiente di accoppiamento  $\alpha_G = C_2/C_T$ , possiamo definire la tensione di soglia al floating gate  $V_{T_{FG}}$  come la tensione per la quale il canale del dispositivo viene invertito avendo posto  $V_{DS} = 0$ . La tensione rispettiva al control gate risulta pertanto:

$$V_{T_{GS}} = \frac{1}{\alpha_G} V_{T_{FG}} - \frac{Q}{C_2} \quad (1.2)$$

L'equazione 1.2 esprime la tensione di soglia al control gate come un parametro *technology dependent*, che chiamiamo  $V_{T_0}$ , sommato ad un termine dipendente dalla carica immagazzinata nel floating gate:

$$V_T = V_{T_0} - \frac{Q}{C_2} \quad (1.3)$$

Poiché è possibile controllare la carica nel floating gate inserendola (mediante elettroni caldi) o togliendola (mediante effetto tunnel), riusciamo ad avere due stati ben distinti del dispositivo: il device si dice *programmato* (stato 0) quando nel floating gate è presente una carica  $Q$  ben definita, e *cancellato* (stato 1) quando invece non è presente alcuna carica. La presenza di una carica negativa nel nodo di FG fa sì che la transcaratteristica  $I_d - V_{gs}$  subisca una traslazione rigida lungo l'asse delle ascisse ( $V_{gs}$ ) di quantità  $\Delta V_T = Q/C_2$  come in figura:

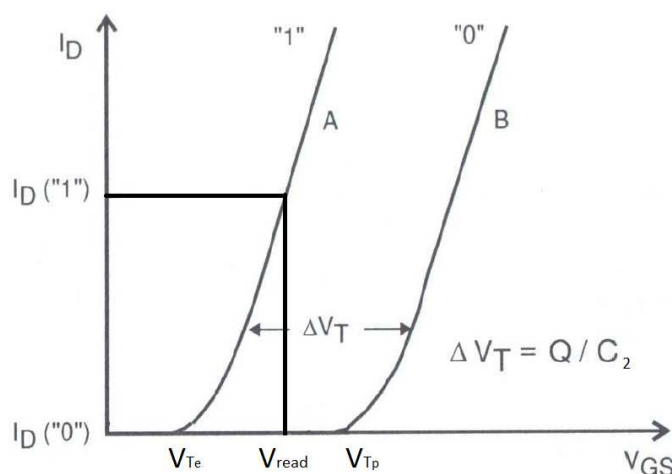


Figura 1.2: transcaratteristiche per lo stato 0 e lo stato 1 della cella di memoria

I due stati del dispositivo presentano perciò due diverse soglie, ovviamente la carica negativa al floating gate fa sì che la soglia  $V_{T_p}$  per il dispositivo programmato sia più alta in quanto sarà più difficile invertire il canale in quel caso:

$$V_{T_e} = V_{T_0} \quad (1.4)$$

$$V_{T_p} = V_{T_0} - \frac{Q}{C_2} \quad (1.5)$$

Per la lettura del dato è sufficiente polarizzare il control gate con una  $V_{read}$  tale che  $V_{T_e} < V_{read} < V_{T_p}$  e leggere la corrente di drain  $I_d$ . Se la  $V_{read}$  è scelta bene, in un caso leggeremo una certa  $I_d > 0$ , nell'altro  $I_d = 0$ , riuscendo a discriminare bene i due stati.

Le tre operazioni basilari che una memoria non volatile deve poter compiere sono la lettura, la scrittura e la cancellazione del dato. La lettura è quella più frequente e per fortuna anche la più semplice come abbiamo visto. Resta da capire come scrivere e cancellare il dato.

Per la scrittura si sfrutta l'iniezione di portatori caldi o l'effetto tunnel a seconda se consideriamo rispettivamente le memorie NOR o NAND. Consideriamo il primo caso: il CG viene portato ad una tensione molto alta (per esempio 12V), diminuendo la barriera energetica tra bulk e floating gate; contemporaneamente anche il drain viene alzato (ad esempio a 6V), in questo modo la velocità dei portatori aumenta e gli elettroni caldi, ovvero con energia uguale o più alta della barriera semiconduttore-isolante, saranno iniettati dal substrato lato drain al floating gate, dove rimarranno intrappolati 1.3.

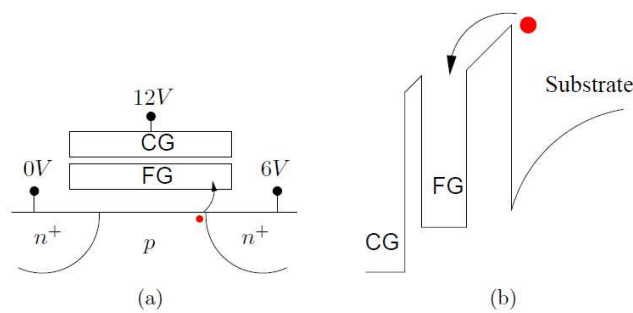


Figura 1.3: Polarizzazione del dispositivo(a) e diagramma a bande(b) durante la fase di scrittura

Iniettando elettroni il potenziale di FG diventa sempre più negativo e tende di conseguenza a saturare l'iniezione di carica, oltre ovviamente a raggiungere l'effetto desiderato di modificare la tensione di soglia della cella.

Per la programmazione si cerca di assicurare che lo shift di soglia indotto sia fisso ed uguale per tutte le celle. La tensione di drain e la lunghezza di canale sono i due vincoli che maggiormente influenzano l'efficienza della programmazione, poiché agiscono direttamente sul campo trasversale che genera gli elettroni caldi. Al fine di ottenere la corretta programmazione, risulta necessaria una verifica, che viene compiuta tramite la lettura ed il confronto del dato con un riferimento memorizzato in un apposito registro. Se il confronto restituisce un'uguaglianza a meno di un piccolo errore la programmazione della cella è terminata, altrimenti risulta necessario un'ulteriore impulso di program. Questo avviene finché lo scarto dal riferimento è piccolo o si superano le iterazioni massime consentite.

Per la cancellazione si sfrutta l'effetto tunnel di Fowler-Nordheim. Per richiamare gli elettroni dal FG al substrato si tiene alto il potenziale nel canale polarizzando ad una tensione tipicamente di 8V il source ed il drain e lasciando il bulk flottante (che si porta ad 8V anch'esso). Viceversa il control gate si polarizza ad una tensione bassa, per esempio -8V. Il forte campo elettrico che si genera tra substrato e FG, fa sì che la barriera energetica tra FG e canale assuma una forma triangolare ed una larghezza piuttosto stretta. Ciò rende possibile l'effetto tunnel che porta gli elettroni a superare tale barriera e tornare così nel substrato. Questa corrente di tunnel scarica il floating gate portando la tensione di soglia ad assumere valori più bassi.

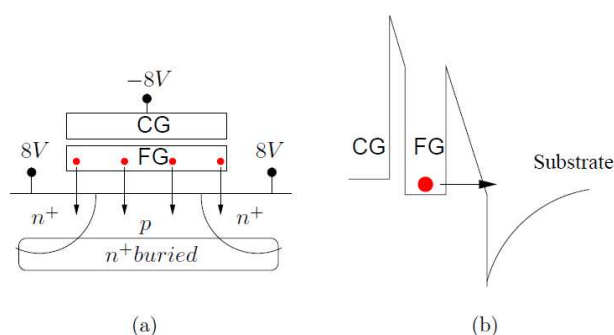


Figura 1.4: Polarizzazione del dispositivo(a) e diagramma a bande(b) durante la fase di cancellazione

È possibile, però, che la tensione di soglia raggiunta dopo la cancellazione non sia quella desiderata e, analogamente a quanto si fa per la programmazione, risulta spesso necessaria una procedura di verifica e ricancellazione, fino a quando non viene raggiunta la tensione di soglia voluta o non viene superato il numero massimo di iterazioni consentite.

### 1.1.2 Flash Multilivello

Abbiamo visto come grazie alla programmazione della cella possiamo introdurre carica all'interno della floating gate. Questa carica, d'altro canto, non è fissata a priori. Essa è infatti ingegnerizzabile, ovvero possiamo mettere più o meno carica a seconda dell'ampiezza e della durata degli impulsi di program. Questo ci dà la possibilità di avere diversi livelli per la stessa cella e perciò più bit. Ovviamente se essa deve memorizzare  $N$  bit, è necessario avere  $2^N$  livelli diversi di programmazione e questo impone dei vincoli più severi sulla larghezza delle distribuzioni e sulla loro separazione (Fig.1.5). In particolare è necessario che le distribuzioni delle soglie dei livelli centrali ("00" e "10" in Fig.1.5) siano molto strette. Questi vincoli hanno portato ad una maggior importanza delle fluttuazioni statistiche della soglia che impattano sia l'affidabilità che la progettazione del dispositivo di memoria.

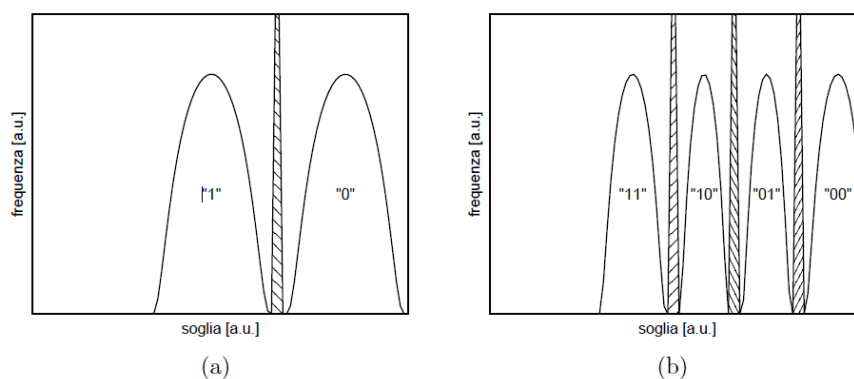


Figura 1.5: Esempio di distribuzione delle soglie per una flash a 2 (a) e a 4 (b) livelli. Tratteggiate le zone in cui i valori di soglia non sono permessi.

## Variability e Scaling

Le fluttuazioni statistiche della tensione di soglia possono essere sostanzialmente di due tipi: statiche o dinamiche. Le prime sono causate da tantissimi fattori, come le variazioni del processo di fabbricazione, le fluttuazioni della carica nell'ossido, le variazioni intrinseche delle dimensioni del gate introdotte dal Line Edge Roughness, le fluttuazioni geometriche in generale, le variazioni della WorkFunction all'interfaccia con l'ossido, ma soprattutto le fluttuazioni dei droganti in numero e posizione (RDF, *Random Dopant Fluctuation*). Per ridurre l'impatto dell'RDF sulle caratteristiche elettriche si adottano determinate geometrie o particolari profili di drogaggio, ma non si può eliminare completamente. Per essere insensibili a queste fluttuazioni è necessario introdurre algoritmi che realizzano operazioni di tipo bit-by-bit program-and-verify. Dopo un impulso di program, si confrontano tutte le celle programmate con i *verify levels*: quelle che risultano fuori specifica necessitano di un ulteriore impulso di programmazione. A differenza della cella a singolo bit precedentemente descritta, le operazioni di programmazione e verifica vengono ripetute sicuramente più volte in quanto la specifica sulla larghezza della distribuzione è più stringente. Questo fa sì che la cella multilivello sia più lenta per quanto riguarda la programmazione. Ci deve essere, infatti, un circuito che ciclicamente scrive e controlla che il valore scritto sia corretto con elevata precisione, questo porta ad una complicazione del sistema oltre che alla presenza di più tensioni di riferimento.

Oltre a questi algoritmi di correzione, nelle memorie sono presenti dei sistemi in grado di rintracciare eventuali errori contenuti nell'informazione memorizzata e dei meccanismi capaci di correggerli attraverso dei codici di correzione. Questi sistemi vengono chiamati ECC (*Error Correction Code*). Facciamo un esempio molto semplice: codifica delle cifre decimali 0-9 con 5 bit. Essendo che il numero totale delle informazioni da codificare è 10, basterebbero 4 bit; il quinto è pertanto ridondante e serve per l'individuazione dell'errore. Esso fa in modo che nella codifica binaria ci siano sempre tre 0 e due 1 e perciò se si verifica un errore il codice se ne accorge. Esso, però, risulta non accorgersi dell'errore nel momento in cui ci siano 2 bit errati, e comunque non mi dice qual' è il bit incriminato nella parola. Gli ECC realmente utilizzati nelle memorie, invece, permettono di rintracciare gli errori anche quando coinvolgono più bit contemporaneamente

della stessa parola, e di correggerli. Il codice più frequentemente usato è l'ECC di Hamming, il quale permette, con 8 bit aggiuntivi, di correggere errori su di un singolo bit per quadword (64 bit) e di rilevare errori doppi.

Più complicati risultano tutti questi algoritmi e codici di correzione più sarà lenta la memoria. È pertanto necessaria una stima a priori delle dispersioni statiche delle tensioni di soglia, in modo da progettarli non eccessivamente complicati ed in grado di correggere tutti gli errori, a meno di una piccolissima percentuale. Per quantificare questa stima sono necessarie simulazioni numeriche 3D su tali dispositivi.

Nonostante siano più lente, abbiano specifiche più stringenti per le distribuzioni delle tensioni di soglia ed abbiano una circuiteria più complicata, le celle multilivello (MLC, *Multi Level Cells*) hanno il forte vantaggio di poter essere realizzate, a parità di numero di bit, con una tecnologia sicuramente più rilassata e perciò meglio fabbricabili e meno soggette a fluttuazioni di ogni genere. La scalabilità per le celle Flash è infatti molto critica. Esse vanno incontro a problemi e limiti non presenti nelle tecnologie CMOS standard, quali la non scalabilità delle tensioni di program ed erase: esse infatti dipendono da parametri fisici come la barriera energetica (3.2 eV fra silicio e ossido) per gli elettroni caldi, o i 10-11 MV/cm del campo elettrico necessario per ottenere la corrente di tunnel Fowler-Nordheim appropriata. Inoltre, rispetto ad un normale MOSFET planare, la cella Flash amplifica le variazioni della tensione di soglia a causa dell'accoppiamento capacitivo:  $\Delta V_{TCG} = \alpha_G^{-1} \Delta V_{TFG}$  dove  $\alpha_G \approx 0.6$ . Si preferisce, perciò, aumentare i bit per cella piuttosto che scalare il dispositivo ulteriormente.

Questo discorso non è del tutto vero se consideriamo le fluttuazioni statistiche dinamiche della soglia. Esse sono attribuibili al *Random Telegraph Noise* (RTN), ovvero alla cattura e rilascio degli elettroni da parte di trappole situate all'interfaccia *Si/Ox*. Questo provoca una variazione di conducibilità e perciò di corrente causando di fatto uno shift di soglia  $\Delta V_T$  nella transcaratteristica, che allarga le distribuzioni delle soglie (Fig.1.7). Se la cattura avviene dopo la fase di program-and-verification e se la  $\Delta V_T$  è abbastanza grande, può accadere che la lettura del dato sia errata. In altre parole: se la  $\Delta V_T$  indotta dall'RTN risulta simile alla differenza delle soglie ideali, il dato può essere corrotto e non è possibile

correggerlo. Analogamente a quanto accade per l'RDF, per l'RTN si adottano particolari profili di drogaggio e forme della cella per ridurre il suo impatto a livello elettrico. Lo shift di soglia dinamico causato dall'RTN ha ovviamente un impatto maggiore nelle MLC in quanto i verify levels sono più vicini. Infatti lo scaling della cella Flash nel range decananometrico ha mostrato che le fluttuazioni della tensione di soglia causate dal Random Telegraph Signal costituiscono uno dei problemi principali dello scaling delle memorie Flash. Risulta fondamentale, pertanto, una previsione della dispersione della  $\Delta V_T$  indotta dall'RTN attraverso simulazioni numeriche 3D della cella. [8] [9] [10]

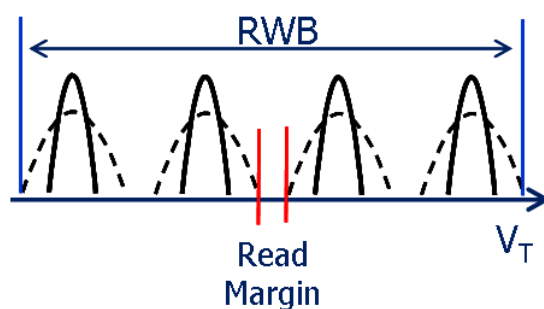


Figura 1.6: Effetto dell'RDF sulle soglie della cella multilivello

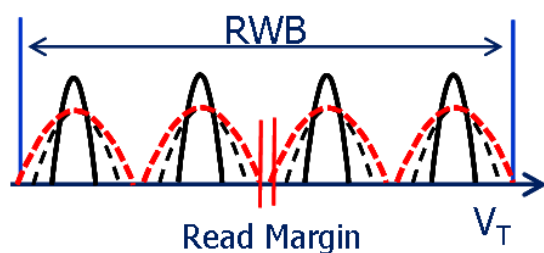


Figura 1.7: Effetto dell'RTN sulle soglie della cella multilivello



## 1.2 Variability

Come abbiamo visto, risulta necessario uno studio approfondito sulla variabilità della tensione di soglia per la cella Flash. I fenomeni fisici maggiormente influenti a riguardo sono l'RDF per quanto riguarda le fluttuazioni statiche e l'RTN per quelle dinamiche. In questo capitolo descrivo tali fenomeni mostrando anche alcuni semplici modelli analitici a riguardo.

### 1.2.1 Random Dopant Fluctuation (RDF)

Il drogaggio ideale-continuo è un'astrazione sempre meno valida allo scalare dei dispositivi. Nella realtà, infatti, il drogaggio è atomistico, ovvero è composto da atomi ionizzati inseriti all'interno del reticolo cristallino in maniera sostituzionale. A seconda del numero di droganti all'interno del dispositivo e della loro posizione, la transcaratteristica risulterà diversa. Dispositivi perciò nominalmente identici, a causa dell'RDF (*Random Dopant Fluctuation*), risultano avere transcaratteristiche “shiftate” tra loro e perciò soglie diverse. Questo sparpagliamento delle caratteristiche elettriche si può quantificare valutando la dispersione della tensione di soglia  $\sigma_{V_T}$ .

Il numero di droganti presenti nel dispositivo segue la statistica di Poisson. Inoltre la conduttività di canale risente fortemente dei fenomeni percolativi della corrente, che verranno spiegati ed analizzati nella prossima sezione in relazione all'RTN. Ciò che determina maggiormente la variabilità della soglia, infatti, è la probabilità che la posizione degli atomi droganti permetta la formazione di un percorso di corrente tra source e drain. Regioni con molti atomi droganti saranno più difficili da invertire, perciò avranno una soglia locale maggiore; mentre regioni sprovviste di droganti avranno al contrario soglie locali minori comportandosi come dispositivi equivalenti a basso drogaggio. Si verifica, inoltre, che mediamente i dispositivi hanno una tensione di soglia minore rispetto alla  $V_T$  di quello ideale, in quanto la formazione di alcuni percorsi percolativi permette alla cella di “accendersi” prima in media. Questi fenomeni di RDF risultano rilevanti per dispositivi scalati ( $L_{ch} < 100nm$ ), mentre per dispositivi “rilassati” risultano avere poco impatto e si può considerare la conduzione del canale uniforme.

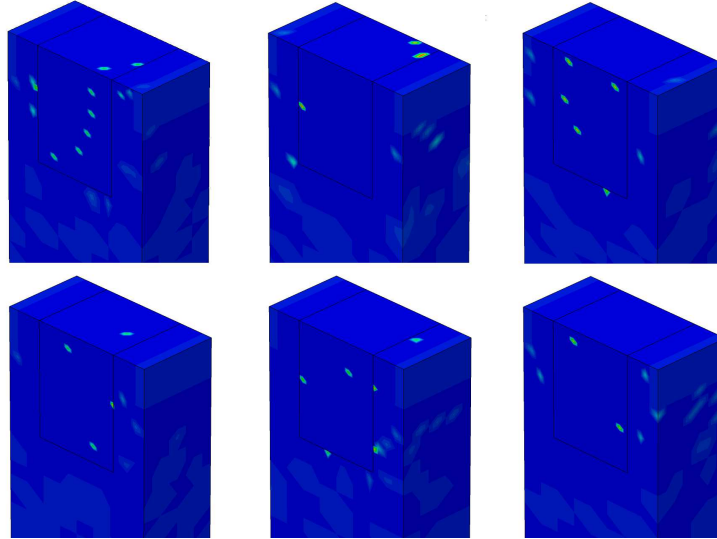


Figura 1.8: Esempio di sei configurazioni atomistiche di drogaggio di boro per una cella floating gate avente profilo ideale accettore costante.

### Modelli analitici

Per il fenomeno dell'RDF esistono molti modelli analitici per un transistor in generale. Essi si differenziano a seconda delle approssimazioni fatte e del dispositivo a cui viene applicato. Nel mio elaborato di tesi propongo tre diversi modelli: il modello di Stolk, quello di Takeuchi ed il modello di Asenov. I primi due sono modelli analitici veri e propri in quanto vengono direttamente dall'equazione di Gauss, mentre il terzo è un modello empirico ottenuto da simulazioni su dispositivi a canale molto corto. Il modello di Stolk e quello di Takeuchi hanno una radice comune che qui in seguito descrivo.

Consideriamo innanzitutto le equazioni mediante le quali si trovano campo elettrico  $E(x)$  e potenziale superficiale  $\phi_s$  per una zona svuotata:

$$E(x) = \frac{q}{\epsilon_{Si}} \int_x^{W_d} N(x) dx \quad (1.6)$$

$$\phi_s = \frac{q}{\epsilon_{Si}} \int_0^{W_d} \int_x^{W_d} N(x') dx' dx = \frac{q}{\epsilon_{Si}} \int_0^{W_d} xN(x) dx \quad (1.7)$$

dove  $W_d$  è la larghezza della zona svuotata ed  $N(x)$  è il profilo della concentrazione di doping.

Consideriamo ora il caso di un drogaggio non uniforme, per esempio con un profilo a scalino come in Fig.1.9

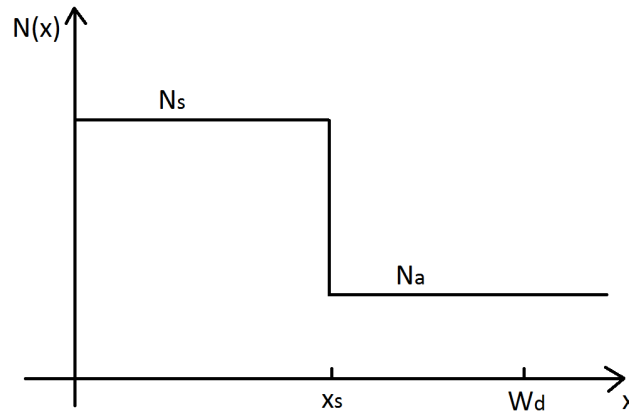


Figura 1.9: profilo di drogaggio a scalino

Integrando l'equazione 1.7 per il profilo di drogaggio in questione otteniamo la 1.8:

$$\phi_s = \frac{q}{\epsilon_{Si}} N_s x_s^2 + \frac{q N_A}{2\epsilon_{Si}} (W_d^2 - x_s^2) \quad (1.8)$$

$$W_d = \sqrt{\frac{2\epsilon_{Si}}{q N_A} \left( \phi_s - \frac{q(N_s - N_A)x_s^2}{2\epsilon_{Si}} \right)} \quad (1.9)$$

L'equazione 1.9 non è nient'altro che la 1.8 esplicitando la larghezza della depletion-layer. Da questa equazione capiamo bene come rispetto al caso di doping uniforme uguale ad  $N_A$ , abbiamo un aumento di carica per unità di area pari a  $q(N_s - N_A)x_s$  e contemporaneamente una diminuzione del potenziale superficiale di  $\frac{q(N_s - N_A)x_s^2}{2\epsilon_{Si}}$ . Ricordandoci l'espressione della tensione di gate

$$V_{gs} = V_{fb} + \phi_s + \frac{\sqrt{2\epsilon_{Si}qN_A\phi_s}}{C_{ox}} \quad (1.10)$$

e che a soglia il potenziale superficiale  $\phi_s = 2\phi_B = 2\frac{kT}{q}\ln(N_a/n_i)$ , calcoliamo la tensione di soglia con il profilo di drogaggio in questione:

$$V_T = V_{fb} + 2\phi_B + \frac{1}{C_{ox}} \sqrt{2\epsilon_{Si}qN_A \left( 2\phi_B - \frac{q(N_s - N_A)x_s^2}{2\epsilon_{Si}} \right)} + \frac{q(N_s - N_A)x_s}{C_{ox}} \quad (1.11)$$

Questo risultato può essere in un certo senso generalizzato ad altri profili di doping. L'aggiunta di densità di drogaggio pari a  $(N_s - N_A)$  su  $(0, x_s)$  equivale ad avere una delta function di drogaggio con dose pari a  $D_I = (N_s - N_A)x_s$  centrata in  $x_c = x_s/2$ . Questo è vero in quanto entrambi gli integrali 1.6 e 1.7 su  $(0, W_d)$  non cambiano. Sostituendo otteniamo:

$$V_T = V_{fb} + 2\phi_B + \frac{1}{C_{ox}} \sqrt{2\epsilon_{Si}qN_A \left( 2\phi_B - \frac{qD_I x_c}{\epsilon_{Si}} \right)} + \frac{qD_I}{C_{ox}} \quad (1.12)$$

La fluttuazione di doping  $\sigma_{dN}$  in  $(x,y,z)$  può essere pensata come un uniforme impianto (rispetto a W ed L) di drogaggio a delta  $\delta D_I$  che si sovrappone ad un background di drogaggio costante pari ad  $N_A$ . Per tal motivo differenziamo la 1.12 rispetto alla dose:

$$\begin{aligned} \delta V_T &= \frac{q\delta D_I}{C_{ox}} \left( 1 - \frac{x}{2\epsilon_{Si}} \sqrt{\frac{2\epsilon_{Si}qN_A}{2\phi_B - \frac{qD_I x}{\epsilon_{Si}}}} \right) \\ &\approx \frac{q\delta D_I}{C_{ox}} \left( 1 - \frac{x}{2\epsilon_{Si}} \sqrt{\frac{2\epsilon_{Si}qN_A}{2\phi_B}} \right) \\ &\approx \frac{q\delta D_I}{C_{ox}} \left( 1 - \frac{x}{W_d} \right) \end{aligned} \quad (1.13)$$

dove l'approssimazione fatta è dettata dalla ragionevole ipotesi che il potenziale superficiale non viene modificato in maniera significativa.

Il numero dei droganti in ogni punto del dispositivo segue la statistica di Poisson, perciò possiamo affermare che:

$$\delta N = (N_A dx dy dz)^{1/2} \rightarrow \delta D_I = \frac{(N_A dx dy dz)^{1/2}}{WL} \quad (1.14)$$

L'equazione 1.13 esprime quanto impatta la variazione di un drogante alla posizione (x,y,z) sulla soglia del transistor. A questo punto ci calcoliamo il valor quadratico medio di tal variazione elevando al quadrato la 1.13 e sostituendo la  $\delta D_I$  con l'espressione 1.14:

$$\langle \Delta V_T^2 \rangle = \delta V_T^2 = \frac{q^2 N_A}{C_{ox}^2 L^2 W^2} \left(1 - \frac{x}{W_d}\right)^2 dx dy dz \quad (1.15)$$

Essendo che la fluttuazione dei droganti in tutti i punti del dispositivo noi la supponiamo incorrelata e completamente randomica, risulta applicabile la sovrapposizione degli effetti. Per questo motivo integriamo in tutto il volume la 1.15 ottenendo il valor quadratico medio totale che coincide con la varianza della tensione di soglia:

$$\sigma_{V_T}^2 = \frac{q^2 N_A}{C_{ox}^2 L^2 W^2} \int_0^W \int_0^L \int_0^{W_d} \left(1 - \frac{x}{W_d}\right)^2 dx dy dz \quad (1.16)$$

La dispersione della tensione di soglia in termini di deviazione standard risulta pertanto[8]:

$$\sigma_{V_T} = \frac{q}{C_{ox}} \sqrt{\frac{N_A W_d}{3LW}} \quad (1.17)$$

Nell'analisi appena descritta abbiamo assunto che il potenziale superficiale è uniforme in entrambe le direzioni del dispositivo W ed L. In altre parole non si tiene conto dell'impatto della natura atomistica del drogaggio sul potenziale superficiale, che quindi dipende solo dalla carica media (o totale) di tutta la depletion region. Questa assunzione nella regione di soglia-sottosoglia non è valida, in quanto l'iniezione di carica è dominata dalla barriera di potenziale più bassa piuttosto

che dal suo valor medio. Come abbiamo anticipato, infatti, i fenomeni percolativi della corrente risultano determinanti. Inoltre nell'analisi appena svolta abbiamo assunto un profilo di drogaggio costante e una geometria semplice. Nella realtà abbiamo forme 3D della cella anche complicate e drogaggi sicuramente non costanti. In generale, perciò, si preferisce adottare simulazioni numeriche 3-D per il calcolo delle fluttuazioni della tensione di soglia piuttosto che un semplice calcolo da modelli analitici approssimati. Nonostante questo, spesso è interessante avere una prima rapida stima della dispersione della  $V_T$ . In questi casi tali modelli tornano utili.

Dall'equazione 1.17 derivano i modelli di Stolk e Takeuchi. Da questa espressione della  $\sigma_{V_T}$  esplicitando infatti la larghezza della regione svuotata a soglia  $W_d = \sqrt{\frac{2\epsilon_{Si}}{qN_A} 2\phi_B}$  otteniamo:

$$\sigma_{V_T} = \frac{\sqrt[4]{4q^3\epsilon_{Si}\phi_B} T_{ox}}{\sqrt{3}} \frac{\sqrt[4]{N_A}}{\epsilon_{ox} \sqrt{LW}} \quad (1.18)$$

Questa espressione, come già detto, non tiene conto dei picchi di potenziale che si formano nel canale per la presenza degli atomi accettori ionizzati. Spesso, infatti, questo fenomeno non può essere trascurato in regime di soglia, in quanto le fluttuazioni di corrente sono determinate proprio dalla variabilità dei percorsi percolativi. Per tenere conto di ciò è possibile derivare una più accurata espressione analitica [5] in cui compare un termine additivo legato alle fluttuazioni del potenziale che va a sommarsi al termine capacitivo legato alle variazioni del campo elettrico:

$$\sigma_{V_T} = \frac{\sqrt[4]{4q^3\epsilon_{Si}\phi_B}}{\sqrt{3}} \left[ \frac{T_{ox}}{\epsilon_{ox}} + \frac{kT}{q} \frac{1}{\sqrt{4q\epsilon_{Si}\phi_B N_A}} \right] \frac{\sqrt[4]{N_A}}{\sqrt{LW}} \quad (1.19)$$

L'equazione 1.19 rappresenta il modello analitico di Stolk. Questa correzione, però, la maggior parte delle volte risulta essere trascurabile; ovvero il primo termine nelle parentesi quadre (legato alle variazioni del campo elettrico) risulta di solito ben maggiore del secondo (legato alle fluttuazioni del potenziale), portando la 1.19 ad essere approssimata con la precedente 1.18. Ciò è dovuto al fatto

che la correzione tiene conto solo dell'effetto medio sul canale di questi picchi di potenziale, quando in realtà gli effetti che determinano di più le fluttuazioni statistiche sono localizzati. Ciò rende necessario adottare simulazioni numeriche atomistiche 3D della cella soprattutto in regime di soglia-sottosoglia, dove questi effetti percolativi e di picchi di potenziale sono fortissimi e determinanti.

Dall'equazione 1.17 esplicitando  $N_A W_d$  in funzione della tensione di soglia  $N_A W_d = \frac{C_{ox}}{q} (V_T - V_{fb} - 2\phi_B) \approx \frac{C_{ox}}{q} (V_T + 0.1V)$  e moltiplicando tale dispersione per un fattore di correzione empirico 1.2 otteniamo il modello di Takeuchi[6]1.20. Questo modello risulta molto comodo: fissata la geometria del dispositivo ( $W, L, T_{ox}$ ), e calcolata la soglia media (facilmente misurabile, a differenza della  $\phi_B$ ), otteniamo la deviazione standard della  $V_T$ :

$$\sigma_{V_T} = \sqrt{\frac{q}{3\epsilon_{ox}}} \sqrt{\frac{T_{ox} (V_T + 0.1V)}{LW}} \times 1.2 = B_{V_T} \sqrt{\frac{T_{ox} (V_T + 0.1V)}{LW}} \quad (1.20)$$

Il modello di Asenov[7], invece, è puramente empirico ed è valido solo nel caso di canali corti. Per i dispositivi con lunghezza di canale  $L < 100nm$  e con concentrazione di drogaggio  $N_A > 10^{18}cm^{-3}$  non è più vero che  $\sigma_{V_T} \propto N_A^{0.25}$  ma si è empiricamente notato una dipendenza più forte dal drogaggio  $\sigma_{V_T} \propto N_A^{0.4}$ . Questa dipendenza più forte dal doping si spiega con il fatto che i modelli analitici di Stolk e Takeuchi tengono conto solo della fluttuazione del numero dei droganti, ma non includono gli effetti sulla posizione randomica di ciascuno di essi. Il modello di Asenov per la dispersione della tensione di soglia risulta quindi essere:

$$\sigma_{V_T} = 3.19 \times 10^{-8} \frac{T_{ox} N_A^{0.401}}{\sqrt{L_{eff} W_{eff}}} \quad (1.21)$$

Tale modello è stato ottenuto dal fitting dei risultati delle simulazioni atomistiche 3D con dispositivi con  $L_{eff} \in [30nm, 100nm]$ ,  $W_{eff} \in [50nm, 500nm]$ ,  $T_{ox} \in [1nm, 6nm]$ ,  $N_A \in [1 \times 10^{18}cm^{-3}, 5 \times 10^{18}cm^{-3}]$ .

Notiamo come invece la dipendenza dalla geometria è la stessa dei modelli di prima: rispetto all'area  $\sigma \propto \frac{1}{\sqrt{L_{eff} W_{eff}}}$  e rispetto allo spessore dell'ossido  $\sigma \propto T_{ox}$ .

### 1.2.2 Random Telegraph Noise (RTN)

Spesso in natura (ma non solo!) osserviamo fenomeni e processi che oscillano fra diversi stati discreti generando un segnale il cui valore dipende dallo stato in cui si trova il sistema stesso. Quando gli stati sono due il segnale generato da tale processo ricorda il segnale generato dal telegrafo, e per questo viene chiamato *Random Telegraph Signal* (RTS); quando poi il processo in questione risulta randomico/stocastico viene chiamato più propriamente *Random Telegraph Noise* (RTN) per sottolineare la sua natura rumorosa.

Nell'ambito dell'elettronica integrata l'RTN assume particolare rilevanza nei dispositivi transistor MOS o simili, in quanto risulta essere il rumore sicuramente più significativo per tecnologie molto scalate. Data la sua natura dinamica, infatti, nell'ambito delle memorie l'RTN risulta essere difficilmente correggibile e rende complicato lo scaling del dispositivo stesso. Se, inoltre, osserviamo questo fenomeno nel tempo, la combinazione di tanti RTN genera il cosiddetto rumore flicker, chiamato anche rumore  $1/f$ , dal momento in cui presenta uno spettro di potenza inversamente proporzionale alla frequenza e quindi particolarmente rilevante a basse frequenze.

In passato l'elettronica digitale era considerata "immune" dai disturbi di questo tipo, ma non è più vero per tecnologie molto scalate. Ciò ha reso necessario uno studio approfondito del fenomeno con lo scopo di capire come possa influire sul funzionamento dei dispositivi e di comprendere il trend di scaling per i prossimi nodi tecnologici. Infatti, a differenza dell'elettronica analogica, nel digitale non si riesce sempre ad aumentare le dimensioni geometriche dei dispositivi per limitare gli effetti dell'RTN, dal momento che l'obiettivo primario è proprio l'aumento della densità di componenti.

#### Origine del fenomeno e teoria di Mueller e Shulz

La causa di questo rumore è stata oggetto di numerosi studi e numerose pubblicazioni. L'origine è attribuita alla presenza di siti trappola localizzati all'interfaccia *Silicio/Ossido*, la cui presenza è a sua volta causata dalla interrotta



periodicità del reticolo cristallino del Silicio. Queste trappole danno origine a fenomeni di intrappolamento e rilascio dei portatori durante la conduzione facendo rispettivamente abbassare e alzare la corrente (Fig: 1.10(a)). Tale rumore spesso si preferisce vederlo come un discostamento di transcaratteristica, e quantificarlo come “shift” di soglia  $\Delta V_T = V_T^{trappola-occupata} - V_T^{trappola-vuota}$  (Fig. 1.10(b)).

Considerando un transistor MOS di tipo n, un elettrone intrappolato non contribuisce alla conduzione, anzi, esso agisce come centro di scattering abbassando la corrente totale, in quanto risulta essere a tutti gli effetti una carica fissa negativa ostacolante il flusso di elettroni da source a drain.

La presenza di legami spuri ed interrotti all’interfaccia dei due materiali causa, come abbiamo detto, la presenza delle trappole, che così danno origine a livelli energetici localizzati all’interno dell’EnergyGap. Dato che la probabilità di occupazione di tali livelli da parte dei portatori dipende dalla distanza tra il livello del sito trappola e quello di Fermi, la probabilità di occupare una trappola dipenderà anche dalla tensione di gate in cui si trova il dispositivo.

Al fine di minimizzare l’effetto di questo rumore è necessario controllare e rendere la densità di stati superficiali la più bassa possibile. Spesso, inoltre, le trappole non sono situate solo all’interfaccia, ma anche all’interno dell’ossido di gate: esse possono essere raggiunte dai portatori (con minor probabilità) per effetto tunnel, causando analogamente un abbassamento della corrente.

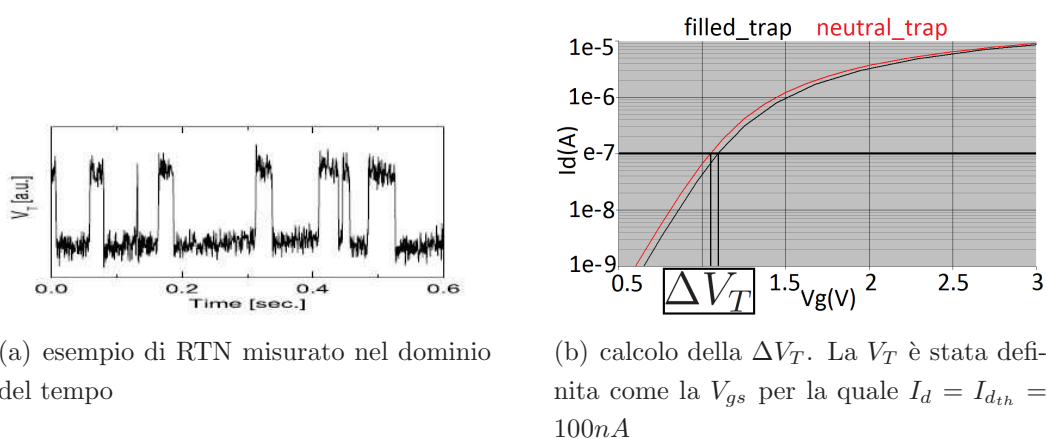


Figura 1.10: Il fenomeno del *Random Telegraph Noise*

Come abbiamo detto, il rumore causato dall'RTN rappresenta un parametro molto critico per la progettazione ed il disegno di circuiti analogici. Negli ultimi anni, inoltre, anche nell'ambito dei circuiti digitali e delle memorie l'RTN sta diventando un fattore determinante. Questo è anche dovuto al fatto che i valori di  $\Delta V_T$  riscontrati nella realtà risultano molto maggiori di quelli previsti dalla teoria. Ma andiamo con ordine.

L'ampiezza del rumore dell'RTN è stata oggetto di lunghe discussioni. Inizialmente si calcolava la variazione relativa della conduttanza causata dall'intrappolamento della carica come la somma di due effetti: la variazione del numero di portatori disponibili per la conduzione e la variazione della mobilità 1.22 [16][17].

$$\frac{\Delta G}{G} = \frac{\Delta N}{N} + \frac{\Delta \mu}{\mu} \quad (1.22)$$

In questo modello molto grezzo si quantifica la variazione di conduttanza, dovuta all'intrappolamento di una singola carica mobile, dalla variazione di un portatore libero nel canale, nel quale la corrente scorre uniformemente, e dalla variazione della mobilità media totale per la creazione di un centro di scattering all'interfaccia. L'ampiezza la stimiamo per i due termini separatamente. Possiamo quantificare la variazione relativa del numero di portatori come

$$\frac{\Delta N}{N} \leq \frac{1}{N_s A} \approx \frac{1}{100} = 1\% \quad (1.23)$$

dove la densità di elettroni  $N_s \approx 4 \times 10^{10} \text{cm}^{-2}$  e l'area di gate  $A = 0.5 \times 0.5 \mu\text{m}^2$  per i campioni riportati nell'articolo di Mueller e Shulz [16][17]. Questo porta a stimare circa 100 cariche nell'area attiva del dispositivo. La 1.23 è però una sovrastima, in quanto non tutta la carica immagine finisce nel canale, ma parte di essa sull'elettrodo di gate e nel substrato. Pertanto  $\Delta N = 1$  risulta essere il "worst case". Con questi dati otteniamo una variazione massima di circa l'1%, il che non spiega il 10% di  $\Delta G/G$  sperimentalmente osservato.

Il termine di mobilità può essere visto in termini di numero medio di eventi di scattering che un portatore incontra durante il suo cammino conduttivo. Noi stimiamo il cambiamento di mobilità  $\Delta\mu$  dovuto ad un addizionale evento di scattering relazionando il tasso di scattering medio  $1/\tau$  con la mobilità  $\mu$ ,

$$\frac{1}{\tau} = \frac{e}{m^*} \frac{1}{\mu} = N v_{th} \sigma_{sc} \quad (1.24)$$

dove  $e$  è la carica elettronica,  $m^*$  è la massa efficace,  $N$  è la densità degli eventi di scattering,  $\sigma_{sc}$  è la cross section dei centri di scattering, e  $v_{th}$  la velocità termica degli elettroni. In una semplice pittura di trasporto di carica possiamo stimare pertanto la mobilità come  $\mu = \frac{e}{m^*} \frac{1}{N v_{th} \sigma_{sc}}$ .

Se gli eventi di scattering relativi all'intrappolamento della carica e tutti gli altri centri di scattering nel canale li assumiamo indipendenti l'uno dall'altro, la risultante mobilità totale nel canale è composta in accordo alla regola di Matthiessen:

$$\frac{1}{\mu} = \frac{1}{\mu_{rest}} + \frac{1}{\mu_{st}} = \frac{1}{\mu + \Delta\mu} + \frac{1}{\mu_{st}} \quad (1.25)$$

dove  $\mu_{st}$  è la mobilità dovuta solo alla singola trappola, e  $\mu_{rest} = \mu + \Delta\mu$  è la mobilità dovuta a tutti gli altri eventi di scattering nel canale.

Ora, risolvendo l'eq.1.25 per  $1/\mu_{st}$ , inserendo il risultato dentro l'eq.1.24 e sostituendo  $N\sigma_{sc,st} = \sigma'_{sc,st}/(WL)$ , otteniamo un'espressione per la larghezza del centro di scattering  $\sigma'_{sc,st}$  richiesta per causare una fluttuazione di mobilità dovuta all'RTN pari a, per esempio, l' 1%:

$$\sigma'_{sc,st} = \frac{WLe}{m^* v_{th}} \left( \frac{1}{\mu} - \frac{1}{\mu + \Delta\mu} \right) = 370nm \quad (1.26)$$

dove  $WL = 0.5 \times 0.5 \mu m^2$ ,  $\mu = 300 cm^2/Vs$  e  $T = 293K$ . Per spiegare un  $\Delta\mu/\mu$  dell' 1% l'addizionale centro scattering indotto dall'intrappolamento della carica dovrebbe avere una irragionevole larghezza di 370nm, che è dello stesso ordine di grandezza della lunghezza del dispositivo stesso. Un realistico centro di scattering, infatti, risulta avere una larghezza di valore ben più basso, per esempio  $\approx 10nm$ .

Come abbiamo visto, il modellino analitico 1.22 non spiega affatto una fluttuazione della conduttanza del 10% sperimentalmente osservata, ma dell'1%, con

il contributo dominante dovuto alla fluttuazione del numero di portatori mobili. Questo ci fa capire che tal modello risulta non sufficiente e mal descrittivo della realtà. Ciò ha spinto Mueller e Shulz a cercare modelli e considerazioni fisiche differenti. Il trasporto di carica attorno ad un potenziale locale è in genere spiegato, infatti, in due diversi modi[16]: in termini di mobilità, come discusso sopra, per la quale i portatori sono scatterati alla perturbazione di potenziale, o in termini di flusso continuo. Il primo caso è apprezzabile quando il libero cammino medio dei portatori tra gli eventi di scattering  $\lambda$  è molto maggiore dell'estensione della perturbazione del potenziale  $\approx \sigma'_{sc}$ , mentre il secondo accade quando  $\lambda \ll \sigma'_{sc}$ , ovvero quando consecutivi eventi di scattering si sovrappongono nello spazio e nel tempo dando origine ad un flusso continuo attorno alla perturbazione. Nel nostro caso il libero cammino medio risulta  $\lambda \approx 0.5\sigma'_{sc}$ , il che rende un'approssimazione migliore la seconda metodologia.

Per stimare la  $\Delta G/G$  in questa maniera, si è simulato il problema numericamente. La simulazione “worst case” dell'impatto di una carica repulsiva all'interfaccia modella la trappola carica come un disco di conduttività nulla nel canale. Se un disco di conduttività zero viene posizionato in un canale in cui la corrente è assunta uniforme, il flusso elettronico verrà deviato attorno alla perturbazione abbassando la corrente totale. Le distribuzioni della densità di corrente e del potenziale possono essere ottenute usando la legge di Kirchhoff e risolvendo l'equazione di Poisson in un modello geometrico 2D. Un modello analitico semplice che approssimi questa condizione è considerare la carica intrappolata come un disco di conduttività nulla ubicato in un piano uniformemente conduttivo ed infinitamente esteso. Questa ipotesi semplificativa porta a stimare la variazione di conduttanza relativa:

$$\frac{\Delta G}{G} = \frac{4 A_T}{\pi A_{ch}} = \frac{(2r_T)^2}{W^2} = \left( \frac{I_{loc}}{I_d} \right)^2 \quad (1.27)$$

dove  $A_T$  è l'area del disco e  $A_{ch}$  è quella del canale. Nell'eq1.27 abbiamo assunto che l'area del canale sia quadrata  $W \times W$ , che  $A_T \ll A_{ch}$  e soprattutto che la densità di corrente nel canale con trappola neutra sia costante ovunque, che in

altri termini vuol dire  $\frac{I_d}{W} = \frac{I_{loc}}{2r_T} = j_0$ . Questo modello quantifica la  $\Delta G/G$  come il quadrato del rapporto tra la corrente deviata dal disco e la corrente totale fluente nel canale.

Il modello analitico e quello numerico portano a risultati differenti di meno del 30% per raggi del disco pari a  $r_T = 100nm$ , ma questa differenza risulta sempre minore tanto più il raggio è assunto piccolo. A raggi ragionevoli ( $r_T \approx 10nm$ ) i metodi calcolano la fluttuazione della conduttanza nel “worst case” pari a  $\Delta G/G \approx 0.5\%$ , il che spiega ancora meno di prima i risultati sperimentalmente ottenuti.

Abbiamo finora considerato due diversi modelli per quantificare la variazione di conduttanza dovuta all'intrappolamento di una singola carica, ma in entrambi i casi si giunge a risultati inferiori di un ordine di grandezza rispetto alle misure sperimentali. Una prima ipotesi esplicativa può essere dovuta al fatto che i risultati sperimentali sono frutto del cosiddetto *Complex RTN*. Questo fenomeno, inizialmente osservato da N.Tega [18], prevede la combinazione di più processi di intrappolamento e rilascio dei portatori legati a diversi siti trappola. Esso può in effetti causare un'oscillazione della tensione di soglia su più livelli e per questo una fluttuazione più ampia, ed inoltre causa una dipendenza  $1/f^2$  in regioni distinte nel dominio delle frequenze. In molti casi, però, abbiamo evidenze sperimentali di solo due livelli di tensione e perciò di una sola trappola. Il *Complex RTN*, d'altronde, si manifesta solo per dispositivi piuttosto grandi: per una tecnologia floating gate di 60nm di lunghezza di canale, infatti, in media una cella contiene un numero di trappole  $< 1$  (o non ne contiene o ne contiene una sola insomma), e per questo non potrà presentare alcun fenomeno di *Complex RTN*.

Nella realtà, perciò, abbiamo evidenze sperimentali di *Giant RTN* dovuti all'intrappolamento di una carica causata dalla presenza di una singola trappola. Per spiegare finalmente tale fenomeno abbiamo bisogno di due diversi concetti: la percolazione della corrente e gli effetti di bordo.

Mueller e Schulz avevano ipotizzato che gli elevati valori del disturbo potessero essere dovuti a perturbazioni localizzate intorno ai siti trappola in un canale del MOSFET non uniforme. Nei modelli precedenti, infatti, avevamo sempre ipotizzato una conduzione uniforme, ma nella realtà osserviamo una forte disomogeneità. Essa è dovuta alla presenza di cariche fisse interfacciali, ovvero i

droganti, e nell'ossido, responsabili di variazioni locali della conduttanza di canale e quindi di densità di corrente, allo stesso modo delle cariche fisse intrappolate. Questo provoca valli e picchi per il potenziale, e la corrente tende a scorrere tra le valli infiltrandosi tra le cariche fisse in percorsi praticamente unidimensionali, come mostrato in Fig.1.11

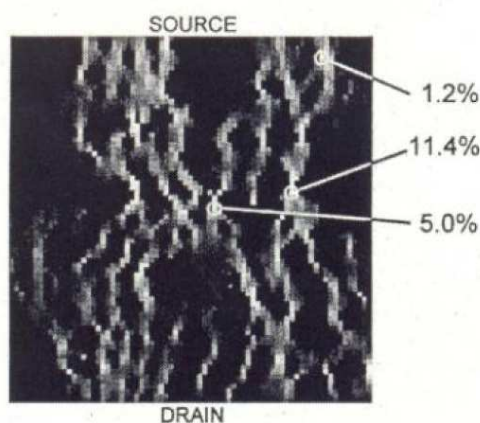


Figura 1.11: Densità di corrente simulata in un canale con distribuzione casuale di carica fissa (densità  $4 \times 10^{11}$ ). Le percentuali indicano la variazione di corrente dovuta all'effetto di una trappola in quel punto.

Questo fenomeno risulta rilevante quando le dimensioni del dispositivo diventano piccole, in quanto i percorsi conduttivi risultano essere pochi. Se per esempio abbiamo un MOS che presenta una decina di cammini percolativi a soglia, nei quali ipotizziamo una corrente  $I_{loc} \approx I_d/10$  per ciascuno, intrappolando un elettrone lungo uno di essi, avremo ragionevolmente un abbassamento del 10% circa della corrente totale, in quanto localmente in quel punto la corrente va sostanzialmente ad annullarsi. Quando la trappola è strategicamente ubicata lungo uno dei filamenti conduttivi, perciò, la corrente di drain subisce un forte calo, il che risulta quantitativamente in accordo con i dati sperimentalmente osservati. In altre parole, una conduzione non uniforme nel canale provoca picchi di corrente locale  $I_{loc}$  e, perciò, ricordando l'eq.1.27  $\frac{\Delta G}{G} = \left(\frac{I_{loc}}{I_d}\right)^2$ , dei valori di  $\frac{\Delta G}{G}$  indotti dall'intrappolamento di una carica mobile più elevati di quelli predetti dalla teoria precedente.

Per spiegare pienamente il fenomeno dell'RTN, dunque, dobbiamo ricorrere all'RDF, o meglio, non possiamo assolutamente escludere gli effetti dell'RDF sull'intrappolamento di una carica durante la conduzione. Ovviamente quando il dispositivo risulta avere grandi dimensioni, la corrente può ragionevolmente essere considerata uniforme, in quanto i percorsi percolativi presenti sono numerosi.

L'altra causa dell'aumento anomalo dell'RTN è da ricercarsi nel *current crowding* ai bordi dell'area attiva causato dall'effetto punta del campo elettrico. Il campo elettrico, infatti, risulta essere più alto nei punti in cui il raggio di curvatura è minore. Tipicamente questo accade sui due bordi dell'area attiva con l'STI, il che provoca un abbassamento locale della soglia. Ciò ha come conseguenza la formazione dei percorsi preferenziali della corrente (appunto *current crowding*) in queste due regioni, e perciò, analogamente a prima, *Giant RTN* quando le trappole catturano elettroni lungo uno di questi due cammini.

### Scaling dell'RTN

Finora abbiamo visto come la variabilità statistica causata dal fenomeno dell'RTN sia di assoluta importanza e rilevanza, ma risulta altrettanto importante capire come le distribuzioni di  $\Delta V_T$  dipendano dalle caratteristiche del dispositivo (geometria, doping, tensioni applicate), al fine di prevedere la variabilità indotta dall'RTN per tecnologie future. In una prima approssimazione 1D del problema, possiamo considerare un elettrone intrappolato in uno stato localizzato nell'ossido di tunnel come una carica in eccesso nel piatto del condensatore ideale  $C_{ox}$ . Questo induce una  $\Delta V_T = q/C_{ox}$ , ovvero:

$$\Delta V_T = \frac{qt_{ox}}{\alpha_G \epsilon_{ox} WL} \quad (1.28)$$

dove  $\alpha_G$  è il fattore di accoppiamento capacitivo tra control-gate e floating-gate,  $t_{ox}$  è lo spessore dell'ossido di tunnel e  $W$  ed  $L$  sono rispettivamente la larghezza e la lunghezza del dispositivo. È significativo notare come allo scalare della tecnologia lo shift di soglia indotto dall'RTN, previsto in 1.28, aumenti, il che è

concorde con le aspettative. In realtà, però, questo semplice modello elettrostatico risulta troppo approssimativo, perché considera il dispositivo ideale, ovvero con drogaggio continuo, e perciò trascura gli effetti di percolazione della corrente e di current crowding, il che, come abbiamo visto nei paragrafi precedenti, è inaccettabile per dispositivi scalati. Ciò è evidente in quanto lo shift di soglia previsto non risulta dipendente dalla posizione in cui la trappola è ubicata all'interfaccia. Nonostante questo, la 1.28 risulta utile per capire grossolanamente alcune dipendenze. Nelle memorie Flash, infatti, possiamo prevedere che l'ampiezza della  $\Delta V_T$  sia maggiore rispetto a quella di altre memorie (per esempio DRAM). Innanzitutto il coefficiente  $\alpha_G$  risulta  $\approx 0.6$ , perciò  $< 1$ , il che aumenta l'entità dello shift di soglia. Inoltre, in tali memorie, non è possibile scalare più di tanto lo spessore dell'ossido di tunnel da una tecnologia all'altra, in quanto è necessario garantire un mantenimento del dato da parte della floating-gate per circa 10 anni, il che porta ad avere specifiche di ritenzione del dato estremamente limitanti (per esempio per una tecnologia NAND a 32 nanometri è richiesto che le perdite di carica immagazzinata nella floating-gate non eccedano un elettrone all'anno).

Nelle memorie Flash l'RTN si traduce in un'oscillazione della tensione di soglia fra due livelli ed è di fondamentale importanza che l'entità di questo "salto" durante l'operazione di lettura non produca in uscita un dato diverso da quello memorizzato nella cella. In pratica l'effetto delle oscillazioni di RTN è quello di allargare le distribuzioni delle tensioni di soglia risultanti dalle operazioni di programmazione e cancellazione (1.5) con delle code esponenziali ([11][12][13][14]) per cui la probabilità che all'incertezza data da queste operazioni si aggiunga una variazione  $\Delta V_T$  è data dalla densità:

$$p(|\Delta V_T|) = \frac{1}{\sigma} \exp\left(-\frac{|\Delta V_T|}{\sigma}\right) \quad (1.29)$$

la cui cumulativa corrispondente risulta:

$$1 - F(|\Delta V_T|) = \exp\left(-\frac{|\Delta V_T|}{\sigma}\right) \quad (1.30)$$

Questo allargamento limita il margine di lettura (read margin), ossia quella finestra in cui si può avere la ragionevole certezza di leggere il dato correttamente. Se la variazione  $\Delta V_T$  per una data cella è troppo ampia, questo margine svanisce



e non è più possibile definire una tensione di gate con cui leggere senza ambiguità il dato memorizzato. I dati sperimentali raccolti su una memoria NOR Flash a 65 nanometri ([11][12]) hanno confermato il comportamento esponenziale delle fluttuazioni della soglia, come risulta evidente dal grafico della funzione di distribuzione cumulativa in Figura 1.12, dove sono rappresentati i risultati ottenuti fra la prima e l'ennesima acquisizione dei dati attraverso un accesso sequenziale a tutte le celle dell'array NOR.

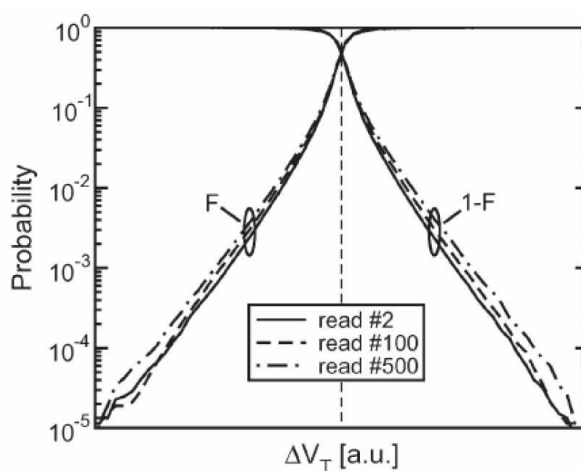


Figura 1.12: Funzione di distribuzione cumulativa estrapolata al primo, al centesimo e al cinquecentesimo accesso in lettura all'array. I valori sperimentali sono riportati in arbitrary units ([a.u.]), in quanto si tratta di dati riservati.

Tipicamente, perciò, la distribuzione cumulativa  $1 - F(|\Delta V_T|)$  (che mi dice in un punto  $x$  la percentuale dei valori aventi  $\Delta V_T > x$ ) segue un andamento esponenziale decrescente. Un parametro caratterizzante tale distribuzione è la pendenza della retta risultante in grafico semilogaritmico ( $x$  vs  $\log(y(x))$ ), ovvero la lambda  $\lambda$  (misurata in mV/decade). Questo parametro ci dice quanti mV sono necessari affinché la distribuzione  $1 - F(|\Delta V_T|)$  diventi più piccola di una decade, ed è legata alla deviazione standard della distribuzione in maniera proporzionale attraverso il fattore  $\ln(10)$ .

$\lambda$  rappresenta il parametro fondamentale per la caratterizzazione statistica dell'RTN e quindi è necessario capire come dipenda dai parametri del dispositivo ( $L, W, t_{ox}, N_A, V_d \dots$ ). Per far ciò si è ottenuta una formulazione empirica[15],

risultato di numerose simulazioni numeriche ottenute facendo variare i parametri separatamente in dispositivi randomizzati secondo una procedura di tipo *Monte Carlo* (metodologia discussa nel prossimo capitolo) per una tecnologia NOR a 65nm. Da questo modello lambda risulta, in perfetto accordo con i dati sperimentali,

$$\lambda = \frac{K t_{ox}^\alpha \sqrt{N_A}}{\alpha_G W \sqrt{L}} \quad (1.31)$$

Ora andiamo ad esplorare le dipendenze di  $\lambda$  dai parametri separatamente, avendo in mente sempre l'equazione 1.31. Simulando i dispositivi imponendo diverse tensioni di drain (100mV e 1V) ci si accorge infatti come la distribuzione, pur mantenendo la stessa pendenza, si stringe con l'aumentare della  $V_d$  (Fig.1.13 a)).

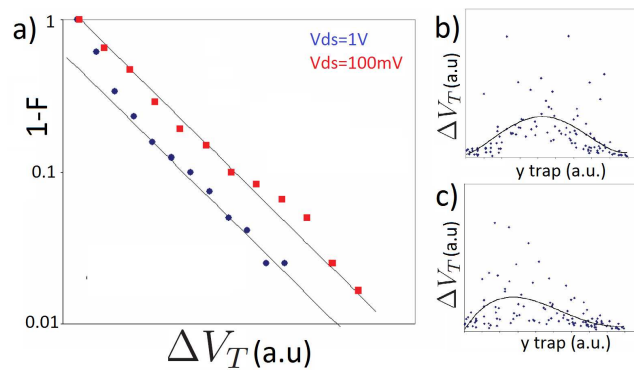


Figura 1.13: a) Probabilità cumulative i cui valori simulati sono stati ottenuti con  $V_d = 100mV$  e  $V_d = 1V$ . Sono poi plottate le ampiezza delle fluttuazioni  $\Delta V_T$  in funzione della posizione lungo il canale per b)  $V_d = 100mV$  e c) per  $V_d = 1V$

Questo è dovuto al fatto che aumentando la tensione di bias al drain si riducono le trappole RTN attive. A basse  $V_d$  le trappole vicine al source e al drain risultano avere un effetto trascurabile sull'ampiezza della  $\Delta V_T$ , mentre quelle al centro hanno un effetto ben maggiore, facendo sì che la distribuzione  $\Delta V_T(y)$  (dove  $y$  è la coordinata spaziale lungo la lunghezza del canale  $L$  nella quale posiziono la trappola) sia simmetrica (Fig.1.13 b) ). Ad alte  $V_d$ , invece, il picco si sposta

verso il source (Fig.1.13 c) ), in quanto al drain la zona svuotata risulta essere più ampia e, non limitandosi all'interfaccia, fa sì che la corrente scorra anche più in profondità nel substrato. Con una  $V_d$  più grande le trappole rimaste fuori dalla regione di drain continueranno a indurre la loro variazione di soglia e perciò l'effetto che si ottiene è quello di avere una "densità di trappole equivalente" ridotta, che induce semplicemente una traslazione rigida della distribuzione di probabilità cumulativa. Da questo se ne deduce che la pendenza della distribuzione non cambia variando la tensione applicata al drain, e perciò la  $\lambda$  risulta indipendente dalla  $V_d$ .

Per quanto riguarda la dipendenza dall'area,  $\lambda$  risulta  $\propto \frac{1}{W\sqrt{L}}$  ovvero  $\propto \frac{1}{Area^{0.75}}$ . Rispetto al modello 1.28 riscontriamo due differenze: la dipendenza dall'area è più debole e soprattutto la dipendenza da  $W$  è diversa da quella da  $L$ . La dipendenza più forte dalla larghezza del dispositivo invece che dalla lunghezza è spiegabile tenendo in considerazione l'effetto della percolazione della corrente sull'ampiezza della  $\Delta V_T$ . Più il dispositivo è largo, infatti, più c'è la possibilità per la corrente di disperdersi in tanti filamenti, mentre un dispositivo con  $W$  piccolo risulta avere pochi percorsi percolativi e perciò aumenta la probabilità di avere *Giant RTN*. Variando  $L$ , invece, i cammini percolativi della corrente cambiano poco, e dunque la dipendenza di  $\lambda$  da questo parametro risulta essere minore.

Per quanto riguarda la dipendenza dallo spessore dell'ossido di tunnel riscontriamo  $\lambda \propto t_{ox}^\alpha$  con  $\alpha \in (0, 1)$ . Con ossidi sottili la corrente scorre uniformemente nel canale; quando invece allontaniamo l'elettrodo di gate dall'interfaccia la densità di corrente risulta essere sempre più affetta da *current crowding* (picchi di corrente sugli spigoli del dispositivo), che induce *Giant RTN* qualora una trappola fosse posizionata sui bordi della cella (Fig.1.14). Questo spiega come ossidi spessi aumentano la probabilità di avere ampi  $\Delta V_T$ .

Infine la dipendenza di  $\lambda$  dal drogaggio risulta essere  $\propto N_A^{0.5}$ . Questo è attribuibile ancora una volta alla percolazione della corrente causata dalla natura discreta degli atomi droganti. La presenza degli ioni causa picchi e valli per il potenziale e perciò disuniformità nella densità di corrente, rendendo possibile la presenza di ampi  $\Delta V_T$  qualora le trappole fossero posizionate sopra uno dei percorsi preferenziali della corrente. Più il drogaggio è elevato più grande sarà la dispersione nei valori di  $\Delta V_T$ , in quanto ci saranno meno cammini percolativi della corrente. Ciò

ostacola lo scaling dimensionale del dispositivo, in quanto, per limitare l'effetto DIBL (abbassamento di barriera indotto dal drain) e l'estensione eccessiva della zona svuotata fra drain e substrato, si cerca di aumentare il drogaggio con lo stringimento delle dimensioni geometriche.

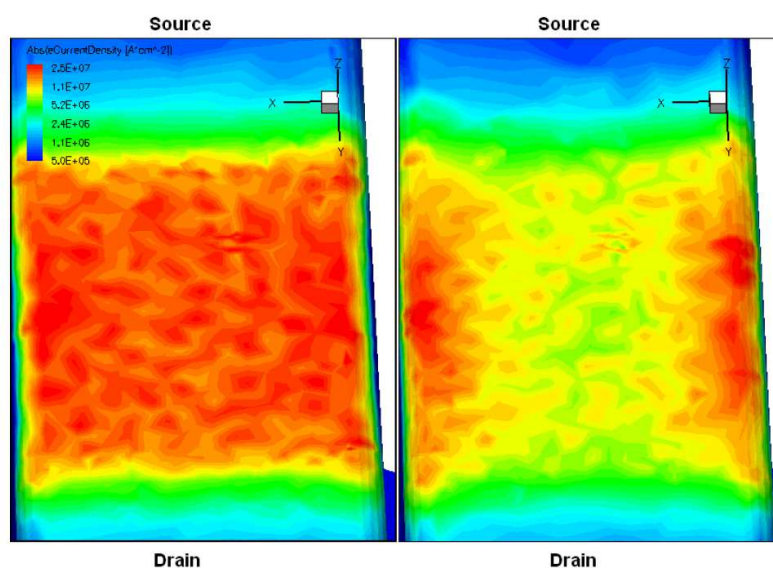
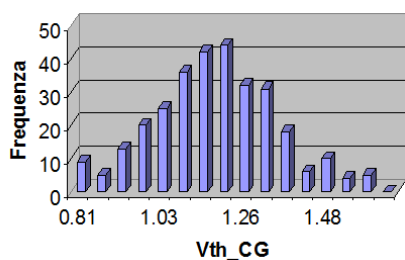


Figura 1.14: Densità di corrente per cella a 45 nm con drogaggio continuo e spessore dell'ossido di 2nm (a sinistra) e 8nm (a destra). Si nota come l'effetto del current crowding sia più elevato per dispositivi con ossidi spessi.

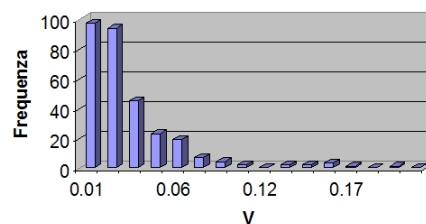
In questo elaborato di tesi non tratto volutamente l'analisi del *Random Telegraph Signal* nel dominio delle frequenze; ed inoltre considero sempre solo l'effetto di una singola trappola, non faccio alcun cenno all'effetto indotto dalla cattura di più elettroni da diversi siti trappola contemporaneamente. Questo perché sia il rumore  $1/f$  che il fenomeno del *MultiTrap* esulano dall'argomento di tesi.

### 1.3 Riduzione delle fluttuazioni

Nei paragrafi precedenti abbiamo visto come le dispersioni della soglia dovute al fenomeno dell'RDF e le fluttuazioni dell'RTN hanno impatto nei dispositivi decananometrici, in particolare abbiamo analizzato il dispositivo in Fig.1.1, il cui drogaggio accettore è uniforme in tutto il substrato. Dispositivi nominalmente identici hanno transcaratteristiche diverse per il solo fatto che gli atomi droganti ionizzati sono in quantità diversa ed hanno una distribuzione spaziale differente da dispositivo a dispositivo, il che provoca una dispersione statica “naturale” della tensione di soglia (Fig.1.15(a)). Dinamicamente, invece, accade che per la presenza di un sito trappola all'interfaccia  $Si/Ox$  si ha una fluttuazione su due stati della corrente, il che vuol dire in altri termini una  $\Delta V_T$ , la cui ampiezza è diversa a seconda della configurazione atomistica di drogaggio e della posizione della trappola (Fig.1.15(b)).



(a) fluttuazione naturale della soglia dovuta ai differenti profili atomistici di drogaggio



(b) fluttuazione della  $\Delta V_T$  dovuta alle diverse coppie configurazione di drogaggio - trappola

Figura 1.15:

Abbiamo visto come la distribuzione di probabilità cumulativa dell'ampiezza della  $\Delta V_T$  segue un andamento esponenziale decrescente, la cui pendenza in grafico semilogaritmico è quantitativamente descritta dal parametro  $\lambda$  (eq.1.31). L'idea è quella di modificare l'andamento di questa distribuzione cambiando opportunamente il drogaggio o la geometria del dispositivo. In tal caso vedremo come le dispersioni di RTN e di RDF si modificano e come queste due fluttuazioni statistiche sono correlate. Si ingegnerizza opportunamente la geometria e il dro-

gaggio della cella per cercare ottimizzazioni tecnologiche al fine di abbattere le fluttuazioni statistiche dell'RDF e soprattutto dell'RTN.

### 1.3.1 Drogaggi disuniformi

In figura 1.16 sono mostrati i profili di drogaggio disuniformi investigati in questo lavoro e comparati con il caso uniforme di riferimento avente  $N_A = 2 \times 10^{18} \text{cm}^{-3}$ . Per un confronto sensato, le concentrazioni di doping in Fig.1.16 sono scelte al fine di ottenere in tutti i casi la stessa  $V_T$  media a trappola scarica. Nel seguito sono mostrati i risultati ottenuti attraverso la metodologia *Monte Carlo* [19].

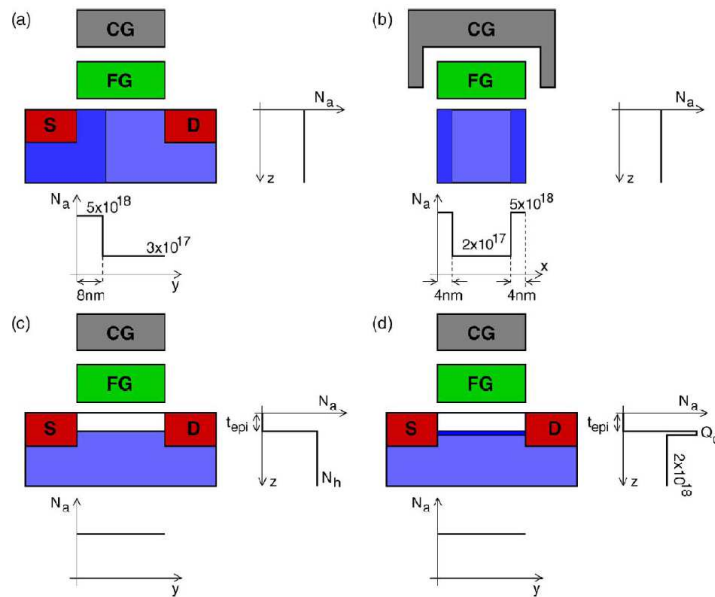


Figura 1.16: Schematici per i profili di doping investigati: a) drogaggio disuniforme lungo L, b) drogaggio disuniforme lungo W, c) drogaggio retrogrado, e) drogaggio a  $\delta$ -shape

A differenza del caso uniforme, la distribuzione di probabilità cumulativa risultante per il caso di drogaggio disuniforme lungo L (Fig.1.16a)) è una doppia pendenza nel grafico semilogaritmico (Fig.1.17a)). La slope per il caso di riferimento risulta minore per bassi valori di  $\Delta V_T$  e maggiore per alti, che vuol dire una distribu-

zione più larga a basse probabilità e più stretta ad alte. Questo è spiegabile dal fatto che per il dispositivo 1.16a) la regione lato source molto più drogata richiede una tensione al gate maggiore per essere invertita e quindi ha maggior controllo sulla conduzione source-drain e sulla soglia del dispositivo rispetto alla regione low-doped. Una trappola in quest'ultima zona che cattura un elettrone dal flusso conduttivo ha perciò poco impatto sulla soglia, causando un valore di  $\Delta V_T$  minore rispetto al caso uniforme. Poiché la regione più ampia è di gran lunga la regione meno drogata e che le trappole sono distribuite uniformemente sul canale, la distribuzione per alte probabilità si stringe. A basse probabilità, invece, le  $\Delta V_T$  sono principalmente determinate dalle trappole ubicate nella regione high-doped e la distribuzione si appiattisce per il maggior impatto sulla soglia di tali trappole rispetto al caso di riferimento.

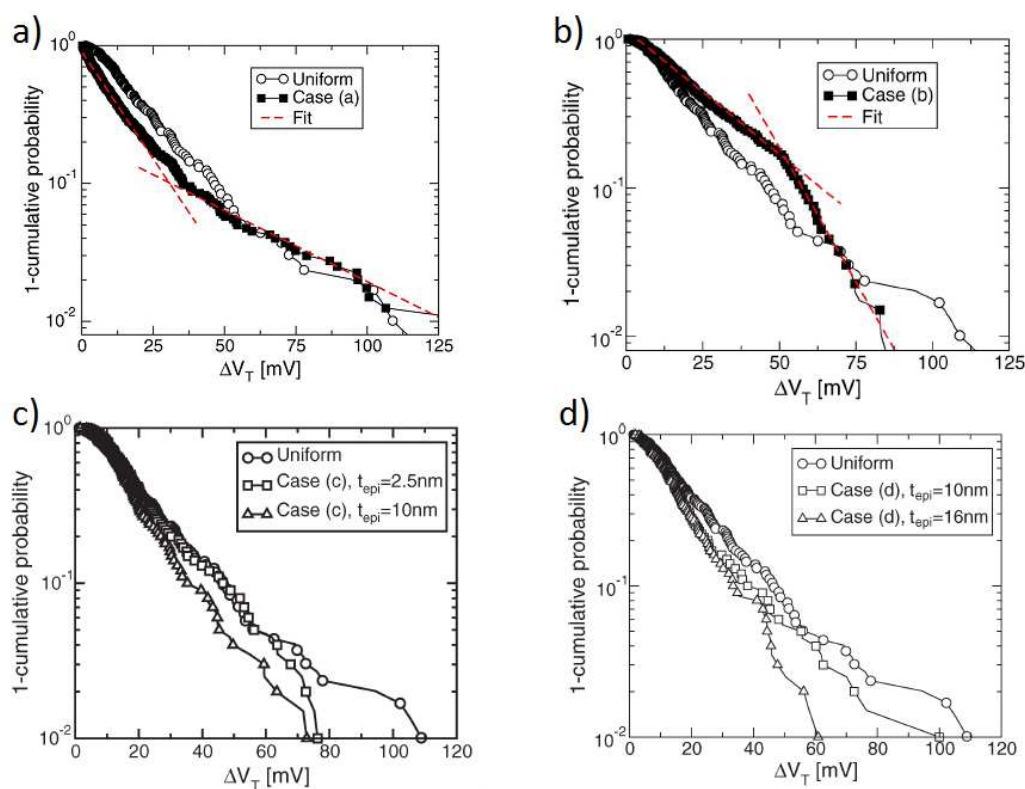


Figura 1.17: distribuzioni di probabilità cumulativa per i dispositivi a), b), c), e d) della Fig.1.16 confrontate con la cumulata nel caso di drogaggio uniforme

Nel caso invece di doping disuniforme lungo la larghezza della cella (Fig.1.16b)), abbiamo un drogaggio maggiore nei due bordi del dispositivo per alzar localmente la soglia in queste regioni. Ciò viene fatto per ridurre l'effetto del *current crowding* dovuto all'intensificazione del campo elettrico negli spigoli della cella. In questo modo la corrente scorre poco sui bordi e molto nel centro del dispositivo. Nel caso di doping uniforme le  $\Delta V_T$  più ampie, ovvero quelle a basse probabilità, sono dovute principalmente proprio a trappole situate nei bordi del dispositivo, ora invece questi siti hanno molto meno impatto. Ciò ha come effetto che la distribuzione per basse probabilità si stringe, come visibile in Fig.1.17b). Discorso contrario ad alte probabilità: le trappole ubicate nel centro del dispositivo, infatti, hanno impatto maggiore rispetto al caso uniforme, dal momento in cui la corrente scorre principalmente proprio nel centro della cella.

L'effetto di questo profilo di drogaggio sull'RTN è sicuramente benevolo, al contrario del dispositivo precedente, in quanto a basse probabilità troviamo valori di shift di soglia ben più piccoli rispetto al caso uniforme; ciò rende questa soluzione molto praticata nella progettazione della cella.

Abbiamo visto come le cause di *Giant RTN* siano da ricercarsi nel *current crowding* ai bordi della cella e nella natura percolativa della corrente. Adottando un drogaggio disuniforme lungo la larghezza del canale  $W$ , riusciamo a limitare il primo fenomeno, ottenendo perciò un netto miglioramento nella distribuzione delle  $\Delta V_T$ . Nelle figure 1.16 c) e d) troviamo invece dispositivi i cui drogaggi adottati hanno il fine di uniformizzare la corrente nel canale, rendendo la conduzione meno percolativa. La causa della percolazione, come detto più volte, è da ricercarsi nella presenza del doping stesso. L'idea, perciò, è semplicemente quella di svuotare il dispositivo dai droganti nella regione del canale, dove avviene la conduzione e dove perciò gli effetti del doping sono più rilevanti. Questo è il caso del dispositivo con drogaggio accettore retrogrado (Fig.1.16 c)), per il quale si sono simulati i seguenti casi:  $t_{epi} = 2.5nm$  con  $N_h = 2.5 \times 10^{18}cm^{-3}$ ,  $t_{epi} = 5nm$  con  $N_h = 3.4 \times 10^{18}cm^{-3}$ ,  $t_{epi} = 10nm$  con  $N_h = 5.7 \times 10^{18}cm^{-3}$  e  $t_{epi} = 16nm$  con  $N_h = 1 \times 10^{19}cm^{-3}$  (per  $t_{epi}$  più grandi si è considerato un drogaggio  $N_h$  più elevato per mantenere la stessa soglia media). L'idea di svuotare il canale o comunque di adottare un basso drogaggio risale agli anni '90 e lo scopo iniziale era la ricerca di un miglioramento di mobilità e transconduttanza e di un modo di controllare la tensione di soglia e la *subthreshold slope*[8]. Il substrato, però non è totalmente



svuotato dai droganti, ma è necessario impiantare una concentrazione  $N_h$  da una certa profondità in poi per evitare che il dispositivo sia suscettibile agli effetti di canale corto. Spesso, inoltre, anche la soluzione retrogrado può presentare alcune controindicazioni: adottare un  $N_h$  elevato in profondità aumenta le capacità parassite di source e drain e riduce la tensione di breakdown per quelle giunzioni. Per questo motivo la soluzione ideale e più spesso utilizzata è rappresentata dal profilo a  $\delta$ -shape, raffigurato in Fig.1.16 d), per il quale si inserisce una “delta” di doping per  $x = x_s$  e in profondità si tiene la concentrazione più bassa.

L’effetto complessivo di questi due profili descritti, perciò, è quello di ridurre mediamente l’ampiezza delle  $\Delta V_T$  rispetto al caso uniforme, il che si traduce in un abbassamento della  $\lambda$ , come visibile nelle figure 1.17 c) e d). Adottando queste soluzioni la forma esponenziale a singola “slope” rimane, ma la pendenza (in termini di mV/dec) diminuisce e decresce sempre di più man mano che si adottano  $t_{epi}$  più grandi. Man mano che i droganti vengono “spinti” più in profondità, infatti, il loro campo Coulombiano ha meno impatto sul potenziale superficiale e per questo la distribuzione si stringe.

Il drogaggio retrogrado e a delta risultano essere i profili più utilizzati a livello industriale dato il loro positivo impatto sulla variabilità statistica dell’RTN. A differenza della soluzione 1.16 b), inoltre, essi provocano un netto miglioramento anche per quanto riguarda la fluttuazione statistica “naturale” della tensione di soglia dovuta all’RDF, come chiaramente visibile in figura 1.18(a) per il caso  $t_{epi} = 10nm$ , in quanto la causa di tale dispersione è da ricercarsi proprio nella presenza degli atomi droganti ionizzati.

Con questi profili di drogaggio disuniformi verticalmente migliora la dispersione statistica della  $V_T$  dovuta all’RDF ( $\sigma_{V_T}$  più piccola) e migliora la dispersione statistica della  $\Delta V_T$  causata dall’RTN ( $\lambda$  più piccola); questi due fenomeni risultano perciò correlati in modo positivo, come visibile in figura 1.18(b) [19]. La forte connessione tra la variabilità della  $V_T$  e l’RTN per questi profili di drogaggio è facilmente visibile dai modelli di lambda 1.31 e di  $\sigma_{V_T}$  1.21, per cui  $\lambda \propto N_A^{0.5}$  e  $\sigma_{V_T} \propto N_A^{0.4}$ . Svuotando il canale dal doping si ha un abbassamento (idealmente un azzeramento) del drogaggio  $N_A$  e quindi una forte diminuzione sia di  $\lambda$  che della dispersione della soglia  $\sigma_{V_T}$ .

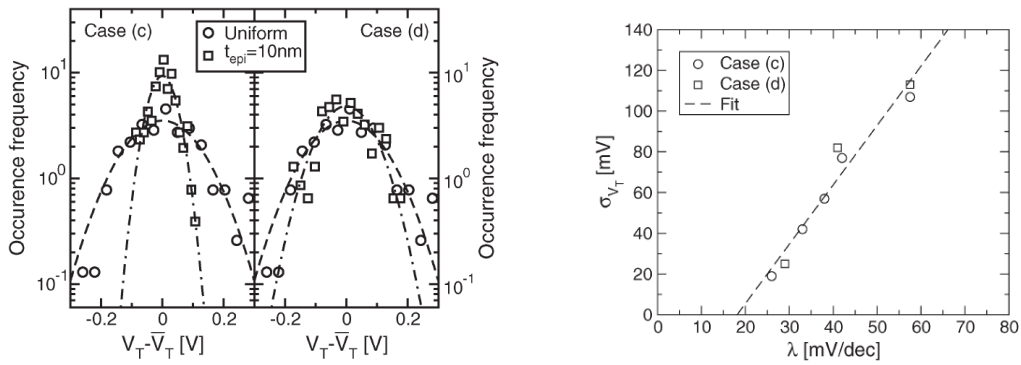
Un semplice modello per la fluttuazione “naturale” della tensione di soglia causata dall’RDF per il drogaggio retrogrado è presente in letteratura [8], e deriva dai modelli precedentemente discussi. La tensione di soglia e l’ampiezza della zona svuotata risultano rispettivamente:

$$V_T = V_{fb} + 2\phi_B + \frac{qN_A}{C_{ox}} \sqrt{\frac{4\epsilon_{Si}\phi_B}{qN_A} + x_s^2} - \frac{qN_A x_s}{C_{ox}} \quad (1.32)$$

$$Z_{dep} = \sqrt{\frac{4\epsilon_{Si}\phi_B}{qN_A} + x_s^2} \quad (1.33)$$

Con questi dati la 1.15 rimane valida, mentre nella 1.16 gli estremi di integrazione in  $dx$  diventano  $x_s$  e  $Z_{dep}$ , così che la dispersione della soglia risulta sensibilmente ridotta:

$$\sigma_{V_T} = \frac{q}{C_{ox}} \sqrt{\frac{N_A Z_{dep}}{3LW}} \left(1 - \frac{x_s}{Z_{dep}}\right) \quad (1.34)$$



(a) distribuzioni statistiche della  $V_T$  per i dispositivi c), e d) della Fig.1.16 confrontate con quella del caso di drogaggio uniforme

(b) correlazione tra  $\lambda$  e  $\sigma_{V_T}$  per drogaggio retrogrado c) e a delta d)

Figura 1.18:

### 1.3.2 Celle con bordi arrotondati

In questo paragrafo studiamo l’impatto che geometrie differenti da quella vista in precedenza hanno sull’instabilità dell’RTN. Le geometrie riportate sono trattate in [23] e raffigurate in Fig.1.19. Esse differiscono per il fatto di avere i bordi più o meno arrotondati: il caso (A) è il caso planare già studiato, il (B) riguarda una struttura con i bordi arrotondati e il caso (C) una struttura completamente arrotondata. Spesso si aumenta il raggio di curvatura ai bordi del dispositivo per diminuire l’effetto punta del campo elettrico e dunque il current crowding, rendendo la corrente più uniforme. Ciò ha come effetto il minor impatto di siti trappola situati in queste zone, i quali nel caso (A), invece, erano i principali responsabili di *Giant RTN* e perciò delle code della cumulata. Da quanto detto ci aspettiamo una distribuzione cumulativa a basse probabilità più stretta man mano che si adottano raggi di curvatura più grandi. Simulando classicamente, infatti, ciò accade, come visibile in Fig.1.20. Tenendo conto degli effetti quantistici, però, la conduzione del canale appare ben più uniforme di quella simulata classicamente, il che rende le cumulate più strette, in particolare quella riguardante la struttura planare. Dalla Fig.1.20 notiamo che tenendo conto degli effetti quantistici le distribuzioni cumulative per i casi (A) e (B) risultano uguali, mentre solo per il caso (C) si nota un miglioramento.

Riguardo invece l’RDF, il dispositivo planare (A) si “accende” prima rispetto alle strutture (B) e (C), in quanto ai bordi risulta avere una soglia locale molto più bassa. Inoltre, per l’intensificarsi della corrente in queste due regioni, l’RDF ha effetto solo in tali zone e non anche centralmente, come invece accade per i casi (B) e (C). Arrotondando i bordi della cella, infatti, la corrente scorre in zone più estese e la fluttuazione degli atomi droganti risulta impattare di più. Ciò è confermato dai risultati delle simulazioni:  $\sigma_{V_T}(A) = 141mV$ ,  $\sigma_{V_T}(B) = 151mV$  e  $\sigma_{V_T}(C) = 159mV$ . Adottare geometrie con bordi arrotondati abbatta le fluttuazioni di RTN ma ha anche l’effetto di aumentare la dispersione naturale della soglia dovuta all’RDF. I due fenomeni risultano pertanto anti-correlati, come visibile in Fig.1.21. In questa figura si fa anche riferimento ai risultati di simulazioni per dispositivi “Sourrounding”. Questi dispositivi (Fig.1.22) sono stati simulati per mostrare come le diverse capacità FG-canale per le tre strutture auto-allineate viste in 1.19 non impattano sui risultati discussi.

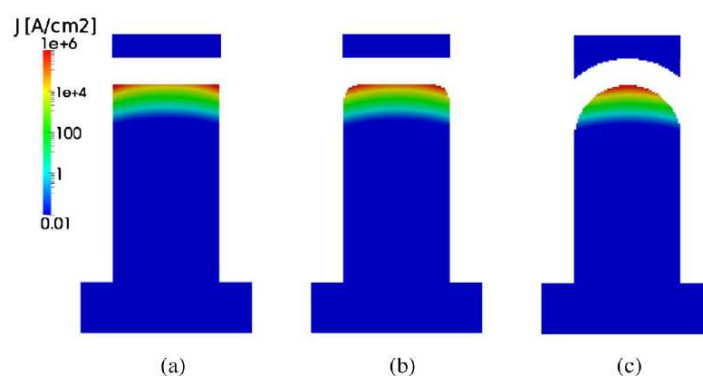


Figura 1.19: Cross section lungo  $W$  delle tre strutture analizzate: caso planare (a), struttura con bordi arrotondati (b) e completamente arrotondata (c)

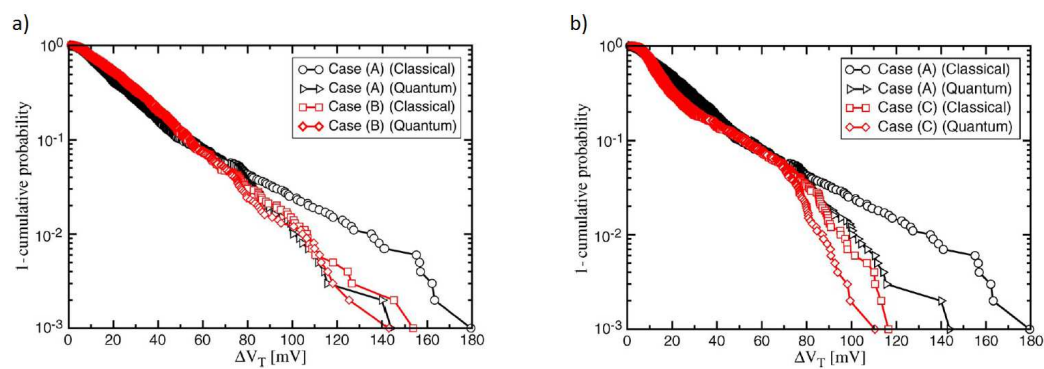


Figura 1.20: a) distribuzione di probabilità cumulativa ottenute da simulazioni classiche e quantistiche per casi (B) e (A) della Fig.1.19. b) Analogico per casi (C) e (A) della Fig.1.19

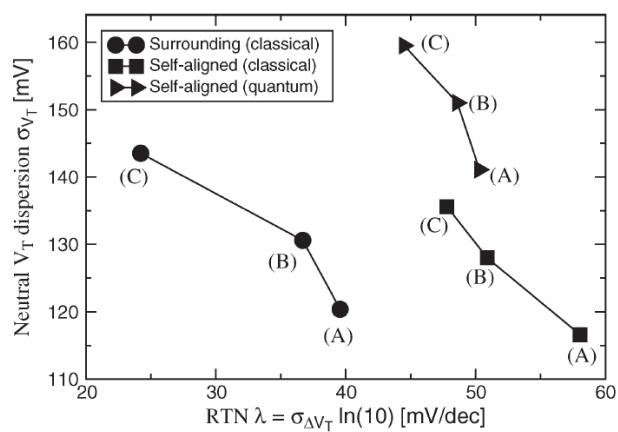


Figura 1.21: dispersione della  $V_T$  in funzione della  $\lambda$ . Le strutture simulate sono le “Self-aligned” di Fig.1.19 e le “Sourounding” di Fig.1.22

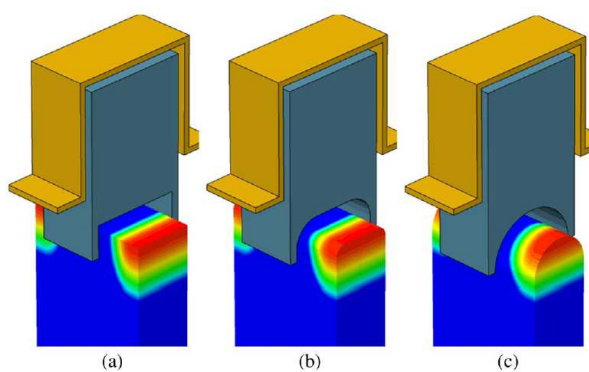


Figura 1.22: Strutture analoghe a quelle della Fig.1.19, con stavolta il FG abbracciante il canale per mantenere la stessa capacità FG-channel per i tre casi

# Capitolo 2

## Metodologie di simulazione

### 2.1 Introduzione

Nei dispositivi ultrascalati il numero degli atomi di drogante presenti nell'impianto di canale è particolarmente esiguo: ad esempio, in un nMOSFET con area attiva quadrata, larghezza e lunghezza efficace di canale  $W_{eff} = L_{eff} = 65nm$ , drogaggio  $N_A = 10^{18}cm^{-3}$  e zona svuotata sotto l'ossido di gate  $W_{dep} = 40nm$ , sono presenti mediamente 170 atomi accettori (tipicamente boro). Dal momento che l'impiantazione sotto l'area attiva causa una dispersione del numero e della posizione degli accettori, avremo una fluttuazione statica della tensione di soglia e una dispersione dell'ampiezza dell'RTN non trascurabili. Come abbiamo visto nel capitolo precedente, infatti, la natura discreta del drogaggio risulta essere la maggior causa delle fluttuazioni statistiche per i dispositivi deca-nanometrici.

In particolare, data la rilevanza e l'enorme impatto delle fluttuazioni dell'ampiezza dell'RTN sulla progettazione dei dispositivi, è necessario prevedere come saranno distribuiti i valori delle  $\Delta V_T$ . Come abbiamo visto, però, i modelli analitici che descrivono questo fenomeno risultano essere inadatti ed insufficienti per dispositivi scalati, in quanto essi trattano la corrente come un flusso continuo ed uniforme, quando in realtà la conduzione risente moltissimo della natura atomistica del drogaggio e degli effetti di bordo. Per questo motivo è necessario ricorrere a simulazioni 3D di tipo numerico per tener conto con la giusta precisione di tali effetti.

## 2.2 Metodo Monte Carlo

La metodologia più semplice per la costruzione di una statistica delle  $\Delta V_T$  è il procedimento di tipo *Monte Carlo* (abbreviato in “MC”) [19][20][21][22]. Questo metodo prevede di costruire una statistica delle  $\Delta V_T$  simulando con il modello *Drift-Diffusion* (acronimo “DD”) tutti i dispositivi “atomistici” (facenti parte della statistica stessa) con la trappola occupata e con la trappola scarica. Questo metodo ha il forte vantaggio di essere fisicamente accurato, data la validità del modello DD ampiamente riconosciuta, ma estremamente lento perché richiede un gran numero di simulazioni 3D.

### 2.2.1 Il Modello Drift-Diffusion

Il modello *Drift-Diffusion* è utilizzato per la simulazione numerica dei dispositivi a semiconduttore fin dall’inizio degli anni ’80, sia in ambito industriale, sia in ambito accademico. Esso rappresenta il momento di ordine zero nell’espansione dell’equazione del trasporto di Boltzmann (BTE Boltzmann Transport Equation) ed è basato su tre equazioni fondamentali: l’equazione di Poisson (2.1) e le equazioni di continuità per gli elettroni (2.2) e le lacune (2.3).

$$\nabla \epsilon \nabla \phi = -q(p - n + N_D - N_A) - \rho_{trap} \quad (2.1)$$

$$\nabla \vec{J}_n = q \frac{\partial n}{\partial t} + qR \quad (2.2)$$

$$\nabla \vec{J}_p = -q \frac{\partial p}{\partial t} - qR \quad (2.3)$$

dove  $n$  e  $p$  sono le concentrazioni di elettroni e lacune,  $N_A$  e  $N_D$  le concentrazioni di droganti, rispettivamente accettori e donori,  $\phi$  è il potenziale elettrostatico,  $\epsilon$  la costante dielettrica del mezzo,  $\rho_{trap}$  la densità di carica distribuita fra cariche fisse e trappole,  $q$  la carica elementare ed infine  $R$  rappresenta il tasso di generazione/ricombinazione fra elettroni e lacune. Le densità di corrente di elettroni e lacune  $\vec{J}_n$  e  $\vec{J}_p$ , la cui somma è ovviamente la corrente totale  $\vec{J}_{tot}$ , sono date

dalle relazioni costitutive che descrivono i principali meccanismi di trasporto nei semiconduttori: la deriva (drift, prima parte delle equazioni 2.4 e 2.5), causata dal campo elettrico, e la diffusione, originata dalla presenza di gradienti di concentrazione dei portatori (seconda parte delle 2.4 e 2.5)

$$\vec{J}_n = q\mu_n n \vec{E} + qD_n \nabla n \quad (2.4)$$

$$\vec{J}_p = q\mu_p p \vec{E} - qD_p \nabla p \quad (2.5)$$

in cui  $\vec{E}$  è il campo elettrico e  $D_n$  e  $D_p$  sono i coefficienti di diffusione, legati alla mobilità e alla tensione termica  $V_{th}$  tramite la relazione di Einstein:

$$D_{n,p} = V_{th}\mu_{n,p} \quad (2.6)$$

Il modello *Drift-Diffusion* è quindi classico, nel senso che descrive attraverso equazioni della fisica classica il trasporto di carica nel semiconduttore, ma si può avvalere in alcuni casi di termini di correzione quantistica. Le simulazioni in questo lavoro sono state realizzate con questo modello su una struttura tridimensionale di una cella Flash opportunamente dimensionata.

### 2.2.2 Il Metodo

Per modellizzare l'effetto della natura discreta del doping si è seguita la metodologia introdotta da A. Asenov [24] del drogaggio "atomistico" nella regione di svuotamento sottostante il canale, in cui l'effetto della discretizzazione influenza maggiormente le caratteristiche del dispositivo, mentre le restanti zone del device vengono trattate con un drogaggio continuo (Fig.2.1). Portato all'estremo, questo metodo associa alle posizioni di ciascun atomo drogante una funzione a "delta" di Dirac di concentrazione di accettori  $N_A$ , il cui integrale sia uguale alla densità di carica di un singolo atomo. Dal momento, però, che una simile funzione (anzi,



distribuzione) non è descrivibile in un calcolatore a causa della risoluzione finita della griglia (e non sarebbe neanche fisicamente rigorosa), essa viene approssimata con una concentrazione finita diversa da zero solo in un piccolo volume di spazio centrato su un nodo della mesh. Il simulatore commerciale utilizzato in questo lavoro [29] presenta la possibilità di piazzare direttamente il doping discreto specificando la posizione degli atomi droganti. L'implementazione di ciò è fatta concentrando lo ione accettore in un vertice della mesh, il quale è il centro del proprio *Voronoi box*. In tale volumetto la concentrazione è scelta in modo che il suo integrale restituisca la carica di un drogante ionizzato  $N = 1/V_{box}$ , ovvero coincide con l'inverso del volume del *Voronoi box*. Il resto della regione di svuotamento ha invece drogaggio  $N = 0$ . La generazione di una mesh adattativa permette di infittire i nodi della griglia in corrispondenza di queste zone, che a livello pratico sono cubi di 2 nanometri di lato. Qui di seguito riporto i passaggi eseguiti per la composizione della procedura costruttrice la statistica dell'RTN.

Inizialmente si crea la struttura e la geometria del dispositivo con il proprio drogaggio continuo-ideale. Da esso un largo numero di dispositivi "atomistici" vengono generati per ciascun profilo di drogaggio rimpiazzando il doping continuo nella regione svuotata sotto il gate con atomi discreti, il cui numero totale è in accordo con la statistica di Poisson e la loro distribuzione spaziale in accordo con il profilo continuo, come mostrato in Fig.2.1. Dopodiché viene calcolata la transcaratteristica  $I_D - V_{CG}$  per ciascuna cella risolvendo numericamente le equazioni di Poisson e drift-diffusion per  $V_D = 0.1V$  (nel nostro caso) ed incrementando  $V_{CG}$ , calcolando poi la tensione di soglia  $V_T^0$  come la tensione  $V_{CG}$  per la quale la corrente al drain raggiunge  $I_{dth} = 100nA \frac{W}{L}$ . Infine per ogni cella si posiziona in modo randomico una singola carica negativa all'interfaccia *Si/Ox* per riprodurre lo stato in cui abbiamo l'intrappolamento di un elettrone da parte di un sito trappola e calcoler poi la transcaratteristica  $I_D - V_{CG}$  e la  $V_T^-$  corrispondente. In questo modo abbiamo la  $\Delta V_T = V_T^- - V_T^0$  per ogni cella, ovvero per ogni coppia *configurazione di drogaggio - trappola*, e da questi valori possiamo costruirci la distribuzione di probabilità cumulativa.

Il metodo MC è stato totalmente implementato (costruzione della geometria, della mesh, simulazione fisica ed elaborazione dei risultati) in "Sentaurus Workbench" (acronimo swb) attraverso i diversi tool della Synopsys (in particolare "Sentaurus Device", abbreviato in Sdevice, calcola la corrente di drain ad ogni bias point risolvendo numericamente il modello DD).

La cella Flash investigata in questo lavoro è mostrata in Fig.2.1: spessore dell'ossido di tunnel pari a 8nm, floating gate in polisilicio a 70nm, stack di dielettrico interpoly ONO (oxide-nitride-oxide) di spessore 4-3-5nm, lunghezza  $L$  e larghezza  $W$  della cella settate a 32nm ed il profilo di drogaggio donore continuo e gaussiano con picco di concentrazione di  $10^{20}cm^{-3}$ . Come già detto, per quanto riguarda il profilo di doping accettore e la posizione della trappola (qualora ci fosse) è stata applicata la procedura *Monte Carlo*. Questo device non è indicativo di alcun processo di produzione o tecnologico, ma è stato creato appositamente per studiare la dispersione statistica della  $\Delta V_T$ .

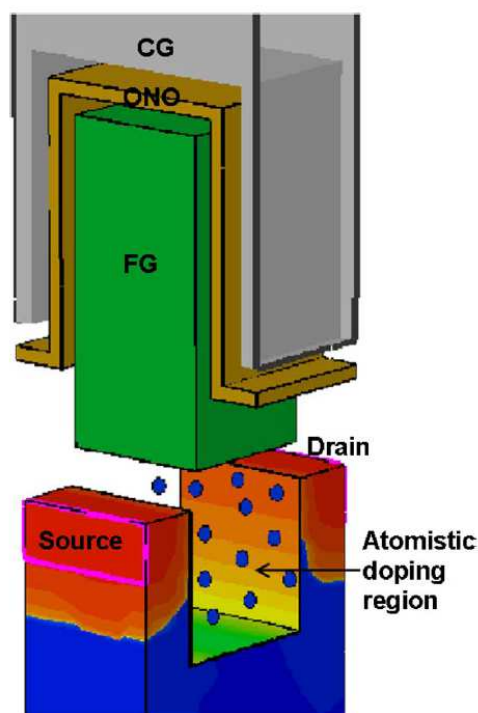


Figura 2.1: Schematico per il template della cella Flash a floating-gate investigata in questo lavoro. La regione di doping atomistico è evidenziata, mentre le regioni di ossido non sono mostrate.

Il metodo *Monte Carlo* è anche rozzamente chiamato “metodo a forza bruta” in quanto per costruirsi la statistica simula con il modello drift-diffusion tutti i dispositivi facenti parte della statistica stessa. Ciò rende il metodo molto preciso, ma estremamente lento e “time consuming”. Il largo numero di dispositivi pre-

sentì in un chip, inoltre, richiede l'esplorazione della distribuzione statistica dell'ampiezza delle fluttuazioni dell'RTN sotto probabilità molto basse, difficilmente raggiungibili con il metodo appena descritto. Con il procedimento *Monte Carlo*, infatti, per avere  $N$  valori dello shift di soglia ho bisogno di  $2 \times N$  simulazioni ( $N$  per i dispositivi con trappola neutra,  $N$  per i dispositivi con trappola riempita). Essendo che ogni simulazione richiede un tempo di  $\approx 10 - 15min$  e che  $N$  dev'essere un numero elevatissimo, questo metodo non è assolutamente praticabile. Ciò ha spinto la ricerca di metodi numerici alternativi, che richiedono meno sforzo computazionale. In questa tesi mostro come utilizzando il metodo sIFM (descritto nella prossima sezione) offro un pratico e conveniente compromesso tra accuratezza e sforzo computazionale nell'analisi statistica dell'ampiezza dell'RTN.

In questo paragrafo ho descritto la procedura MC per l'RTN. Il procedimento per l'RDF è analogo, con la differenza che non ho due variabili aleatorie (configurazione di drogaggio e trappola) da cui costruire gli  $N$  dispositivi, ma solamente una (i profili di doping), in quanto la trappola è da considerarsi scarica. Dopo averli simulati ed aver calcolato la soglia per ognuno di essi, si calcola la  $\sigma_{V_T}$  totale risultante.

## 2.3 Impedance Field Method

A causa del Random Dopant Fluctuation, dispositivi nominalmente identici risultano avere comportamenti elettrici diversi. Se parliamo di transistori o dispositivi Floating Gate vediamo che le transcaratteristiche (ovvero le curve  $I_d - V_{gs}$ ) risultano particolarmente “shiftate” l’una con l’altra. Per questo motivo possiamo pensare l’RDF come una fluttuazione statica, ovvero come un rumore in continua (a frequenza nulla).

L’Impedance Field Method (abbreviato con “IFM”) è un metodo intelligente per la modellizzazione delle fluttuazioni statiche ai contatti del dispositivo dovute a perturbazioni locali di vario genere, come la fluttuazione degli atomi droganti in numero e posizione (RDF), la fluttuazione della WorkFunction, le fluttuazioni geometriche del dispositivo etc ...

In questo lavoro di tesi useremo l’IFM solo per la prima (e più importante) fonte di variabilità statistica di quelle elencate poc’ anzi, ovvero per il Random Dopant Fluctuation.

Il concetto base dietro l’IFM sta nel trattare le varie sorgenti di fluttuazioni come una perturbazione del dispositivo di riferimento (il dispositivo con drogaggio continuo). Dopo aver risolto l’equazione non lineare di Poisson e le equazioni drift-diffusion per il dispositivo di riferimento, si usa la soluzione del potenziale elettrico per calcolarsi la risposta impulsiva ai contatti per una piccola perturbazione locale di drogaggio in ogni punto del dispositivo. Questa risposta impulsiva è la *Green Function* o *Impedance Field* per l’RDF  $\mathbf{G}_c^s(r)$ : dove  $c$  è il contatto considerato ed  $s$  la specie di drogaggio. Data una configurazione atomistica di drogaggio  $N_\nu(r)$  si può risalire alla deviazione di corrente dal dispositivo di riferimento (con doping continuo) attraverso la funzione di Green e le eq.2.7-2.8. Da queste equazioni possiamo interpretare la  $\mathbf{G}_c^s(r)$  come la “funzione peso” o la “sensitivity” del doping: ci dice quanto una variazione di drogaggio in quel punto si ripercuote al contatto considerato in termini di corrente o di tensione (a seconda di quale Green Function stiamo considerando).

$$\delta I_{\nu,c} = \sum_s \int \mathbf{G}_c^s(r) \delta N_\nu(r) d^3r \quad (2.7)$$

$$\delta N_\nu^s(r) = N_\nu^s(r) - N_{ref}^s(r) \quad (2.8)$$

Sdevice randomizza il drogaggio (col numero di atomi droganti seguente la statistica di Poisson e la posizione di essi in accordo con il drogaggio continuo) creando i diversi dispositivi “atomistici”. Inizialmente per ogni dispositivo  $\nu$  sdevice calcola il discostamento di drogaggio  $\delta N_\nu^s(r)$  dalla configurazione continua di doping  $N_{ref}^s(r)$  secondo l’equazione 2.8, e, inoltre, fissato il bias point  $V_{gs}$  si calcola la  $\mathbf{G}_c^s(r)$  dalla soluzione DD del dispositivo di riferimento. Dopodiché calcola l’integrale di volume 2.7 per ogni dispositivo  $\nu$ , ottenendo così i vari discostamenti di corrente al contatto  $c$  rispetto al dispositivo di riferimento a bias point fissato. Questo procedimento viene ripetuto ad ogni step di  $V_{gs}$ .

Il grande vantaggio di questo metodo sta nel fatto che risolve il modello drift-diffusion (DD), e perciò un sistema di equazioni differenziali alle derivate parziali fortemente non lineari, solo una volta (ovvero solo per il dispositivo di riferimento con drogaggio continuo) e non per ognuna delle configurazioni di drogaggio come invece adoperava il MC. Infatti l’IFM, dopo aver calcolato la  $\mathbf{G}_c^s(r)$ , semplicemente calcola mediante un integrale di volume il discostamento di corrente per ogni configurazione  $\nu$  dal dispositivo di riferimento. Il metodo è un metodo lineare, di “piccolo segnale”, inoltre la Green Function non dipende dalla configurazione atomistica di drogaggio randomizzata (ma dipende solo dal drogaggio continuo di riferimento!), e viene infatti calcolata, fissato il bias point, solo una volta indipendentemente da quanti profili atomistici di doping voglio considerare. Da quanto spiegato, possiamo dire che la  $\delta \mathbf{I}_{\nu,c}$  è **la deviazione della corrente in risposta lineare alla perturbazione  $\delta \mathbf{N}_\nu^s(\mathbf{r})$** . Il metodo risulta approssimato e al contempo molto veloce. [1] [2] [3] [4]

L’Impedance Field Method, in realtà, ha diverse versioni per la modellizzazione dell’RDF: il “noise-like” IFM e lo “statistical” IFM. La prima tratta questa dispersione statistica come un rumore, inteso proprio come rumore in frequenza, mentre la seconda fa uso dell’equazione 2.7 per la ricostruzione delle transcaratteristiche dei singoli dispositivi. In questo lavoro useremo questa seconda versione del metodo per la simulazione dell’RTN. Per quanto riguarda l’RDF, invece, le due varianti del metodo danno gli stessi risultati in termini di dispersione di corrente e di tensione, risultando pertanto equivalenti.

### 2.3.1 Metodo di risoluzione della funzione di Green

In questo paragrafo mostrerò le basi teoriche del metodo della funzione di Green per la risoluzione di equazioni differenziali lineari, ed alcune sue semplici applicazioni analitiche.

Consideriamo un operatore differenziale lineare nelle variabili  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  non singolare  $L_{\mathbf{x}}$ , associato ad un'equazione differenziale lineare:

$$L_{\mathbf{x}}[u(\mathbf{x})] = f(\mathbf{x}) \quad (2.9)$$

Analogamente a quanto accade per l'algebra lineare, vorremmo calcolare "l'operatore inverso"  $L_{\mathbf{x}}^{-1}$  tale che ci dia la funzione soluzione del problema:

$$u(\mathbf{x}) = L_{\mathbf{x}}^{-1}[f(\mathbf{x})] \quad (2.10)$$

Ora consideriamo il problema più generale:

$$L_{\mathbf{x}}[u(\mathbf{x})] = \sum_k f_k(\mathbf{x}) \quad (2.11)$$

Se  $u_k(\mathbf{x})$  è la soluzione del "problema parziale"

$$L_{\mathbf{x}}[u_k(\mathbf{x})] = f_k(\mathbf{x}) \quad (2.12)$$

la linearità del problema assicura che la soluzione è

$$u(\mathbf{x}) = \sum_k u_k(\mathbf{x}) \quad (2.13)$$

Questo suggerisce di usare una base per le possibili funzioni  $f$ , ed una base siffatta è fornita da (attenzione, qui  $\mathbf{y}$  è un indice e non denota derivazione):

$$f_{\mathbf{y}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{y}) := \delta(x_1 - y_1) \dots \delta(x_n - y_n) \quad (2.14)$$

In effetti, ogni funzione  $f(\mathbf{x})$  può essere scritta come:

$$f(\mathbf{x}) = \int A(\mathbf{y}) f_{\mathbf{y}}(\mathbf{x}) d\mathbf{y} = \int A(\mathbf{y}) \delta(\mathbf{x} - \mathbf{y}) d\mathbf{y} \quad (2.15)$$

scegliendo come "coefficienti" la funzione  $A(\mathbf{y}) = f(\mathbf{y})$ .

In questo senso, ogni problema (lineare) non omogeneo può essere ridotto alla considerazione di problemi non omogenei in cui il termine noto sia una funzione

delta. Queste idee euristiche trovano la loro espressione nella teoria (e nel metodo) della funzione di Green.

Ciò detto, considerando il problema iniziale 2.9, la **funzione di Green** viene definita come quella funzione associata all'operatore  $L_x$  che risulta soluzione del problema associato:

$$L_x G(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}) \quad (2.16)$$

A questo punto ci viene in aiuto un importante teorema che lega la funzione di Green alla soluzione del nostro problema lineare.

**Teorema.** *Se la funzione  $G(x, y)$  è soluzione della 2.16, allora la funzione 2.17 è soluzione della 2.9 (problema formulato in 1-D per semplicità).*

$$u(x) = \int_{-\infty}^{+\infty} G(x, y) f(y) dy \quad (2.17)$$

**Dimostrazione.** *Dato che  $L_x$  agisce sulle variabili  $x$  ma non sulle  $y$ ,*

$$L_x u(x) = \int_{-\infty}^{+\infty} L_x [G(x, y)] f(y) dy = \int_{-\infty}^{+\infty} \delta(x - y) f(y) dy = f(x) \quad (2.18)$$

**Osservazione.** *La soluzione più generale della 2.16 si scriverà come:*

$$G(x, y) = g(x, y) + \sum_{k=1}^n c_k(y) \varphi_k(x, y) \quad (2.19)$$

dove  $g(x, y)$  è una soluzione particolare di 2.16, detta anche parte singolare della funzione di Green, le  $\varphi_k(x, y)$  sono  $n$  soluzioni (indipendenti, con  $n$  l'ordine di  $L_x$  come operatore differenziale se  $x$  è uno-dimensionale) della omogenea della 2.16 e le  $c_k(y)$  sono determinate dalle condizioni al contorno.

## Operatori del prim'ordine

Vediamo cosa succede per operatori del prim'ordine molto semplici, con  $x$ -uno dimensionale:

$$L_x = a(x) \frac{d}{dx} + b(x) \quad (2.20)$$

se  $a(x)$  e  $b(x)$  sono finite e  $a(x) \neq 0 \forall x$ , diremo che l'operatore è regolare; assumiamo che sia questo il caso.

La soluzione generale di

$$L_x[\varphi(x)] = 0 \quad (2.21)$$

risulta essere, come ben noto,

$$\varphi(x) = \varphi_0 \exp \left[ - \int_{x_0}^x \frac{b(z)}{a(z)} dz \right] \quad (2.22)$$

dove  $\varphi_0$  è una costante. La soluzione di

$$L_x[G(x, y)] = \delta(x - y) \quad (2.23)$$

risulterà quindi essere

$$G(x, y) = g(x, y) + C(y) \exp \left[ - \int_{x_0}^x \frac{b(z)}{a(z)} dz \right] = g(x, y) + C(y) e^{-F(x, x_0)} \quad (2.24)$$

dove ho scritto per comodità  $F(x, x_0) = \left[ \int_{x_0}^x -\frac{b(z)}{a(z)} dz \right]$ . Concentriamoci ora sulla soluzione particolare  $g(x, y)$ ; per  $x \neq y$  la 2.23 si riduce alla 2.21, quindi la soluzione particolare della 2.23, sarà nella forma:

$$g(x, y) = \begin{cases} A_-(y) e^{-F(x, y)} & x < y \\ A_+(y) e^{-F(x, y)} & x > y \end{cases} \quad (2.25)$$

Abbiamo quindi (con  $\epsilon > 0$ ) che

$$\lim_{\epsilon \rightarrow 0} g(y \pm \epsilon, y) = A_{\pm}(y) \quad (2.26)$$

I valori delle costanti (rispetto alla variabile  $x$ )  $A_{\pm}(y)$  non sono arbitrari, ma sono legati da una relazione che discende dalla 2.23. Infatti

$$[g(x + \epsilon, y) - g(x - \epsilon, y)] = \int_{x-\epsilon}^{x+\epsilon} \frac{\partial g(\alpha, y)}{\partial \alpha} d\alpha \quad (2.27)$$

d'altra parte, la 2.23 assicura che

$$\frac{\partial g(x, y)}{\partial x} = \frac{1}{a(x)} \delta(x - y) + \frac{b(x)}{a(x)} g(x, y) \quad (2.28)$$

e quindi integrando la 2.28 in  $dx$  e specializzandola per  $x = y$  otteniamo

$$[g(y + \epsilon, y) - g(y - \epsilon, y)] = \int_{y-\epsilon}^{y+\epsilon} \frac{\partial g(x, y)}{\partial x} dx = \frac{1}{a(y)} \quad (2.29)$$



Quindi abbiamo dimostrato l'uguaglianza

$$A_+(y) - A_-(y) = \frac{1}{a(y)} \quad (2.30)$$

Possiamo ora scegliere  $A_{\pm}(y)$  in modo arbitrario - il che corrisponde a cambiare la costante arbitraria  $C(y)$  nella 2.24 - purché la differenza  $A_+(y) - A_-(y)$  sia fissata. Per comodità scegliamo  $A_+(y) = \frac{1}{a(y)}$  e  $A_-(y) = 0$ , che porta ad avere

$$g(x, y) = \begin{cases} 0 & x < y \\ \frac{1}{a(y)} e^{-F(x,y)} & x > y \end{cases} = \Theta(x - y) \frac{1}{a(y)} e^{-F(x,y)} \quad (2.31)$$

Abbiamo dimostrato che per l'equazione differenziale

$$a(x)u'(x) + b(x)u(x) = f(x) \quad (2.32)$$

la funzione di Green corrispondente risulta essere

$$G(x, y) = \left[ \frac{\Theta(x - y)}{a(y)} + C(y) \right] \exp \left[ - \int_y^x \frac{b(z)}{a(z)} dz \right] \quad (2.33)$$

Quanto detto è verificabile facilmente, integrando la 2.17 con la funzione di Green 2.33

$$\begin{aligned} u(x) &= \int_{-\infty}^{+\infty} G(x, y) f(y) dy \\ &= \alpha \exp \left[ \int_x^{x_0} \frac{b(z)}{a(z)} dz \right] + \int_{x_0}^x \frac{f(y)}{a(y)} \left( \exp \left[ \int_x^y \frac{b(z)}{a(z)} dz \right] \right) dy \end{aligned} \quad (2.34)$$

che è la ben nota formula per la soluzione generale della equazione lineare del primo ordine.

### Operatori del second'ordine

Ora vediamo cosa succede per operatori del second'ordine, con x-uno dimensionale:

$$L_x = a(x) \frac{d^2}{dx^2} + b(x) \frac{d}{dx} + c(x) \quad (2.35)$$

e la più generale equazione lineare del secondo ordine si scriverà come

$$L_x[u(x)] = a(x)u''(x) + b(x)u'(x) + c(x) = f(x) \quad (2.36)$$

Assumeremo che  $L_x$  sia regolare, cioè che le funzioni  $a(x)$ ,  $b(x)$  e  $c(x)$  siano limitate, con inoltre  $a(x) \neq 0$  per qualsiasi  $x$ .

Un'equazione del tipo 2.36, come è ben noto dai corsi base di Analisi, si può sempre ridurre in "forma canonica"

$$u''(x) + V(x)u(x) = \Phi(x) \quad (2.37)$$

e perciò ci si limita a studiare operatori del second'ordine del tipo

$$L_x = \frac{d^2}{dx^2} + V(x) \quad (2.38)$$

Ricordiamo che sia per l'equazione generale  $L_x[u(x)] = f(x)$  che per la omogenea associata, non è possibile scrivere una formula generale di soluzione; non è quindi possibile pensare di determinare la più generale  $G(x, y)$ . D'altra parte sarà possibile determinare  $G(x, y)$  una volta che si conoscano due soluzioni indipendenti dell'omogenea associata. Conosciute tali soluzioni, infatti, abbiamo un importante teorema che ci dice come calcolare la funzione di Green.

**Teorema.** *Sia  $L_x$  l'operatore definito dalla 2.38. Se  $\phi_1(x)$  e  $\phi_2(x)$  sono due soluzioni indipendenti della omogenea  $L_x[u(x)] = 0$  e  $W(x)$  il loro Wronskiano, allora la funzione di Green  $G(x, y)$  per l'operatore  $L_x$  è fornita da:*

$$G(x, y) = \frac{A_1(y)\phi_1(x) + A_2(y)\phi_2(x) - \Theta(x-y)\phi_1(x)\phi_2(y) + \Theta(y-x)\phi_1(y)\phi_2(x)}{W(x)} \quad (2.39)$$

**Dimostrazione.** *Il Wronskiano di  $m$  funzioni è il determinante della matrice  $m \times m$  che ha nella riga  $k$ -ma le derivate di ordine  $k-1$  delle funzioni; nel caso di due funzioni si ha 2.40. È noto, poi, come tale funzione sia una costante non banale, come dimostrato dall'eq.2.41.*

$$W(x) = \phi_1(x)\phi_2'(x) - \phi_1'(x)\phi_2(x) \quad (2.40)$$

$$W'(x) = \phi_1(x)\phi_2''(x) - \phi_1''(x)\phi_2(x) = V(x)(\phi_1\phi_2 - \phi_1\phi_2) = 0 \quad (2.41)$$

Usando l'espressione 2.40 e ricordando che  $\Theta'(x) = \delta(x) = \delta(-x)$ , possiamo calcolare la derivata seconda rispetto ad  $x$  della funzione di Green

$$\begin{aligned} \frac{\partial^2 G(x, y)}{\partial x^2} &= A_1(y)\phi_1''(x) + A_2(y)\phi_2''(x) \\ &\quad - \frac{\Theta(x-y)\phi_1''(x)\phi_2(y) + \Theta(y-x)\phi_1(y)\phi_2''(x)}{W(x)} \\ &\quad - \delta(x-y)\frac{\phi_1'(x)\phi_2(y) - \phi_1(y)\phi_2'(x)}{W(x)} \end{aligned} \quad (2.42)$$

Ricordando che  $\phi_i'' = -V\phi_i$  e l'espressione della  $G(x, y)$  (2.39), la somma dei primi due termini risulta  $= -V(x)G(x, y)$ ; mentre esplicitando la funzione  $W(x)$ , secondo la 2.40, il terzo termine implode in  $\delta(x-y)$ . Pertanto la 2.42 risulta essere:

$$\frac{\partial^2 G(x, y)}{\partial x^2} = -V(x)G(x, y) + \delta(x-y) \quad (2.43)$$

che vuol dire  $L_x[G(x, y)] = \delta(x-y)$ , ovvero che la  $G(x, y)$  è realmente la funzione di Green per l'operatore  $L_x$  dato dalla 2.38.

### Equazioni a derivate parziali

Il modello DD, però, è un sistema di equazioni alle derivate parziali, e, dunque, dobbiamo complicare ulteriormente la nostra analisi. Consideriamo per questo operatori di derivazione parziale del secondo ordine a coefficienti costanti. Avremo in generale le variabili  $(x_1, \dots, x_n)$  e scriveremo

$$L_{\mathbf{x}} = a_0 + \sum_{i=1}^n a_i \frac{\partial}{\partial x_i} + \sum_{i=1}^n b_i \frac{\partial^2}{\partial x_i^2} \quad (2.44)$$

$$L_{\mathbf{x}}[u(\mathbf{x})] = f(\mathbf{x}) \quad (2.45)$$

La funzione di Green in questo caso deve soddisfare la

$$L_{\mathbf{x}}[G(\mathbf{x}, \mathbf{y})] = \delta_n(\mathbf{x} - \mathbf{y}) \quad (2.46)$$

Ricordiamo che

$$\delta(x-y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp[ik(x-y)] dk \quad (2.47)$$

ed in analogia che  $\delta_n(\mathbf{x}-\mathbf{y}) := \delta(x_1-y_1) \dots \delta(x_n-y_n)$ . La delta multidimensionale si può rappresentare pertanto come

$$\delta_n(\mathbf{x} - \mathbf{y}) = \frac{1}{(2\pi)^n} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \exp \left[ i \sum_{j=1}^n k_j (x_j - y_j) \right] d\mathbf{k} \quad (2.48)$$

dove  $d\mathbf{k} = dk_1 \dots dk_n$ . Definiamo per semplicità le funzioni  $P(\mathbf{k}, \mathbf{x} - \mathbf{y})$  e  $Q(\mathbf{k})$  come:

$$P(\mathbf{k}, \mathbf{x} - \mathbf{y}) = \exp \left[ i \sum_{j=1}^n k_j (x_j - y_j) \right] \quad (2.49)$$

$$Q(\mathbf{k}) = a_0 + \sum_{j=1}^n a_j (ik_j) + \sum_{j=1}^n b_j (ik_j)^2 \quad (2.50)$$

**Teorema.** *Nel caso di operatori di secondo ordine 2.44 puramente differenziali ( $a_0 = 0$ ), una soluzione particolare della 2.46, e dunque la parte singolare  $g(\mathbf{x}, \mathbf{y})$  della funzione di Green per l'operatore  $L_{\mathbf{x}}$ , è fornita da*

$$g(\mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^n} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \frac{P(\mathbf{k}, \mathbf{x} - \mathbf{y})}{Q(\mathbf{k})} d\mathbf{k} \quad (2.51)$$

### 2.3.2 Funzione di Green nel metodo IFM

Ora, sempre considerando il problema 2.9, decidiamo di perturbare il termine noto  $f(x)$  in  $f(x) + \delta f(x)$ . Quindi

$$L_x[u(x) + \delta u(x)] = f(x) + \delta f(x) \quad (2.52)$$

che, data la linearità dell'operatore  $L_x$ , diventa

$$L_x[\delta u(x)] = \delta f(x) \quad (2.53)$$

Possiamo adesso immaginare di applicare questa teoria all'equazione di Poisson:  $L_x$  rappresenta la 2.1,  $\delta f(x) = \delta N_\nu(x)$  è la perturbazione di concentrazione di doping dal drogaggio continuo per il dispositivo  $\nu$ , e  $\delta u(x) = \delta V_\nu(x)$  è la deviazione del potenziale elettrico risultante.

Il problema, però, è tutt'altro che lineare: l'equazione di Poisson risulta fortemente non lineare data la dipendenza esponenziale delle concentrazioni di portatori dal potenziale. L'idea, perciò, è quella di **linearizzare** il modello in un intorno della soluzione esatta. Per ogni bias point si calcola la soluzione esatta con il

modello DD, dopodiché si sviluppa il modello linearizzato  $L_x$  per l'equazione di Poisson in un intorno di questa soluzione, e da esso si calcola la *Funzione di Green*. Quindi per ogni dispositivo atomistico  $\nu$  si calcola la deviazione approssimata di potenziale nel punto  $x$  come (dalla 2.17):

$$\delta V_\nu(x) = \int G_\phi^V(x, x') \delta N_\nu(x') dx' \quad (2.54)$$

La funzione di Green  $G_\phi^V(x, x')$  descrive la risposta del potenziale  $V$  in  $x$  in seguito ad una perturbazione in  $x'$  nella parte destra dell'equazione di Poisson (causata da  $\delta N_\nu$ ) per la variabile soluzione  $\phi$ .

Provo qui a fare un semplice esempio 1-dimensionale del problema in questione. Supponendo che la costante dielettrica sia costante ovunque e che non ci siano trappole cariche ( $\rho_{trap} = 0 \forall x$ ), l'equazione di Poisson diventa:

$$\epsilon \frac{\partial^2 \phi(x)}{\partial x^2} = -q (p(x) - n(x) + N_D(x) - N_A(x)) \quad (2.55)$$

Ricordando ora che le concentrazioni dipendono dal potenziale in maniera esponenziale, ovvero che  $p(x) \propto e^{(\phi_p - \phi)/V_{th}}$  e che  $n(x) \propto e^{+(\phi - \phi_n)/V_{th}}$ , la concentrazione di doping perturbata  $\overline{N_A(x)} + \delta N_A(x)$  induce la risposta lineare del potenziale  $\overline{\phi(x)} + \delta \phi(x)$  secondo

$$\epsilon \frac{\partial^2 (\overline{\phi(x)} + \delta \phi(x))}{\partial x^2} = -q \left( \overline{p(x)} - \overline{p(x)} \frac{\delta \phi(x)}{V_{th}} + \right. \\ \left. - \overline{n(x)} - \overline{n(x)} \frac{\delta \phi(x)}{V_{th}} + \overline{N_D(x)} - \overline{N_A(x)} - \delta N_A(x) \right) \quad (2.56)$$

che, semplificandola, si traduce in

$$\frac{\partial^2 \delta \phi(x)}{\partial x^2} - \left( \frac{q (\overline{p(x)} + \overline{n(x)})}{\epsilon V_{th}} \right) \delta \phi(x) = \frac{q}{\epsilon} \delta N_A(x) \quad (2.57)$$

dove nella 2.56 ho adottato la linearizzazione al prim'ordine secondo Taylor sulla soluzione  $(\overline{\phi(x)}; \overline{n(x)}; \overline{p(x)})$  del modello DD. Ora, adottando variabili adimensionali si giunge a

$$\frac{\partial^2 (\delta\widehat{\phi}(x))}{\partial x^2} - C(x) (\delta\widehat{\phi}(x)) = \delta\widehat{N}_A(x) \quad (2.58)$$

che coincide perfettamente con l'equazione 2.37, la cui risoluzione è stata già oggetto di discussione.

Detto ciò, faccio notare come trovarsi la risposta  $G(x, y)$  alla perturbazione di drogaggio  $\delta N_A(x) = \delta(x - y)$  voglia dire trovarsi la risposta (in approssimazione lineare) del potenziale in  $x$  dovuta ad un atomo ionizzato posizionato in  $y$  (la cui carica è appunto  $q\delta(x - y)$ ). Possiamo interpretare pertanto la funzione di Green  $G(x, y)$  come la *risposta impulsiva* del potenziale in  $x$  per un atomo drogante in  $y$ , infatti:

$$\delta V_\nu(x) = \int G_\phi^V(x, x') \delta(x' - y) dx' = G(x, y) \quad (2.59)$$

Un procedimento del tutto analogo si può fare per la corrente. Dunque, possiamo calcolare la deviazione di corrente al contatto  $c$   $\delta I_{c,\nu}$  per il dispositivo  $\nu$  ad ogni bias point  $V_g$  dalla funzione di Green  $G_\phi^I(x')$  (ho omesso  $x$ , in quanto coincide con  $c$ ) come 2.60; espressione che coincide con la 2.7.

$$\delta I_{c,\nu} = \int G_\phi^I(x') \delta N_\nu(x') dx' \quad (2.60)$$

Numericamente, fissato il bias point, il simulatore deve calcolare una volta sola la funzione di Green dalla soluzione del modello DD per il dispositivo con drogaggio continuo. Il peso computazionale di ciò è esponenzialmente dipendente dalla mesh. Più la mesh è fine più il simulatore ci metterà tempo per il calcolo di essa. Facendo riferimento al dispositivo studiato in questo lavoro di tesi (2.1), per una mesh in cui la zona svuotata è stata suddivisa in cubi di  $2nm$  di lunghezza per lato il tempo necessario per il calcolo della  $G_\phi^I(x')$  è di 7-9 secondi circa, il che è davvero poco. Una volta calcolata la Green Function, infatti, per calcolare tutte le deviazioni di corrente per i dispositivi atomistici  $\nu$  da quello con drogaggio

continuo, è sufficiente calcolare gli integrali di volume 2.7 (2.60), il che avviene molto velocemente.

Sdevice ci permette di visualizzare la funzione di Green, e un esempio è mostrato dalla figura 2.2. Purtroppo non c'è il modo di dire al simulatore di calcolarsi la funzione solo in una determinata regione, ma ciò non è particolarmente un problema in quanto il tempo di calcolo sarebbe minore solo di poco.

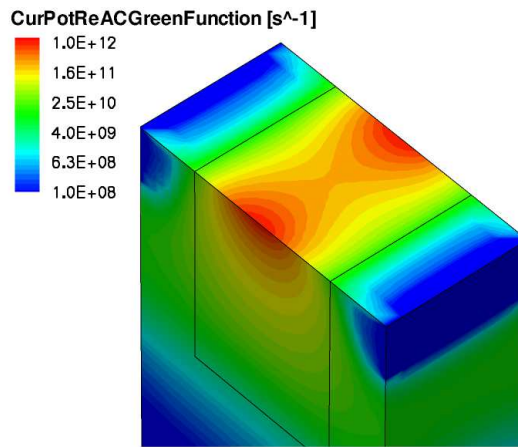


Figura 2.2: funzione di Green per il contatto di drain calcolata a soglia per il dispositivo con drogaggio accettore continuo uniforme pari a  $N_A = 2 \times 10^{18} \text{cm}^{-3}$

La Green Function ci dice molto dell'impatto del drogaggio atomistico sul dispositivo. Come già detto, infatti, più la funzione di Green è alta, più in quel punto una variazione di drogaggio provoca un'ampia deviazione di corrente al contatto. È in sostanza una "funzione peso" del drogaggio. Questo è ben visibile dalla Fig 2.2: abbiamo dei picchi della  $G_\phi^I(x)$  sui bordi del dispositivo, dove infatti una fluttuazione di drogante provoca un'ampia dispersione della  $I_d$  in quanto per effetto del *current crowding* una gran quantità di corrente scorre su tali regioni, rendendole più sensibili a fluttuazioni locali di doping.

Poiché il metodo è lineare, ovvero assume che la dipendenza della corrente sia lineare con il discostamento del doping, la funzione di Green risultante agisce come funzione peso lineare (come il coefficiente angolare di una retta) di questa dispersione, e dunque non dipende dal  $\delta N_\nu$ , ovvero non dipende dal drogaggio

atomistico considerato. È pur vero, però, che il calcolo della funzione di Green è stato fatto sulla soluzione DD del dispositivo ideale di riferimento, e perciò essa dipende fortemente dal doping continuo e dalla posizione della trappola occupata (qualora ci fosse). In 2.3a) è mostrato il dispositivo con trappola neutra, mentre in 2.3b) con trappola carica. Le Green Function a soglia risultanti sono rispettivamente 2.4a) e 2.4b) e sono appunto diverse. La trappola è stata posizionata sul bordo alto del dispositivo, annullando di fatto la corrente in quella zona. Una riduzione della  $I_{loc}$  provoca un abbassamento della sensibilità dal drogaggio in quella zona, e perciò la Green Function, che è una sorta di *sensitivity* del doping, subisce una forte riduzione in tale regione a seguito dell'intrappolamento della carica. Ciò detto è ben visibile in Fig.2.4. La diversità delle Green Function con/senza trappola mi assicura di ottenere  $\Delta V_T \neq 0$

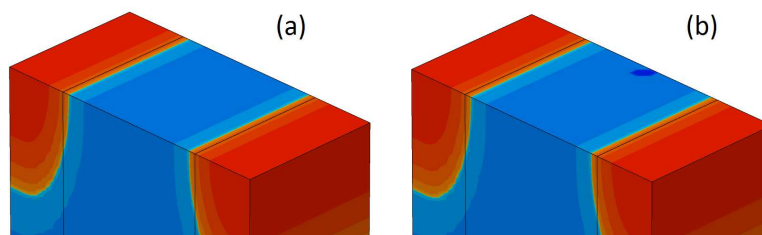


Figura 2.3: concentrazione di doping per dispositivo con drogaggio continuo ed uniforme nel caso di a) trappola neutra e b) trappola carica

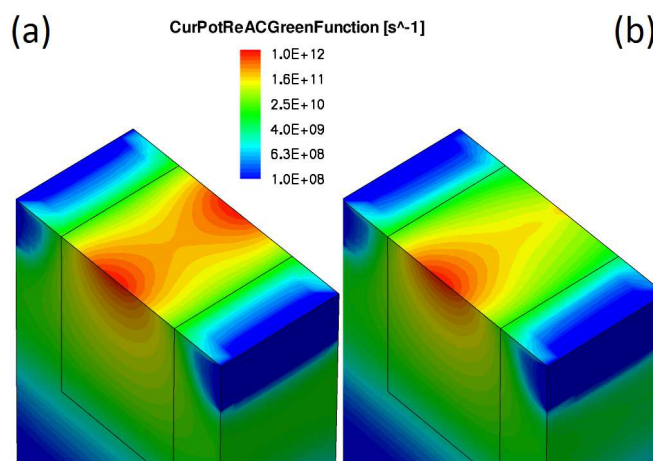


Figura 2.4:  $G_{\phi}^I(r)$  calcolate a soglia per i dispositivi a) e b) della Fig:2.3



### 2.3.3 “Noise-Like” IFM

La variante noise-like IFM (abbreviata nl-IFM) discussa qui modella le fluttuazioni randomiche del drogaggio in maniera molto simile all’analisi di rumore. Per l’RDF la media (campionara) statistica delle fluttuazioni di drogaggio risulta essere nulla in tutto il dispositivo  $\langle \delta N_\nu^s(r) \rangle = 0$ , e, dunque, in approssimazione lineare coerentemente con l’equazione 2.7, la media statistica della fluttuazione di corrente al terminale  $c$  risulta nulla anch’essa  $\langle \delta I_{\nu,c} \rangle = 0$ . Perciò l’informazione sugli effetti statistici dell’RDF in termini di corrente ai contatti è contenuta solo in momenti statistici più alti. Le equazioni 2.61 e 2.62 descrivono rispettivamente il momento di second’ordine della fluttuazione di drogaggio e di corrente:

$$K^s(r_1, r_2) = \langle \delta N_\nu^s(r_1) \delta N_\nu^s(r_2) \rangle = N_{ref}^s(r_1) \delta(r_1 - r_2) \quad (2.61)$$

$$\begin{aligned} \langle \delta I_{\nu,c_1} \delta I_{\nu,c_2} \rangle &= \sum_s \left\langle \int G_{c_1}^s(r_1) \delta N_\nu^s(r_1) d^3r_1 \int G_{c_2}^s(r_2) \delta N_\nu^s(r_2) d^3r_2 \right\rangle \\ &= \sum_s \int \int G_{c_1}^s(r_1) K^s(r_1, r_2) G_{c_2}^s(r_2) d^3r_1 d^3r_2 \end{aligned} \quad (2.62)$$

L’uguaglianza 2.61 è veritiera assumendo che le fluttuazioni randomiche di drogaggio siano spazialmente incorrelate.

Facendo la trasformata di Fourier nel tempo del momento di second’ordine delle  $\delta I_{\nu,c}$  otteniamo la densità spettrale di rumore in corrente che scorre tra i contatti  $c_1$  e  $c_2$ . Qui di seguito è anche riportata, analogamente, la trasformata di Fourier di  $K^s(r_1, r_2)$ :

$$\mathcal{F}\{K^s(r_1, r_2)\} = N_{ref}^s \delta(r_1 - r_2) \delta(\nu) \approx N_{ref}^s \delta(r_1 - r_2) \frac{\Theta(0.5Hz - |\nu|)}{1Hz} \quad (2.63)$$

$$\begin{aligned} S_{c_1,c_2}^I(\nu) &= \langle \delta I_{\nu,c_1} \delta I_{\nu,c_2} \rangle \delta(\nu) \approx \langle \delta I_{\nu,c_1} \delta I_{\nu,c_2} \rangle \frac{\Theta(0.5Hz - |\nu|)}{1Hz} \\ &\approx \sum_s \int \int G_{c_1}^s(r_1) N_{ref}^s(r_1) \delta(r_1 - r_2) \frac{\Theta(0.5Hz - |\nu|)}{1Hz} G_{c_2}^s(r_2) d^3r_1 d^3r_2 \\ &\approx \sum_s \int G_{c_1}^s(r_1) \left( N_{ref}^s(r_1) \frac{\Theta(0.5Hz - |\nu|)}{1Hz} \right) G_{c_2}^s(r_1) d^3r_1 \end{aligned} \quad (2.64)$$

Da notare come il rumore dovuto all'RDF sia un rumore in continua, infatti la 2.64 non è nient'altro che la 2.62 moltiplicata per la delta di Dirac centrata in  $\nu = 0$ . Computazionalmente, però, non si può implementare una delta di Dirac e quindi la si approssima con una funzione "rettangolo" tra -0.5Hz e 0.5Hz come appunto quantitativamente descritto dalle equazioni sopra riportate.

A questo punto per il calcolo della varianza della corrente di drain basta integrare per tutte le frequenze la densità spettrale di rumore in corrente. Nel nostro problema questo si semplifica nella moltiplicazione tra  $S_{c_1, c_2}^I(\nu = 0)$  e 1Hz. La deviazione standard della corrente di drain  $\sigma_{I_d}(V_g)$  risulterà essere la radice quadrata di ciò:

$$\sigma_{I_d}(V_g) = \sqrt{S_{c_1, c_2}^I(0)1Hz} \quad (2.65)$$

Quest'equazione quantifica la dispersione della corrente di drain per l'RDF al bias point  $V_g$ . Il passaggio alla dispersione della tensione di gate passa attraverso la transconduttanza tra drain e gate  $y_{d,g}$ :

$$\sigma_{V_g}(V_g) = \frac{\sigma_{I_d}(V_g)}{y_{d,g}} \quad (2.66)$$

Per questo metodo approssimato, la  $y_{d,g}$  è la transconduttanza tra drain e gate del dispositivo di riferimento, ovvero quello con drogaggio continuo:  $y_{d,g} = \partial I_{ref,d} / \partial V_g$ .

È banale a questo punto il calcolo della dispersione della tensione di soglia, parametro caratterizzante le fluttuazioni delle caratteristiche elettriche del dispositivo, che si calcola come la  $\sigma_{V_g}(V_g)$  valutata a soglia:  $\sigma_{V_g}(V_{th}) = \sigma_{V_g}(V_g = V_{th})$ . In questo lavoro di tesi la tensione di soglia è stata definita come la tensione di gate per la quale la corrente di drain raggiunge i  $100nA \frac{W}{L}$ .

Il noise-like IFM risulta essere immediato e veloce. Sdevice simula il dispositivo di riferimento, calcola la Green Function e poi, per ogni bias point, si calcola la densità spettrale di rumore al drain per una frequenza indicata (equazione 2.64).

Ovviamente questa frequenza deve essere tale che  $|\nu| < 0.5Hz$  se no restituisce valore nullo. Tramite un semplice integrale di volume abbiamo la densità spettrale e perciò, passando attraverso le equazioni 2.65 e 2.66, otteniamo la fluttuazione della tensione di soglia del dispositivo. Questo metodo ha il pregio appunto di essere veloce ed immediato a restituirci le fluttuazioni elettriche in termini di deviazioni standard, ma non ci dà nient'altro, ovvero non ci dice nulla sui singoli dispositivi che compongono la statistica. Il noise-like IFM non ci fa vedere come sono disposte le varie configurazioni atomistiche di drogaggio e quali transcaratteristiche formano. È un metodo estremamente utile per il calcolo della fluttuazione della  $V_{th}$  per l'RDF, ma risulta perciò inutilizzabile per lo studio dell'RTN.

### 2.3.4 “Statistical” IFM

L'idea dietro a questa variante IFM sta nel calcolo delle dispersioni statistiche direttamente dalle  $\delta I$  (equazione 2.7). Un modo per far ciò può essere quello di calcolare ad ogni bias point la dispersione della corrente di drain dalle  $\delta I$  e poi passare alla deviazione standard della tensione di gate (che valutata alla soglia del dispositivo con doping continuo ci dà la dispersione della  $V_{th}$ ) attraverso la transconduttanza  $y_{d,g}$ :

$$\langle \delta I_{\nu,d} \rangle = \frac{1}{N} \sum_{\nu} \delta I_{\nu,d} \quad (2.67)$$

$$\sigma_{I_d}(V_g) = \sqrt{\frac{1}{N} \sum_{\nu} (\delta I_{\nu,d} - \langle \delta I_{\nu,d} \rangle)^2} \quad (2.68)$$

$$\sigma_{V_g}(V_g) = \frac{\sigma_{I_d}(V_g)}{y_{d,g}} \quad (2.69)$$

Per arrivare a tale risultato si può, analogamente, ricostruire tutte le transcaratteristiche  $I_d - V_g$  di ogni configurazione atomistica di drogaggio randomizzata, per poi calcolare per ognuna di esse la tensione di soglia e ottenere così la  $\sigma_{V_g}(V_{th})$ . Per la costruzione della  $I_d - V_g$  per un particolare profilo di drogaggio, si parte dalla transcaratteristica di “riferimento”, ovvero quella per il dispositivo con

doping continuo, per poi, per ogni bias point, sommare la  $\delta I$  relativa a quella particolare  $\delta N$ . Il problema maggiore sta nel fatto che costruendo in questo modo le curve si può arrivare a risultati completamente sbagliati ed evidentemente non fisici, infatti in questo lavoro si è adottato un diverso metodo di ricostruzione delle transcaratteristiche per approssimare il meglio possibile la realtà.

Per la costruzione della  $I_d - V_g$  per una particolare configurazione atomistica di drogaggio  $\nu$ , partendo dalla transcaratteristica di riferimento  $I_{ref,d} - V_g$  e dalle fluttuazioni della corrente di drain per quel profilo di doping  $\delta I_{\nu,d}$ , abbiamo diversi metodi:

$$I_{\nu,d} = I_{ref,d}(V_g) - \delta I_{\nu,d} \quad (2.70)$$

$$I_{\nu,d} = I_{ref,d}(V_g + \delta V_{\nu,g}) \quad (2.71)$$

dove  $\delta V_{\nu,g} = -\delta I_{\nu,d}/y_{d,g}$ .

La 2.70 rappresenta il metodo di ricostruzione più ovvio, che è stato già descritto; mentre la 2.71 rappresenta il metodo duale: partendo dalla transcaratteristica di riferimento ad ogni bias di corrente si somma il discostamento in tensione  $\delta V_{\nu,g}$  per quella particolare configurazione di drogaggio rispetto al dispositivo di riferimento.

I due metodi di ricostruzione sono “duali” perché il primo ricostruisce la transcaratteristica da quella media calcolando lo shift di corrente a parità di  $V_g$ , mentre il secondo calcolando lo shift di tensione a parità di  $I_d$ . Essi risultano essere equivalenti per un sistema idealmente lineare: infatti espandendo la 2.71 secondo Taylor al prim'ordine e ricordandosi che  $y_{d,g} = \partial I_{ref,d}/\partial V_g$  si trova esattamente la 2.70. Ma un dispositivo elettronico è tutt'altro che un sistema lineare e perciò ci converrà usare un metodo piuttosto che l'altro a seconda di quanto è grande l'errore di linearizzazione per i due metodi di ricostruzione, a seconda dei dettagli di questa non linearità. Questo ragionamento ci porterà ad utilizzare un metodo ibrido chiamato “**Weighted Method**”.

### Ricostruzione delle transcaratteristiche

Vediamo inizialmente cosa succede adottando i due metodi di ricostruzione delle singole transcaratteristiche suggeriti dalle equazioni 2.70 e 2.71.

Se ad ogni bias di tensione sommiamo la deviazione di corrente spesso accade che questa  $\delta I$  risulta essere enormemente sovrastimata nella regione di soglia-sottosoglia. Questo errore di linearizzazione porta anche a casi patologici e non fisici, casi in cui  $\delta I > I_{ref,d}$  portando la transcaratteristica ad assumere valori negativi, come si vede in figura 2.5 in cui sono presenti 100 casi ricostruiti.

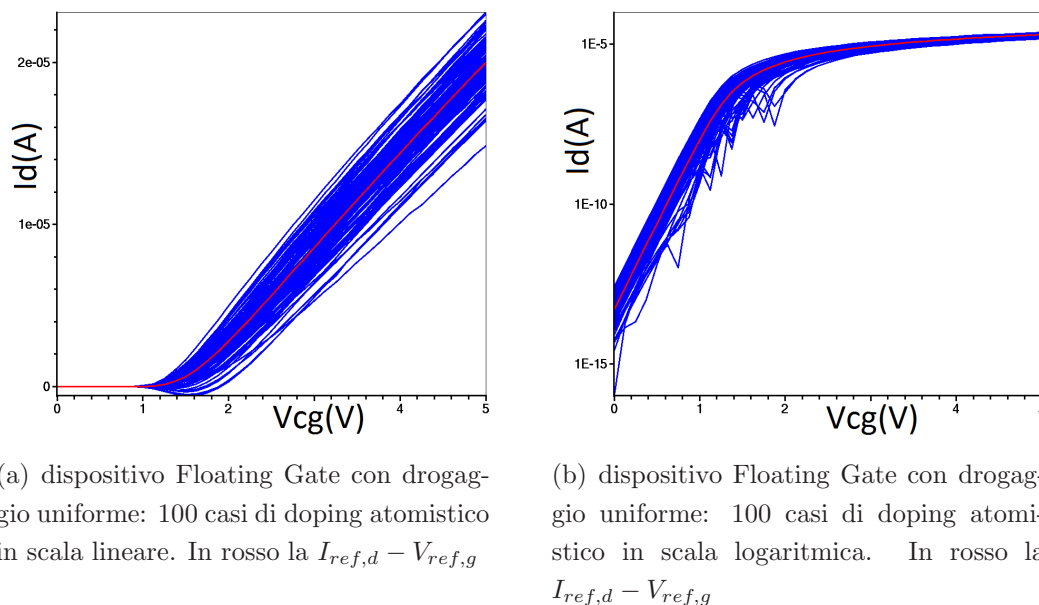


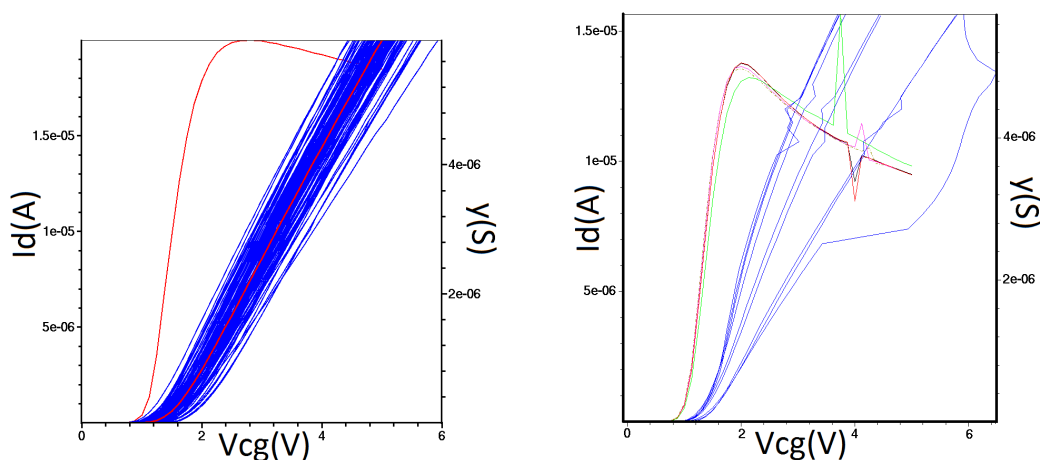
Figura 2.5: ricostruzione curve col metodo  $\delta I$

Notiamo come ci siano dei casi di doping in cui la transcaratteristica risulta essere decrescente con  $V_g$ , fino ad assumere valori negativi di corrente. Il cambio di segno di alcune  $I_{\nu,d}$  si vede bene in scala logaritmica, dove abbiamo dei “picchi” verso il basso di corrente (in scala logaritmica ho considerato il modulo della  $I_{\nu,d}$ ).

Questi casi patologici - non fisici nella regione di sottosoglia sono figli di un grosso errore di linearizzazione in quella regione: ciò che accade è che nella regione in questione la natura della corrente è fortemente percolativa, e perciò fortemente

non lineare. Per questo si commettono errori grossolani nella ricostruzione delle singole transcaratteristiche a basse polarizzazioni.

L'altro modo per la costruzione delle  $I_{\nu,d}(V_g)$  sta nel sommare la deviazione di tensione  $\delta V_{\nu,g}$  dal dispositivo di riferimento a parità di corrente di drain. Quello che succede è che nella regione di sottosoglia otteniamo valori ragionevoli e ben ricostruiti, mentre nel soprasoglia le cose sono più complicate: innanzitutto otteniamo transcaratteristiche con bias point  $V_{\nu,g}$  sopra l'alimentazione e poi, essendo che la  $\delta V_{\nu,g} = -\delta I_{\nu,d}/y_{d,g}$  e che la  $\delta I_{\nu,d}$  tende a saturare mentre la transconduttanza media  $y_{d,g}$  tende a decrescere, lo "shift" di tensione  $\delta V_{\nu,g}$  dal dispositivo di riferimento esplose portando il fascio di curve ad allargarsi in modo non fisico. Inoltre, se ci sono errori numerici nel calcolo della transconduttanza, come picchi ingiustificati, questi si ripercuotono tutti sulla  $\delta V_{\nu,g}$  che esplose od implode per un singolo punto portando la  $V_{\nu,g}$  ad essere non monotona. Tutti questi casi patologici descritti sono ben visibili nelle figure sotto riportate:



(a) dispositivo Floating Gate con drogaggio uniforme: 100 casi di doping atomico in scala lineare. In rosso la  $I_{ref,d} - V_{ref,g}$  e la  $y_{d,g}$

(b) diversi dispositivi Floating Gate: 9 casi patologici di  $V_{\nu,g}$  non monotona (in blu). Mostrate anche le  $y_{d,g}$  relative ad ognuno di essi

Figura 2.6: ricostruzione curve col metodo  $\delta V$

L'idea per non incorrere in tutti questi errori sopra descritti è di utilizzare un metodo di ricostruzione ibrido: nella regione di sottosoglia prevalgono i fenomeni elettrostatici, dunque sarebbe opportuno usare il metodo  $\delta V$ , mentre soprasoglia,

dove prevalgono i fenomeni di scattering e la corrente è più uniforme, la linearizzazione proposta dal metodo  $\delta I$  risulta un'ottima approssimazione. Come punto di congiunzione tra i due metodi viene preso il punto di massima transconduttanza. Questo metodo viene chiamato **“Weighted Method”** e per gli stessi 100 casi riportati nelle figure precedenti otteniamo il fascio di curve 2.7

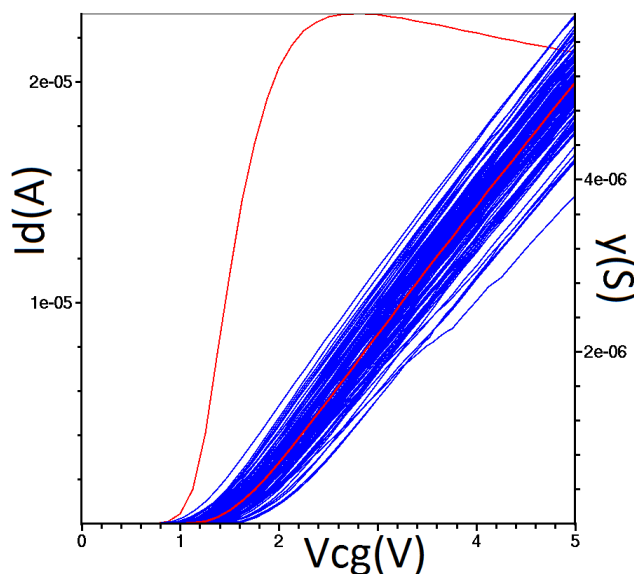
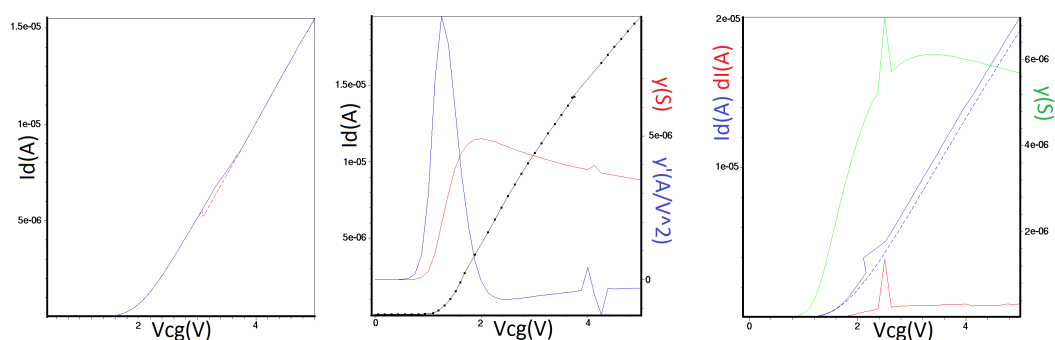


Figura 2.7: ricostruzione delle transcaratteristiche col Weighted Method: in rosso la transconduttanza e la transcaratteristica media

Per implementare questo metodo di ricostruzione si usa il tool della Synopsys “Inspect”. Si costruisce una procedura, chiamata “GetWeights”, da cui il metodo prende il nome, la quale ha come valore d’ingresso la lista contenente i valori della derivata della transconduttanza media ad ogni bias point, e ha come uscita una lista (W) di 1,-1 e 0: se  $\partial y_{d,g}/\partial V_g > 0$  restituisce -1, se  $\partial y_{d,g}/\partial V_g < 0$  restituisce +1, se c’è un cambio di segno nella  $\partial y_{d,g}/\partial V_g$  tra due valori consecutivi restituisce 0. Dopodiché vengono definiti  $W_p = 0.5(1 + W)$  e  $W_m = 0.5(1 - W)$  per tutti i valori della lista W, e otteniamo il risultato che vogliamo imponendo:

$$\begin{cases} I_{\nu,d} &= I_{ref,d}(V_g) + \delta I_{\nu,d} W_m \\ V_{\nu,g} &= V_{ref,g}(I_d) + \delta V_{\nu,g} W_p \end{cases} \quad (2.72)$$

Questo metodo di ricostruzione delle curve è quello che ho utilizzato nel mio lavoro di tesi per la maggior parte dei casi. Nonostante sia molto efficace e non possiede gli evidenti errori descritti in precedenza, non è immune da diversi problemi. Può succedere, infatti, che il raccordo tra i due metodi di ricostruzione sia brusco, e perciò diventa necessario un raccordo “smooth”, più morbido, spalmato fino ad 8 punti. Inoltre, può succedere che ci sia un errore numerico nel calcolo della transconduttanza media e ciò si ripercuote direttamente nella costruzione delle transcaratteristiche, portandole spesso ad avere delle  $V_g$  non monotone se il picco di  $y_{d,g}$  è sufficientemente alto. Se tale sovraelongazione avviene nella regione di soprasoglia non è difficile modificare la procedura “GetWeights” per filtrare tali errori, ma se accade nel sottosoglia è necessario eliminare il problema alla radice, optando per una simulazione elettrica “Extended Precision (128bit)” facendo, però, esplodere il tempo di simulazione. Nelle figure sotto riportate vediamo degli esempi in cui queste tre (seppur rare) problematiche di ricostruzione si manifestano.



(a) raccordo brusco tra i due sotto metodi  
 (b) picco della  $y_{d,g}$  nel so-  
 prasoglia dovuto ad un errore  
 numerico  
 (c) picco della  $y_{d,g}$  nel sot-  
 tosoglia dovuto ad un errore  
 numerico

Figura 2.8: problematiche di ricostruzione utilizzando il Weighted Method



### 2.3.5 sIFM per lo studio dell'RTN

Finora abbiamo visto come il metodo IFM, in particolare la variante “*statistical - Impedance Field Method*”, sia un’ottima metodologia, seppur approssimata, per lo studio delle dispersioni statistiche dovute alle fluttuazioni dei droganti, in vero la Synopsis ha implementato in sdevice questa procedura proprio per tale scopo. Oltre al fenomeno di RDF, inoltre, con l’IFM si possono studiare anche altre fluttuazioni statistiche (fluttuazioni geometriche, LER ecc), e quanto detto precedentemente sul metodo si ripete pari pari anche per gli altri fenomeni. Quello che interessa a noi, però, sono le fluttuazioni di RTN. In questo paragrafo spiego la metodologia di simulazione che ho sviluppato per lo studio dell’RTN e delle fluttuazioni di  $\Delta V_{th}$  attraverso due metodi duali tra di loro: sIFM-“doping” e sIFM-“trap”.

#### sIFM-“doping”

A differenza della variante “noise-like”, con lo “statistical” IFM (abbreviato in “sIFM”) abbiamo la possibilità di costruire, tramite il sopradescritto “weighted method”, le singole transcaratteristiche di ogni configurazione atomistica di drogaggio. Questa importante proprietà del metodo permette di studiare l’RTN. Qui di seguito è spiegato come ho utilizzato lo sIFM per la costruzione di una vasta statistica di  $\Delta V_{th}$ .

Inizialmente per ciascun bias point calcoliamo la Green Function  $G_D^0(r)$  dalla soluzione delle equazioni Poisson/drift-diffusion ottenuta nel caso di doping continuo e trappola neutra. Dopo avere generato  $N_{dop}$  profili di drogaggio randomizzati, ci calcoliamo le  $\delta I_{\nu,d}(V_g)$  (2.7) per ogni bias point  $V_g$  e per ogni configurazione di drogaggio  $\nu$ , costruiamo così le transcaratteristiche relative a ciascun profilo di doping  $\nu$  (2.72) ed otteniamo infine le rispettive tensioni di soglia neutre  $V_{th,\nu}^0$ . Dopodiché, ripetiamo questo procedimento nel caso di trappola “piena”, ovvero nel caso di trappola all’interfaccia Si/Ox caricata negativamente, per le stesse configurazioni di drogaggio: calcolo  $G_D^-(r)$ , calcolo  $\delta I_{\nu,d}$ , costruzione delle transcaratteristiche e calcolo delle  $V_{th,\nu}^-$ . In questo modo abbiamo  $N_{dop}$  valori di

$\Delta V_{th} = V_{th,\nu}^- - V_{th,\nu}^0$  per quella trappola. In altre parole: costruiamo le  $N_{dop}$  transcaratteristiche (ognuna riferita ad un profilo atomistico di drogaggio particolare) con il “weigthed method” per il caso di trappola neutra, dopodiché costruiamo le transcaratteristiche per le stesse  $N_{dop}$  configurazioni di doping ma stavolta con la trappola carica ed infine valutiamo per ognuna di esse lo shift di soglia  $\Delta V_{th}$  causato dalla presenza di quella carica negativa alla superficie del canale. Questa procedura con la trappola occupata la ripetiamo  $N_{trap}$  volte, ognuna relativa ad una diversa posizione randomizzata della trappola alla superficie del canale, sempre mantenendo le stesse  $N_{dop}$  configurazioni atomistiche di drogaggio. In tal modo otteniamo una statistica di  $N_{dop} \times N_{trap}$  valori dello shift di soglia  $\Delta V_{th}$  dovuto alla presenza di una singola trappola carica negativamente all’interfaccia. Quanto detto viene graficamente descritto dalla figura 2.9.

In questo lavoro ho considerato  $N_{trap} = N_{dop} = 100$  (ovvero 100 configurazioni atomistiche di drogaggio e 100 posizioni diverse della trappola in superficie) ottenendo una statistica di  $10^4$  valori nello stesso tempo computazionale con il quale il metodo MC descritto precedentemente ne ottiene solo 200 circa. Da questo dato capiamo perfettamente come lo sIFM permette di ottenere una statistica nettamente più ampia del metodo MC, rendendo possibile l’esplorazione delle bassissime probabilità nelle cumulate dello shift di soglia per l’RTN. D’altro canto, è necessario tener conto che il metodo descritto si basa su approssimazioni/linearizzazioni e pertanto è fondamentale capire quando e in quali casi non è più credibile e ci dà risultati errati.

Ho dato per scontato, inoltre, come costruire una statistica con le stesse 100 configurazioni di drogaggio e le stesse 100 trappole incrociando tutti i casi possibili sia corretto e coincidente con il procedimento descritto precedentemente per il metodo MC. Ciò dovrebbe essere vero, in quanto il profilo atomistico di doping e la posizione della trappola sono due variabili aleatorie indipendenti, e, dunque, le  $10^4$  coppie risultanti dovrebbero comporre una giusta statistica, poiché indipendenti anch’esse. Comunque sia, verificheremo quanto detto nel capitolo dei risultati.

Il metodo appena descritto per la simulazione dell’RTN che fa uso dello statistical-IFM viene chiamato (dal sottoscritto) “sIFM-doping” (o più brevemente “sIFM-dop”), in quanto modella il drogaggio atomistico come piccola

perturbazione dal device di riferimento con trappola neutra o carica.

Lo “sIFM-doping” risulta estremamente veloce nel costruire la statistica desiderata. In vero, per ottenere  $10^4 = N_{dop} \times N_{trap} = N$  valori dello shift di soglia ho bisogno di fare  $1 + N_{trap} = 101$  simulazioni (1 per il dispositivo con trappola scarica, le altre con trappola carica), mentre per il metodo MC ho bisogno di  $(1 + N_{trap}) \times N_{dop} = 10100$  simulazioni (per gli stessi dispositivi). Tenendo conto che la singola simulazione con il metodo IFM dura circa solo il doppio di una Drift Diffusion (per i dispositivi e le mesh utilizzate in questo lavoro di tesi), in quanto necessita il calcolo anche della funzione di Green, il vantaggio in termini di tempo è notevole (fattore  $\approx 50$ ).

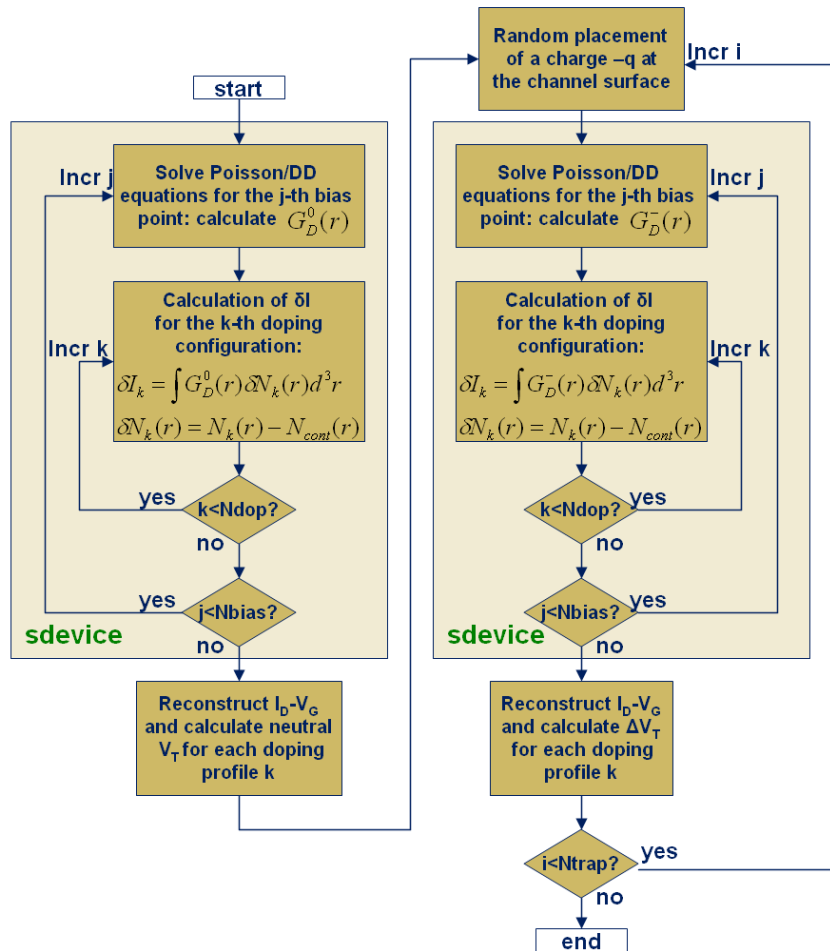


Figura 2.9: diagramma di flusso per il calcolo di  $N_{dop} \times N_{trap}$  valori dello shift di soglia  $\Delta V_{th}$  dovuto al fenomeno dell’RTN

**sIFM-“trap”**

Il metodo descritto nel paragrafo precedente (sIFM-dop) tratta il drogaggio atomistico come una variazione da quello ideale-continuo, la quale a sua volta provoca una  $\delta I$  di corrente in risposta lineare al contatto di drain. Lo sIFM-dop applica il metodo al doping fissata la trappola. Per lo studio dell’RTN abbiamo però due variabili aleatorie indipendenti: la trappola e il doping atomistico. Il metodo “duale” (sIFM-trap) fa esattamente l’opposto: vede la trappola come una piccola perturbazione dal dispositivo di riferimento, che è il device avente un particolare drogaggio atomistico e trappola neutra, e tale variazione provoca linearmente una deviazione di corrente al drain. In altre parole: lo sIFM-trap applica il metodo alla trappola fissata la configurazione di drogaggio atomistico. Tale procedimento è anche più corretto, in quanto ricordiamo che il metodo applicato è un metodo lineare: più la perturbazione dal dispositivo di riferimento è piccola migliore sarà l’approssimazione. Trattare la trappola come “piccola perturbazione” piuttosto che il drogaggio atomistico è sicuramente un’approssimazione migliore.

Purtroppo Sdevice non implementa lo sIFM applicato sulle trappole interfaciali, non c’è il modo di dirgli di calcolarsi le Green Function all’interfaccia che danno le risposte impulsive ai contatti per la presenza di una carica in tale regione. Però l’aver una trappola caricata negativamente all’interfaccia equivale ad avere un atomo accettore ionizzato in quel punto, si tratta in entrambi casi di una carica  $-q$ . Per questo l’idea è quella di sfruttare l’IFM per l’RDF (visto in precedenza) applicandolo sul dispositivo “atomistico” ed andando a considerare la Green Function solo all’interfaccia. Selezioniamo una configurazione atomistica di drogaggio (generata dal metodo) e per essa ci calcoliamo dapprima la transcaratteristica con trappola neutra e poi le  $I_d - V_g$  di tutte le trappole per quella configurazione di drogaggio. Stavolta, però, la transcaratteristica per una particolare trappola  $i$  nella posizione  $r_i$  la si costruisce calcolandosi per ogni  $V_g$  lo shift di corrente  $\delta I_{i,d}$  dal dispositivo con trappola neutra come la Green Function valutata in  $r_i$  (in  $s^{-1}$ ) moltiplicata per la carica puntuale  $q$  (in  $C$ ):

$$\delta I_{i,d} = \int q\delta(r - r_i)G_T(r)dr = qG_T(r_i) \quad (2.73)$$

Calcolati i  $\delta I_{i,d}$  per ogni bias point e per ogni trappola costruisco tutte le transcaratteristiche con il solito metodo 2.72. Dopodiché questo procedimento viene ripetuto per tutti i profili atomistici di drogaggio. Quanto detto viene graficamente descritto dal diagramma di flusso sottostante:

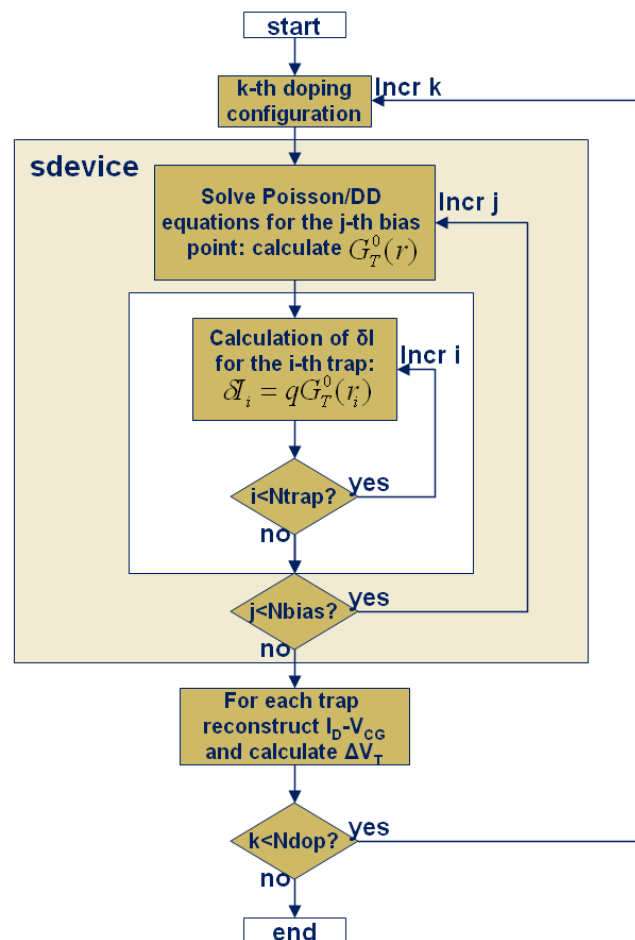


Figura 2.10: diagramma di flusso per il calcolo di  $N_{dop} \times N_{trap}$  valori dello shift di soglia  $\Delta V_{th}$  dovuto al fenomeno dell'RTN attraverso il metodo sIFM-trap

Detto ciò, il tempo computazionale necessario per il calcolo di  $10^4 = N_{dop} \times N_{trap} = N$  valori dello shift di soglia è di solo  $N_{dop}$  simulazioni (una per ogni profilo di doping).

Un altro vantaggio dello sIFM-trap sta nel fatto che il dispositivo con trappola neutra ha la transcaratteristica giusta, ovvero simulata proprio con il modello

Drift Diffusion. Perciò l'errore di approssimazione risiede solo nella transcaratteristica del dispositivo con trappola carica e non in entrambi come prima accadeva. Come si fa ad implementare questo metodo utilizzando i tool della Synopsys? Innanzitutto con il tool *Tdx* o *Data Explorer* riusciamo a separare in  $N_{dop}$  file .msh le configurazioni di drogaggio dall'unico file generato da Sdevice che le conteneva. Ognuno di essi sarà l'input file di Sdevice in cui verrà attivato il metodo sIFM con il calcolo ed il plot delle Green Function. Purtroppo Tdx non possiede un modo per associare i valori dei dataset alle coordinate dei nodi della mesh. Per questo ho utilizzato un programma costruito in azienda chiamato *plotitf*: esso mi scrive in un file di testo tutti i valori del dataset specificato con i nodi della mesh associati per le interfacce dichiarate. Ho generato 4100 comandi (41 bias point x 100 configurazioni di drogaggio) per ottenere tutti i dataset delle Green Function all'interfaccia *Si/Ox*. A noi, però, non interessano tutti i valori che la  $G_T(r)$  assume all'interfaccia, ma solo 100 di essi (perchè abbiamo  $N_{trap} = 100$  posizioni della trappola). Per ognuno di questi 4100 file, con *Inspect* dapprima carichiamo i valori dei nodi della mesh all'interfaccia  $r_j = (x_j, y_j, z_j)$  e le  $G_T(x_j, y_j, z_j)$  relative, poi associamo le posizioni delle trappole  $r_i$  ai nodi più vicini  $r_{j,i}$  (dove è minima la distanza euclidea) e per esse approssimiamo la 2.73 come  $\delta I_{i,d} \approx qG_T(r_{j,i})$ . Ovviamente l'approssimazione è migliore tanto più la mesh è fine nella regione in questione. Tal procedimento lo ripetiamo dapprima per ogni bias point e poi per ogni configurazione di drogaggio con un doppio ciclo for all'interno del command file di Inspect.

Ciò detto viene graficamente descritto dalla sottostante immagine:

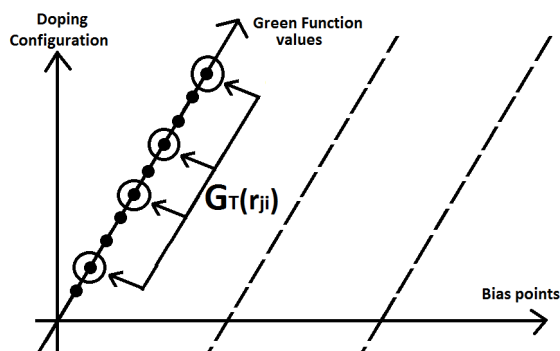


Figura 2.11: metodo sIFM-trap: gestione dei valori all'interno di Inspect

### 2.3.6 IFM in Sdevice

Per attivare l'Impedance Field Method in Sdevice basta inserire le giuste keyword nel command file. Nella sezione *Physics* bisogna specificare la keyword *Noise* con l'opzione *Doping* per attivare il calcolo della densità spettrale di rumore per l'RDF.

```
Physics {...
    Noise(Doping(Mobility BandGapNarrowing))
}
```

Si possono specificare anche le opzioni *Mobility* e *BandGapNarrowing* affinché queste due proprietà fisiche dipendano dalla randomizzazione del drogante nel dispositivo (nel mio lavoro di tesi non l'ho fatto).

Per attivare, invece, il calcolo delle  $\delta I$  (equazione 2.7), è necessario inserire nel cmd file:

```
Math {...
    RandomizedVariation(
        NumberOfSamples = 100
        Randomize = 147
    )
}
```

Inoltre se si vuole assicurare una robusta e veloce simulazione:

```
Math {...
    Method = Blocked
    SubMethod = ILS
    ACMethod = Blocked
    SubACMethod = ILS (Set = 2)
}
```

Il noise-like IFM richiede un' analisi in AC. Per questo motivo è necessario introdurre la sezione *System*, nella quale la sotto-sezione *Solve* contiene il blocco *Quasistationary* seguito da quello *ACCoupled*:

```

Solve {...
  Quasistationary(...
    Goal {Parameter = vgate.dc Voltage = 1.0}
  ) {ACCoupled (
    StartFrequency = 0.1 EndFrequency = 0.1
    NumberOfPoint = 1Decade
    Node(ndrainngatensource)
    ObservationNode(ndrainngate)
    Exclude(vdrainvsourcvegat)
    ACExtract = n@node@_mos
    NoisePlot = n@node@_noise_lin
  ) {Poisson Electron Hole}
  }
}

```

In questo modo vengono generati i file:

- IdVgLin\_n@node@\_nmos\_I\_ndrain.csv
- nmos\_n@node@\_des.tdr
- IdVgLin\_n@node@\_nmos\_ac\_des.plt
- n@node@\_noise\_lin\_nmos\_ndrain\_@bias\_point@\_acgf\_des.tdr

Il primo, attivato dalla keyword *RandomizedVariation* nella *Math* section, contiene in formato CSV (comma separated values) tutte le  $\delta I$ , ai vari bias point e per ogni configurazione di drogaggio; il secondo, attivato dalla *Plot* section, contiene i dataset di tutti i drogaggi atomistici (visualizzabili con *tecplot\_sv*) purché specificata la keyword *RandomizedDoping*; il terzo, attivato dalla keyword *ACExtract* nella *ACCoupled* section, contiene le densità spettrali di RDF ai vari bias point; il quarto, attivato dalla keyword *Noiseplot* nella *ACCoupled* section, contiene, per quel bias point, i valori per ogni nodo della mesh delle GreenFunction e delle



integrande delle densità di rumore (visualizzabili con `tecplot_sv`) specificate nella sezione *NoisePlot*.

Questi output file di Sdevice, insieme a quello contenente i valori della  $I_d - V_g$  del dispositivo di riferimento (`IdVgLin_nmos_n@node@_des.plt`), permettono di realizzare le varianti s-IFM e nl-IFM. Per mezzo di quest'ultimo file e del file `.csv` contenente i valori delle  $\delta I$  ci possiamo ricostruire le transcaratteristiche dei vari dispositivi randomizzati (equazioni 2.72). Attraverso il terzo file, invece, che contiene i valori delle  $S_{I_d}$  e delle  $y_{d,g}$ , riusciamo a costruire la  $\sigma_{V_g}(V_g)$  (mediante le equazioni 2.65 e 2.66). Tutte queste operazioni di caricamento dei valori dai file e costruzione di curve mediante le equazioni citate possono essere fatte dal tool della Synopsys *Inspect* mediante opportuni script.

# Capitolo 3

## Risultati

Dopo aver brevemente mostrato alcuni confronti tra il metodo IFM e il MC per quanto riguarda il fenomeno dell'RDF, in questo capitolo mostrerò i risultati dei metodi *sIFM-doping* e *sIFM-trap* per le fluttuazioni di RTN per diversi profili di doping. Li confronteremo con il riferimento MC, cercando di capire vantaggi e limitazioni al fine di prevedere il loro comportamento in altri dispositivi.

### 3.1 RDF

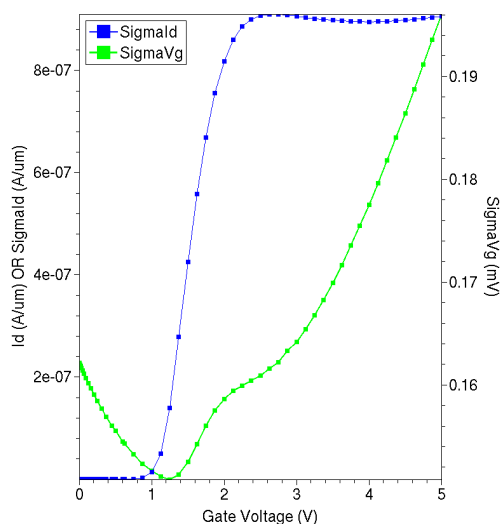
In questa sezione voglio brevemente mostrare come il metodo IFM funzioni discretamente bene per il calcolo delle dispersioni statistiche per quanto riguarda il fenomeno dell'RDF ([1], [2], [3] e [4]).

#### 3.1.1 Calcolo delle dispersioni

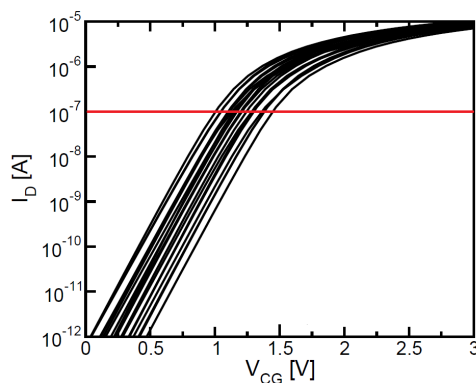
Per il calcolo della  $\sigma_{V_T}$  sono possibili due strade: il *noise-like IFM* (nIFM) e lo *statistical IFM* (sIFM). Il primo metodo calcola la deviazione standard della corrente  $\sigma I_D(i)$  dall'equazione 2.65 ad ogni bias point  $i$ , che, trasformata mediante la transconduttanza media, ci dà la fluttuazione della tensione al gate  $\sigma V_G(i)$  (2.66). Nella figura 3.1(a) riporto il caso di drogaggio uniforme: per ogni bias point  $i$  vengono calcolate le deviazioni standard di corrente di drain e di tensione

di gate a partire dalla densità spettrale di rumore calcolata da sdevice. Dopo aver ottenuto la tensione di soglia come la  $V_G$  per la quale il dispositivo di riferimento (a drogaggio continuo) raggiunge la corrente al drain di  $I_{D_{th}} = 100nA \frac{W}{L} = 100nA$ , valuto la  $\sigma_{V_G}(V_G)$  ( $= \sigma_{V_G}(i)$ ) a soglia ottenendo la  $\sigma_{V_T}$ .

Per quanto riguarda invece lo *statistical IFM* ricostruisco ogni transcaratteristica dai dispositivi randomizzati utilizzando il *weighed method* (2.72) per poi valutare per ognuno di essi la tensione di soglia (Fig.3.1(b)) e calcolare infine la  $\sigma_{V_T}$ . I risultati per le due varianti del metodo IFM coincidono, purché la statistica considerata dallo sIFM sia sufficientemente grande. È importante che i bias point per cui vengono calcolati gli shift di corrente  $\delta I_d$  coincidano con i bias point della transcaratteristica media.



(a)  $\sigma_{V_G}(i)$  e  $\sigma_{I_D}(i)$  nel caso di doping uniforme



(b) fascio di curve calcolate per diverse configurazioni atomistiche di doping con il metodo sIFM per il dispositivo a drogaggio uniforme. Viene evidenziata la  $I_{D_{th}}$  per la quale vengono calcolate le soglie dei campioni

Figura 3.1: varianti nIFM (a) e sIFM (b) per il metodo IFM

### 3.1.2 Risultati

Riporto qui i risultati relativi al confronto tra il MC e l'IFM per diversi casi: drogaggio uniforme e retrogrado 3.2, drogaggio delta shape 3.3(a), drogaggio disuniforme lungo L e W 3.3(b) e geometrie aventi bordi arrotondati 3.4. Nella prima figura, inoltre, ho mostrato anche il confronto con entrambe le varianti IFM (nIFM e sIFM).

Si osserva un buon accordo del metodo IFM con il riferimento MC in tutti i casi riportati. Guardando la Fig.3.2 si nota come anche la tendenza della curva in funzione dello svuotamento del canale sia perfettamente in accordo con quella MC; analogo discorso per la tendenza  $\sigma_{V_T}$  in funzione del raggio di curvatura sullo spigolo di Fig.3.4. Tali confronti sono stati fatti tutti imponendo la  $I_{D_{th}} = 100nA$ . Analoghi risultati sono stati ottenuti anche considerando livelli di correnti minori. Per correnti maggiori, invece, specialmente quando il canale è ben formato, il metodo presenta qualche limitazione, come vedremo dopo al par.3.2.1.

A parte per il caso della Fig.3.2, gli altri risultati sono stati ottenuti con mesh molto diverse tra i metodi. Per questi casi le  $\sigma_{V_T}$  per il metodo MC sono state prese dagli articoli [19] e [23]. Mesh diverse vuol dire aggiungere dispersione nel calcolo della deviazione standard della tensione di soglia. Abbiamo visto come spesso le differenze tra i risultati IFM e quelli MC siano minime, e quindi tali dispersioni possono essere dominanti. Pertanto per i casi di drogaggio retrogrado e uniforme ho preferito fare un confronto più significativo, ovvero con la stessa mesh, in modo tale che l'unica cosa che varia sia effettivamente solo il metodo. Inoltre, poiché l'sIFM permette di selezionare un gruppo limitato e preciso di configurazioni di drogaggio, ho selezionato la stessa statistica di campioni, ovvero gli stessi profili di doping, per il MC e l'sIFM.

Riporto alle Fig.3.5(a) e 3.5(b) il caso di drogaggio uniforme (Fig.3.2 a  $t_{epi} = 0$ ) per quanto riguarda i metodi sIFM e MC: stessa mesh e stessi 392 profili di doping atomistico considerati. L'accordo è molto forte con il MC, molto più di quanto ci aspettavamo dalle figure precedenti: i singoli valori sono abbastanza correlati (coefficiente di correlazione  $\rho = 0.913$ ) e su una linea di tendenza vicino alla bisettrice del primo e del terzo quadrante. Le cumulative corrispondenti sono praticamente coincidenti e le deviazioni standard sostanzialmente identiche:  $\sigma_{V_T}^{MC} = 159.4mV$  e  $\sigma_{V_T}^{sIFM} = 164.3mV$ .

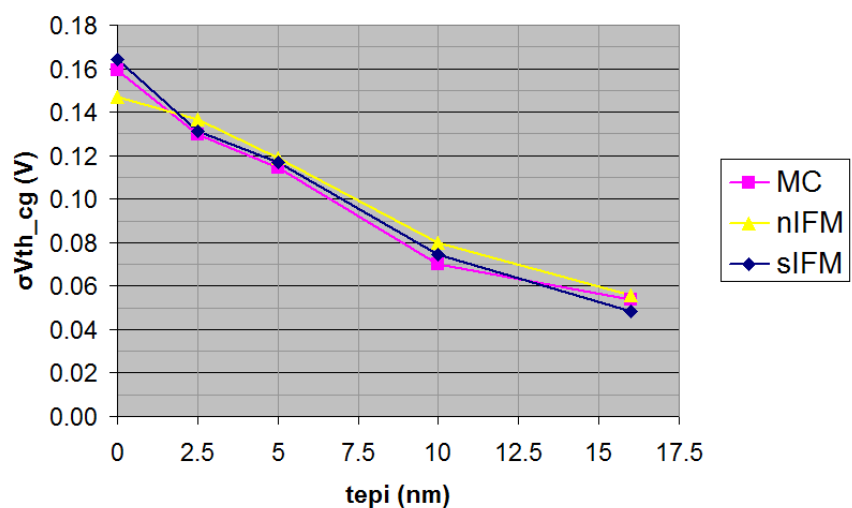


Figura 3.2: dispersioni di soglia nel caso uniforme e retrogrado alle varie  $t_{epi}$  per i metodi MC, nIFM e sIFM

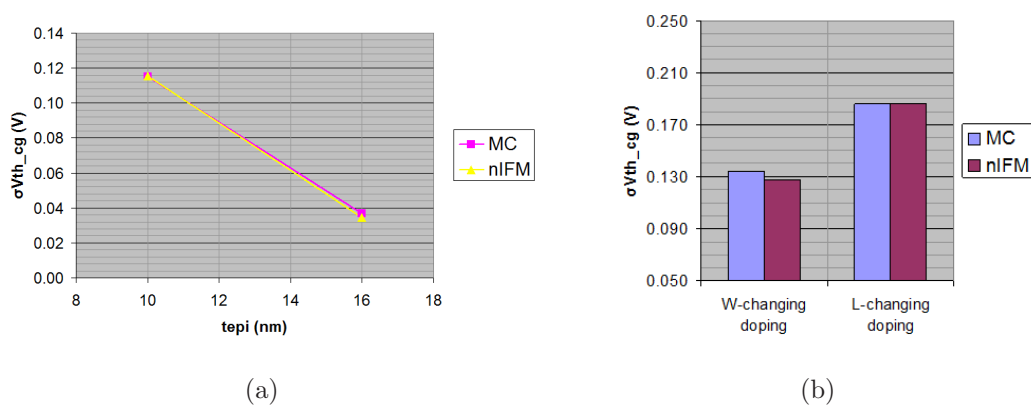


Figura 3.3: dispersioni di soglia nel caso delta doping alle varie  $t_{epi}$  (a) e disuniforme lungo W ed L (b) per i metodi MC e nIFM

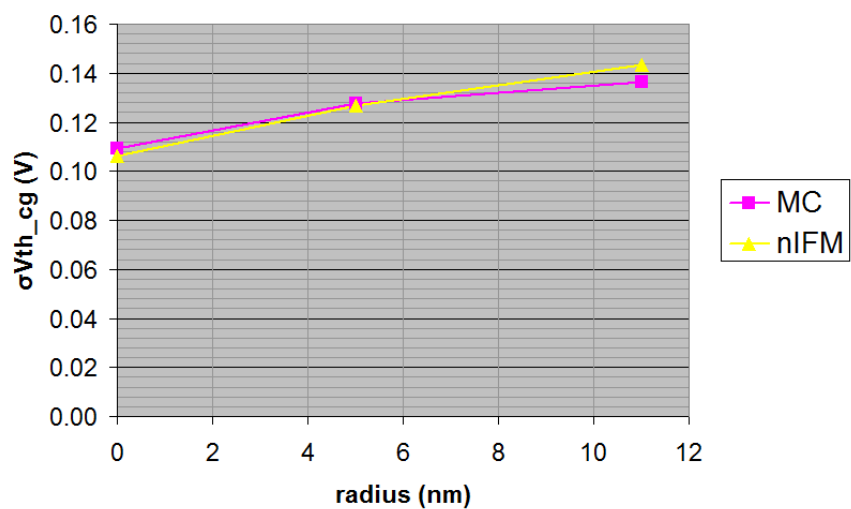
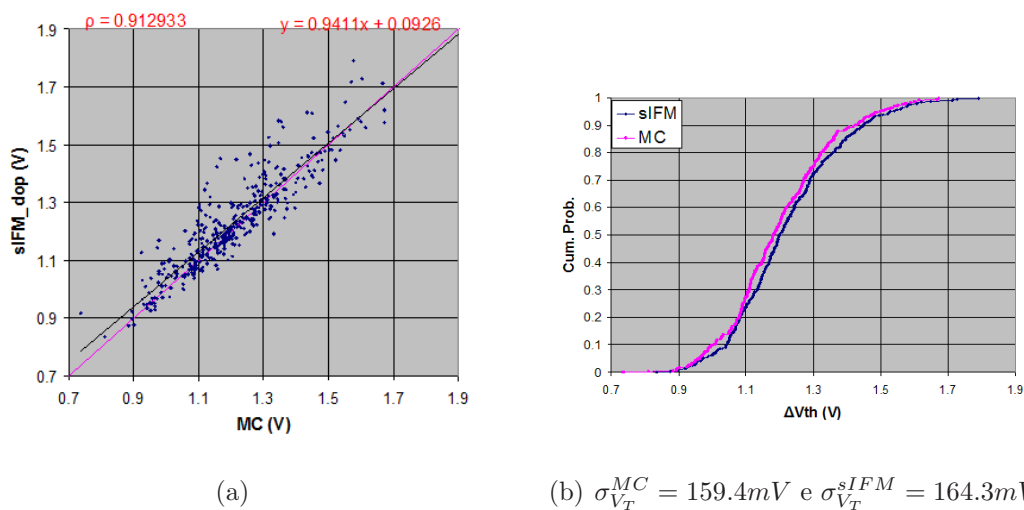


Figura 3.4: dispersioni di soglia nel caso di celle con bordo arrotondato al variare del raggio di curvatura per i metodi MC e nIFM



(a)

(b)  $\sigma_{V_T}^{MC} = 159.4mV$  e  $\sigma_{V_T}^{sIFM} = 164.3mV$ 

Figura 3.5: scatterplots sIFM-MC (a) e cumulative relative ai due metodi (b) per gli stessi 392 dispositivi randomizzati. Caso di doping uniforme

## 3.2 RTN

In questo paragrafo mostriamo i risultati ottenuti sulle fluttuazioni dell'RTN con i metodi “*sIFM-doping*” e “*sIFM-trap*” analizzandoli e confrontandoli con i rispettivi dati MC ottenuti da simulazioni *Drift Diffusion*. Mostrerò questi confronti per diversi dispositivi e cercherò di interpretare tali risultati. Data la natura approssimativa dei due metodi, infatti, un obiettivo da porsi è sicuramente quello di capire i limiti di validità di essi andando ad analizzare ed interpretare come, quanto e quando i metodi funzionano. Molti di questi risultati e molte di queste osservazioni sono stati oggetto di due pubblicazioni ([25] e [26]).

Ricordo innanzitutto che i dispositivi simulati sono celle floating-gate, la cui struttura e profili di doping sono stati descritti nel capitolo precedente. La statistica per i metodi sIFM è stata costruita su  $N_{doping} = 100$  profili atomistici di drogaggio e  $N_{trap} = 100$  posizioni della trappola, ovvero su  $N_{doping} \times N_{trap} = 100 \times 100 = 10^4$  valori di shift di soglia  $\Delta V_T$ , potendo andare ad esplorare le distribuzioni cumulative fino ad una probabilità di  $10^{-4}$ . Data la pesantezza computazionale del MC in termini di tempo, invece, si è simulato un numero molto inferiore di dispositivi per questo metodo, selezionando un piccolo “subset” dei 10000: per esempio simulo DD i 100 devices con trappola neutra, poi per ognuno di essi inserisco randomicamente (una alla volta) 4 delle 100 trappole creando così 400 coppie configurazione di drogaggio-trappola, dopodiché simulo tali 400 dispositivi DD andando poi a calcolare i 400 valori di shift di soglia risultanti.

### 3.2.1 Drogaggio uniforme

Iniziamo a vedere come si comportano i metodi per quanto riguarda il dispositivo con drogaggio uniforme, ovvero per il device il cui doping accettore ideale è continuo e costante ovunque pari a  $N_A = 2 \times 10^{18} \text{cm}^{-3}$ . Qui di seguito riporto per l'appunto un'immagine del dispositivo in questione (Fig.3.6) e le cumulate risultanti dai 3 metodi (Fig.3.7).

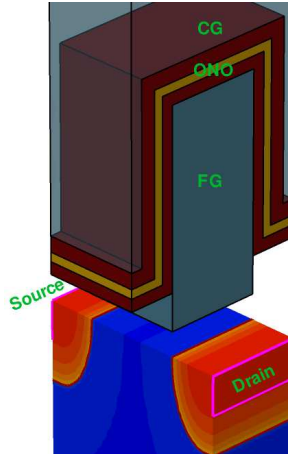


Figura 3.6: dispositivo floating-gate con profilo di drogaggio ideale uniforme pari a  $N_A = 2 \times 10^{18} \text{cm}^{-3}$

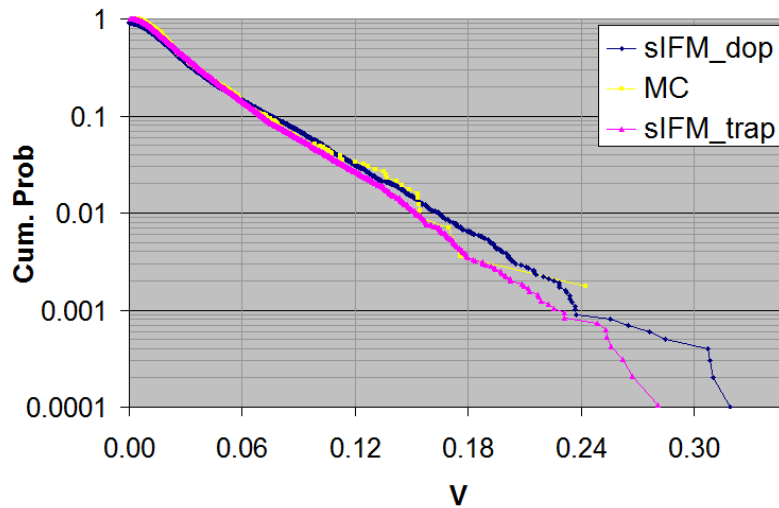


Figura 3.7: distribuzioni di probabilità cumulativa della  $\Delta V_T$  per il caso di drogaggio uniforme costruite su  $N_{trap} \times N_{doping} = 100 \times 100 = 10^4$  valori: in blu il metodo sIFM-doping, in lilla lo sIFM-trap e in giallo il MC (costruito su 568 valori subset dei 10000)



Il MC della Fig.3.7 è stato costruito su 568 valori dello shift di soglia e non su 600 in quanto alcune simulazioni sono fallite per mancata convergenza della risoluzione numerica. È banale dire, infatti, come simulare il comportamento fisico del dispositivo risulta alquanto più difficoltoso dal punto di vista computazionale-numerico per tali dispositivi “atomici”, dove i profili di drogaggio hanno disuniformità molto piccate per la presenza degli atomi ionizzati. Per limitare questi problemi ed evitare il “doping nullo” (che porterebbe ad inevitabili errori e divergenze del codice) si sono simulati tali dispositivi con uno strato aggiuntivo di drogaggio uniforme a bassa concentrazione ( $10^{10} \text{ cm}^{-3}$ ).

Dai risultati della Fig.3.7 per il dispositivo con doping uniforme (Fig.3.6) possiamo ben vedere come i tre metodi praticamente coincidono. Questo è soprattutto vero ad alte probabilità, mentre per basse troviamo un discostamento massimo tra i due metodi sIFM del 2% circa. La teoria (vedi capitoli precedenti) ci dice come la distribuzione di probabilità cumulativa dovrebbe essere un esponenziale decrescente, dunque in un grafico semilogaritmico una retta a pendenza negativa, e così accade per tutti e tre i metodi.

Tutto ciò è parecchio sorprendente, poiché il caso di doping uniforme è un po’ da considerarsi un “worst case”, in quanto i fortissimi effetti non lineari di current crowding e di percolazione della corrente dovrebbero portar i metodi lineari *sIFM-doping* e *sIFM-trap* ad approssimare molto male la realtà. Come infatti ho spiegato nei capitoli precedenti, proprio per limitare questi effetti si adottano geometrie più arrotondate o drogaggi disuniformi.

D’altro canto, però, il fatto che le cumulate risultano molto simili non vuol dire che i valori delle  $\Delta V_T$  per ogni singolo dispositivo siano ben approssimati, tanto meno che i metodi ricostruiscono bene le transcaratteristiche!! Approssimare bene le distribuzioni di probabilità cumulativa del MC, difatti, significa che “in media” i metodi sIFM risultano dare uno shift di soglia in linea con il DD, ma non è detto che ci sia una forte correlazione tra i metodi sIFM e il metodo MC. In altre parole, può essere solo un caso che tali distribuzioni siano sovrapposte, cioè che non si verifica per altri dispositivi. Per vedere l’entità di questa correlazione e su che tendenza avviene, dobbiamo andare a vedere come i metodi sIFM si comportano in ognuno dei 568 casi, costruendo così gli *scatter plot sIFM\_trap-MC* e *sIFM\_doping-MC*. I confronti diretti per questi casi sono riportati nelle figure 3.8(a), 3.8(b) e 3.8(c).

Dalla figura 3.8(c) notiamo come il subset di dispositivi su cui costruiamo la distribuzione cumulativa di 568 valori sia perfettamente in linea con la tendenza della curva costruita su tutti i 10000 casi, il che verifica solo il fatto che il subset di devices in questione sia preso veramente in modo randomico. Da questi andamenti, possiamo aspettarci che gli scatterplots risultanti tra i metodi *sIFM* e il MC siano più o meno centrati sulla bisettrice del primo e del terzo quadrante, ma non sappiamo ancora quanto sia larga la “nuvola” di correlazione.

Le figure 3.8(a) e 3.8(b) mostrano rispettivamente in grafici bilogaritmici gli scatterplots *sIFM\_trap-MC* e *sIFM\_doping-MC* con i coefficienti di correlazione risultanti e le linee di tendenza relative. Ricordo ora brevemente la definizione del coefficiente di correlazione.

Date  $n$  osservazioni congiunte di due variabili  $x$  e  $y$   $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , si definisce il coefficiente di correlazione come

$$\rho = \frac{S_{xy}}{\sqrt{\sigma_x^2 \sigma_y^2}} = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y} \quad (3.1)$$

dove  $\bar{x}$  e  $\bar{y}$  sono i valori medi,  $S_{xy}$  è la covarianza e  $\sigma_x$  e  $\sigma_y$  sono le deviazioni standard per le variabili  $x$  ed  $y$ .

Quello che possiamo dire è che, come ci aspettavamo, le “nuvole” di correlazione posano praticamente sulla bisettrice e risultano piuttosto strette (coefficienti di correlazione  $\rho > 0.85$ ). Notiamo poi che il metodo *sIFM\_trap* risulta essere più correlato con il MC rispetto allo *sIFM\_doping* ( $\rho = 0.899$  contro  $\rho = 0.858$ ), anche se su una linea di tendenza più lontana, seppur di poco, dalla bisettrice  $y = x$ . Osserviamo, inoltre, dalla figura 3.8(b) come purtroppo per bassi shift di soglia la tendenza dell'*IFM\_doping* sia quella di sottostimare sempre di più tale quantità rispetto al DD. Quello che non si vede dal grafico bilogaritmico è come spesso tale metodo addirittura calcoli shift di soglia negativi, il che è assolutamente non fisico dato che l'effetto dell'intrappolamento nel canale di una carica negativa mobile sia quello dell'abbassamento della corrente totale dovuta alla repulsione coulombiana per le cariche dello stesso segno. Ciò è visibile dalla figura 3.9. Vedremo bene come i dispositivi con shift di soglia molto piccoli calcolati DD portano spesso ad una inevitabile  $\Delta V_T < 0$  calcolata con il metodo *sIFM\_doping*.

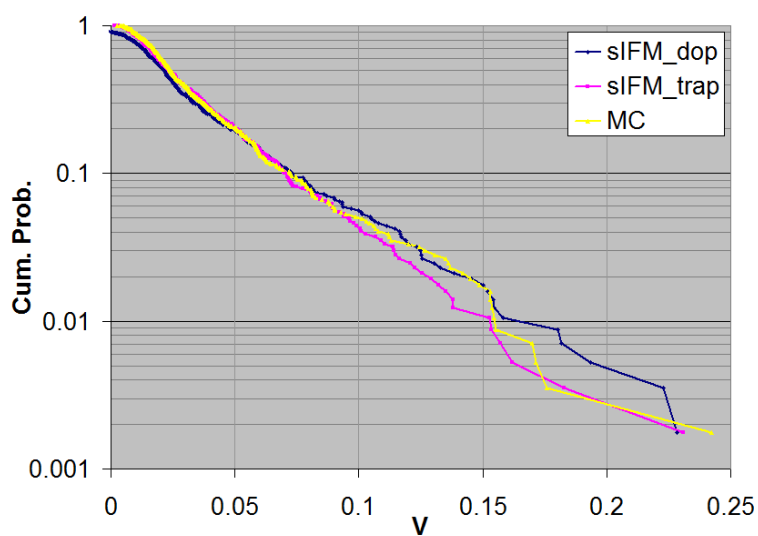
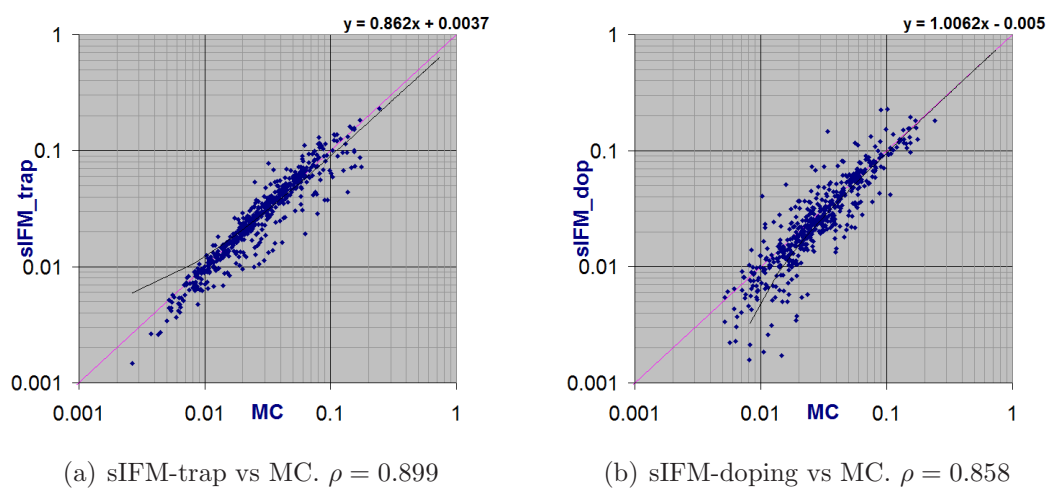


Figura 3.8: scatterplots su 568 valori relativi ai metodi  $sIFM\_trap$  (a) e  $sIFM\_doping$  (b) con il MC nel caso di drogaggio uniforme, e le cumulate corrispondenti (c)

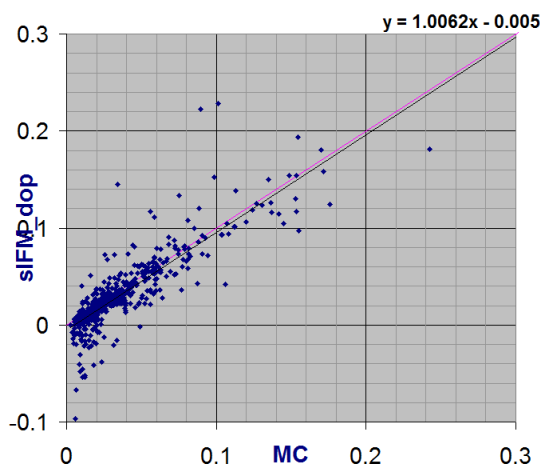


Figura 3.9: sIFM-doping vs MC in grafico lineare. Si vede bene come per bassi shift di soglia lo sIFM-trap calcoli spesso valori negativi

Ovviamente valori di  $\Delta V_T$  negativi non possono verificarsi per il caso sIFM-trap in quanto dalla transcaratteristica del dispositivo con trappola scarica simulata DD viene ricostruita la  $I_d - V_g$  del dispositivo con trappola carica calcolando la corrente come  $I_{d,\nu}(V_g) = I_{d,neutral\_trap}(V_g) - \delta I_{d,\nu}(V_g)$  dove il secondo termine risulta sempre positivo  $\delta I_{d,\nu}(V_g) = qG_d(r_\nu) > 0$ , discorso analogo per la  $V_{g,\nu}^i$  per la costruzione della curva sottosoglia.

Andare a vedere come il metodo sIFM-trap calcola in maniera approssimata le ampiezze delle  $\Delta V_T$  ci dà immediatamente idea di come ricostruisce anche le  $I_d - V_g$  singole. Per come è costruita tale metodologia, infatti, l'errore sta tutto nella ricostruzione della transcaratteristica per il dispositivo con trappola carica, e dunque gli errori nel calcolo degli shift di soglia risiedono solo in tali  $I_d - V_g$ . Il discorso per quanto riguarda il metodo sIFM-doping è invece differente. Un valore di shift di soglia perfettamente approssimante il DD non è sinonimo di una corretta ricostruzione delle due transcaratteristiche. In vero, può accadere che sbagliando la ricostruzione nel caso di trappola neutra e commettendo lo stesso errore nel caso di intrappolamento di elettrone, l'errore totale sulla  $\Delta V_T$  sia sostanzialmente nullo. Pertanto, a differenza del metodo sIFM-trap, per lo sIFM-doping lo scatterplot non ci dice la bontà della ricostruzione delle singole transcaratteristiche. Possiamo dire, perciò, che a parità di dispersione dello scatterplot il metodo sIFM-trap ricostruisce sicuramente meglio le  $I_d - V_g$  rispetto allo sIFM-doping.

Dopo aver visto e commentato gli scatterplot delle pagine precedenti per il caso di drogaggio uniforme ed aver sottolineato questi ultimi due aspetti, possiamo affermare come il metodo *sIFM-trap* approssimi generalmente meglio il riferimento (che qui assumiamo essere il *Monte Carlo-DD*) rispetto allo *sIFM-doping*. Ciò è confermato dalla banale osservazione che l'errore di approssimazione di tal metodo risiede solo nella linearizzazione dell'effetto ai contatti di un singolo atomo, ovvero la trappola, e non dell'effetto complessivo di decine/centinaia di atomi, come invece lo *sIFM-doping* tenta di fare. Questo è il motivo per cui la nuvola della Fig.3.8(a) risulta essere più stretta di quella della Fig.3.8(b). È pur vero, però, che nel calcolo dello shift di soglia indotto dall'intrappolamento di un elettrone, lo *sIFM-trap* linearizza proprio l'effetto di tale intrappolamento, mentre lo *sIFM-doping* no. Quest'ultimo, infatti, può approssimare male l'effetto che i droganti hanno, ma poi in media considerare in maniera corretta l'effetto della trappola all'interfaccia. Come è visibile dagli scatterplot, infatti, la nuvola della Fig.3.8(b) (*sIFM-doping* vs *Monte Carlo*) risulta essere più centrata sulla bisettrice rispetto a quella della Fig.3.8(a) (*sIFM-trap* vs *Monte Carlo*). Non possiamo, inoltre, generalizzare i risultati appena commentati sul drogaggio uniforme per altri dispositivi. Se per esempio consideriamo il caso di un drogaggio fortemente retrogrado o a  $\delta$ -shape, può accadere che lo *sIFM-doping* approssimi meglio il *Monte Carlo* dello *sIFM-trap*, in quanto l'effetto degli atomi droganti risulta essere molto minore e dunque la linearizzazione su di esso può rappresentare un'approssimazione eccellente, mentre ciò non è vero per quanto riguarda l'effetto sulla conduzione di una trappola (infatti è sempre presente il fenomeno del *current crowding*).

Nelle figure seguenti ho riportato alcuni casi particolari della statistica da poco descritta, ovvero ho focalizzato l'attenzione su 4 coppie configurazione di drogaggio - trappola. Quello di Fig.3.10(a) (caso 47-47) è semplicemente un caso particolare scelto in maniera randomica, dove le cose più o meno tornano:  $\Delta V_T^{DD} = 24.245mV$ ,  $\Delta V_T^{sIFM-trap} = 23.99mV$  e  $\Delta V_T^{sIFM-doping} = 28.076mV$ . Da notare come non è stato rappresentato il caso di trappola scarica per il metodo *sIFM-trap* (non c'è la curva blu solida), in quanto coincide con il DD (curva nera solida), come già ampiamente detto e motivato. Il dispositivo di Fig.3.11(a) (caso 84-81), invece, rappresenta il caso in cui il discostamento tra *sIFM-trap* e *Monte Carlo* è massimo ( $\Delta(\Delta V_T) = 103.892mV$ ): la trappola è posizionata a

$2nm$  dal bordo del dispositivo e il metodo sIFM sottostima fortemente lo shift di soglia risultante. Vedremo nei paragrafi successivi il perché di questa sottostima, andando a studiare gli effetti non lineari di una trappola all'interfaccia che fan sì che il metodo di ricostruzione *sIFM-trap* approssimi male il riferimento *Monte Carlo*. Per questo caso il metodo *sIFM-doping* approssima meglio il MC, nonostante la somma degli errori su entrambe le curve non sia di tanto minore. Come si vede dalla Fig.3.11(b), infatti, le curve rosse approssimano male le curve nere (soprattutto il caso a trappola neutra), ma l'errore sullo shift di soglia netto è ben minore di quello calcolato dal metodo *sIFM-trap*. Il device di Fig.3.12(a) (caso 68-65), invece, rappresenta il caso in cui il discostamento tra *sIFM-doping* e *Monte Carlo* è massimo ( $\Delta(\Delta V_T) = 133.268mV$ ): la trappola è posizionata a  $2nm$  dal bordo del dispositivo e il metodo sIFM sovrastima fortemente lo shift di soglia risultante. Ho mostrato questi ultimi due casi estremi per far capire come posizionando strategicamente la trappola su picchi di densità di corrente, per esempio sui bordi del dispositivo dove è forte l'effetto del *current crowding*, la linearità decade bruscamente sia per quanto riguarda l'effetto della trappola che l'effetto dei droganti, e i metodi linearizzati sIFM approssimano malissimo il riferimento *Monte Carlo - Drift Diffusion*. In realtà, lo vedremo, questo non è vero per quanto riguarda il metodo *sIFM-doping* nei casi in cui il drogaggio risulta lontano dalla trappola, per esempio per profili retrogradi. Infine il dispositivo di Fig.3.13(a) (caso 20-15) rappresenta un esempio in cui lo shift di soglia calcolato con il metodo *sIFM-doping* risulta essere fortemente negativo ( $\Delta V_T^{DD} = 6.355mV$ ,  $\Delta V_T^{sIFM-doping} = -66.685mV$ ). Questo dispositivo mostra come nel caso in cui i droganti siano in numero sbilanciati in un lato (quello alto nella figura) bloccando di fatto la conduzione a trappola scarica lungo quel bordo, posizionando la trappola in quella zona lo *sIFM-doping* tende a calcolare shift di soglia negativi. Ciò, come vedremo, è facilmente spiegabile e generalizzabile.

Come già detto, il dispositivo con drogaggio uniforme e spigoli a 90 gradi rappresenta uno dei casi peggiori da approssimare con questi metodi lineari, in quanto gli effetti non lineari di *current crowding* e di percolazione della corrente fan sì che le approssimazioni fatte dai metodi sIFM non risultano più valide per certe coppie configurazione di drogaggio - trappola. Quello che ci aspettiamo è che arrotondando gli spigoli o innalzando il drogaggio sui bordi, il primo di questi effetti (il quale è la principale causa di *Giant RTN*) diminuisca significativamen-

te e dunque lo *sIFM-trap* approssimi meglio il MC (linea di tendenza più vicina alla bisettrice e coefficiente di correlazione più vicino all'unità). Essendo che tal metodo linearizza l'effetto della trappola ai contatti, ci aspettiamo come una conduzione più uniforme nel canale permetta di approssimare meglio il riferimento MC. D'altro canto, prevediamo che adottando drogaggi retrogradi o a  $\delta$ -shape il metodo *sIFM-doping* migliori la propria approssimazione sul DD, in quanto l'effetto che hanno i droganti sui contatti è indubbiamente minore (essendo lontani dal canale) e dunque linearizzare ciò potrebbe risultare un'ottima approssimazione. Viceversa, adottare drogaggi più alti ai bordi o arrotondamenti agli spigoli della cella non è detto che porti miglioramenti nell'approssimazione del metodo *sIFM-doping*, anzi vedremo come addirittura ci siano peggioramenti. Discorso analogo per quanto riguarda drogaggi disuniformi verticalmente: non è detto che svuotando il canale di droganti il metodo *sIFM-trap* approssimi meglio il MC rispetto al caso uniforme, in quanto è vero che la percolazione viene meno, ma non c'è uniformità di conduzione per l'effetto punta sugli spigoli del dispositivo.

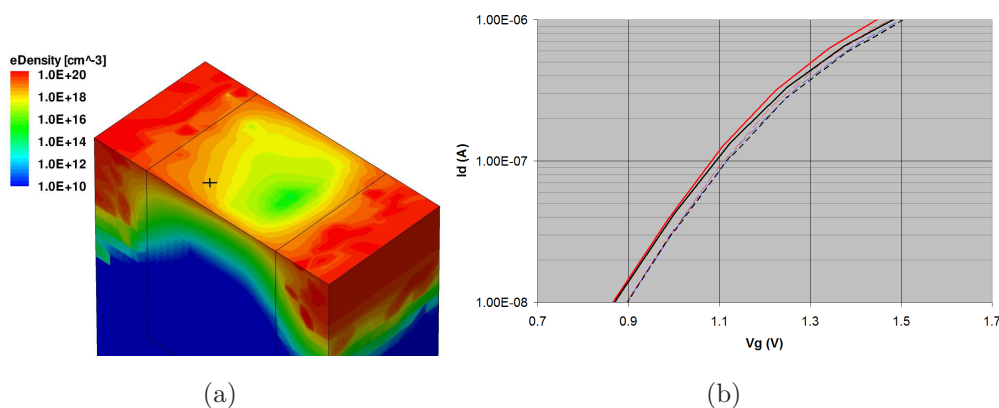


Figura 3.10: confronto diretto dei tre metodi per la coppia doping-trap 47-47. A sinistra è plottata la concentrazione di elettroni a trappola scarica (calcolata DD) e la posizione della trappola, a destra le transcaratteristiche simulate (DD, curve nere) e ricostruite (*sIFM-trap*, curve blu, e *sIFM-doping*, curve rosse): tratteggiati sono i casi in cui la trappola è carica, le curve solide sono a trappola neutra.

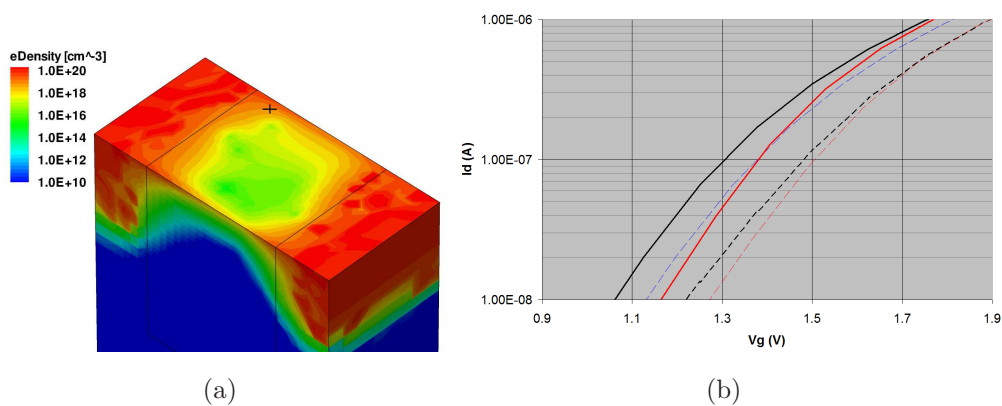


Figura 3.11: stesso delle Fig.3.10(a) e 3.10(b) ma per il caso 84-81

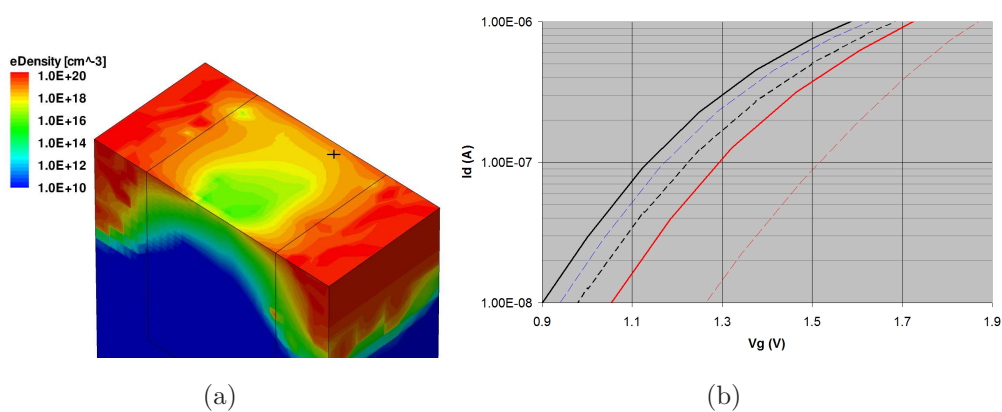


Figura 3.12: stesso delle Fig.3.10(a) e 3.10(b) ma per il caso 68-65

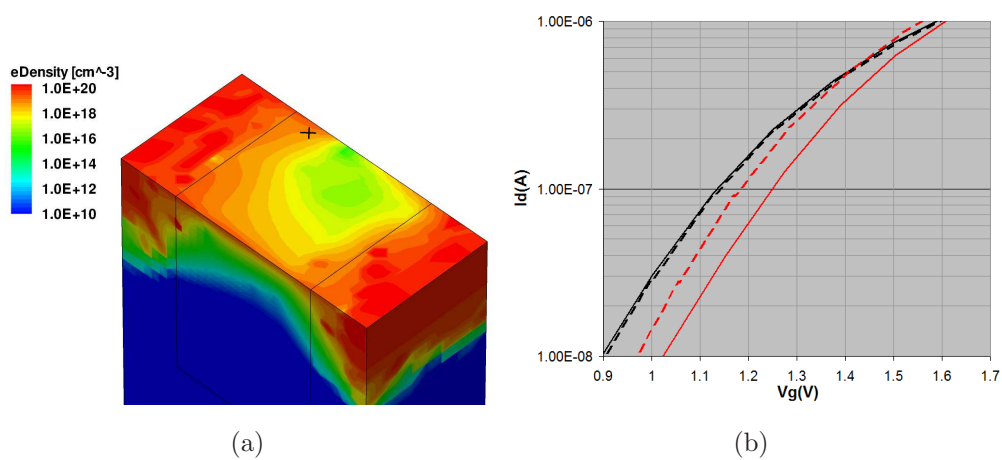


Figura 3.13: stesso delle Fig.3.10(a) e 3.10(b) ma per il caso 20-15



### $\Delta V_{CG}$ sotto e sopra soglia

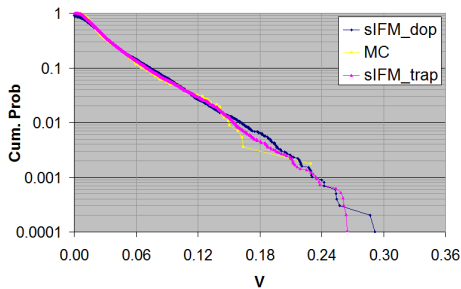
Finora abbiamo calcolato gli shift di soglia per  $I_{d_{th}} = 100nA \frac{W}{L} = 100nA$  ( $W = L = 32nm$ ) ovvero nella zona del “ginocchio” della transcaratteristica. Nessuno ci vieta di considerare gli shift di  $V_{CG}$  ad una  $I_{d_{th}}$  differente. Ho calcolato pertanto le  $\Delta V_{CG}$  per le stesse 568 coppie doping-trap di prima a due correnti diverse:  $I_{d_{th}} = 10nA$ , ovvero molto sotto-soglia, e  $I_{d_{th}} = 5\mu A$ , ovvero nel sopra-soglia. I risultati relativi a ciò sono visualizzabili nelle figure 3.14 e 3.15.

Innanzitutto ricordo, dal *weighted method* di ricostruzione delle  $I_d - V_g$ , come nel sotto-soglia le curve sono ricostruite con il metodo “dV” (2.71), e nel sopra-soglia con il “dI” (2.70). Inoltre, per ricostruire in maniera più fisica le transcaratteristiche, è bene adottare sotto-soglia un’interpolazione tra i bias-point esponenziale e sopra-soglia lineare. Questi due accorgimenti sono necessari e da non trascurare, spesso i risultati risulterebbero molto diversi in caso contrario.

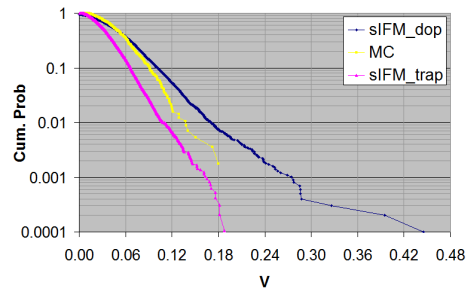
In Fig.3.14 sono mostrate le distribuzioni di probabilità cumulativa sotto-soglia e sopra-soglia per i metodi *sIFM-trap*, *sIFM-doping* e il riferimento MC. Quello che si nota è che per il caso  $I_{d_{th}} = 10nA$  i metodi sIFM approssimano molto bene il MC, mentre ciò non accade per il sopra-soglia.

Quello che ci aspettiamo è che la linearità valga un po’ di più nel sotto-soglia rispetto al caso  $I_{d_{th}} = 100nA$ , in quanto le  $\Delta V_{CG}$  risultano indubbiamente minori, quindi dovremmo trovare scatterplots più correlati e centrati sulla bisettrice del primo e del terzo quadrante. Vedendo le figure 3.15(c) e 3.15(d) e confrontandole con le Fig.3.8(a) e 3.8(b), si nota, infatti, un lieve miglioramento. Le cumulate sIFM relative ad esse sono praticamente coincidenti con quella MC (Fig.3.15(a)). Sopra-soglia ci aspettiamo che le cose funzionino ancora meglio. Infatti la Funzione di Green considerata è quella in corrente (la  $G_\phi^I$ ), e pertanto, dato che nel sopra-soglia il metodo di ricostruzione delle curve è il “dI”, non abbiamo anche gli errori relativi alla  $y_{dg}$ . Inoltre, in tal regione di funzionamento, la linearità dovrebbe risultare un’approssimazione migliore, poiché la conduzione nel canale è più uniforme e l’effetto delle trappole e degli ioni droganti risulta essere minore. Purtroppo, però, non accade proprio ciò che ci aspettavamo (Fig.3.14). Analizziamo il perché focalizzando l’attenzione sulla correlazione tra i metodi in un sottoinsieme di 568 casi (Fig.3.15). Per quanto riguarda il metodo *sIFM-trap* (Fig.3.15(e)), notiamo come non solo c’è meno correlazione con il MC, ma anche che la retta di regressione è ben distante dalla bisettrice. Infatti, tal

metodo sottostima fortemente la  $\Delta V_{CG}$  e accade che i punti della Fig.3.15(e) si posizionano lungo rette diverse a seconda della posizione della trappola. Vedere pertanto il sopra-soglia risulta essere fuorviante per lo *sIFM-trap*, e così è anche per gli altri dispositivi. La ragione di tal malfunzionamento sta nel fatto che la  $G_\phi^I$  calcolata è relativa solo all'equazione di Poisson, e non tiene conto pertanto delle fluttuazioni di scattering sulla corrente dovute agli atomi droganti. Per lo *sIFM-trap* ciò si traduce nel fatto che il metodo non valuta l'effetto della trappola sullo scattering, il quale, ad alti bias point, risulta non trascurabile. Esso sottostimerà sempre la  $\Delta V_{CG}$  sopra-soglia in quanto non considera la maggior causa di diminuzione di corrente in tal regione di funzionamento. Vedendo, invece, lo *sIFM-doping* (Fig.3.15(f)), notiamo come non si veda una gran differenza rispetto ai casi  $I_{dth} = 100nA$  e  $I_{dth} = 10nA$ , anzi osserviamo un lieve miglioramento. Ciò è dovuto al fatto che l'effetto della trappola sullo scattering ad alte polarizzazioni è comunque simulato in maniera DD, anche se su drogaggio continuo. Tal metodo, in vero, approssima solo l'effetto sui droganti, ma non sulla trappola. Se pensiamo per esempio di allontanare i droganti adottando un profilo retrogrado o a delta, dovremmo trovare un ottimo accordo anche nel sopra-soglia.

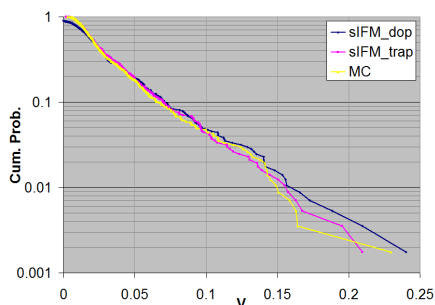


(a) cumulata sotto-soglia

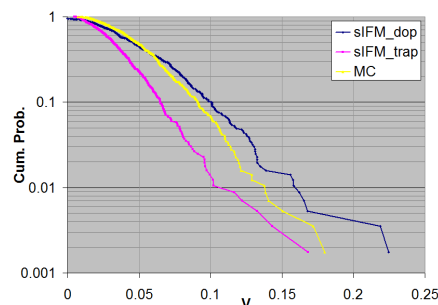


(b) cumulata sopra-soglia

Figura 3.14: distribuzioni di probabilità cumulativa costruite sui 10000 valori per i metodi *sIFM-trap*, *sIFM-doping* e sui 568 valori (subset dei 10000) per il *MC* nel caso di drogaggio uniforme in cui  $I_{dth} = 10nA$  (sotto-soglia) e  $I_{dth} = 5\mu A$  (sopra-soglia)



(a) cumulata sotto-soglia



(b) cumulata sopra-soglia

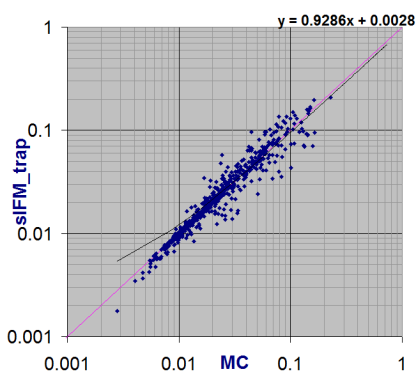
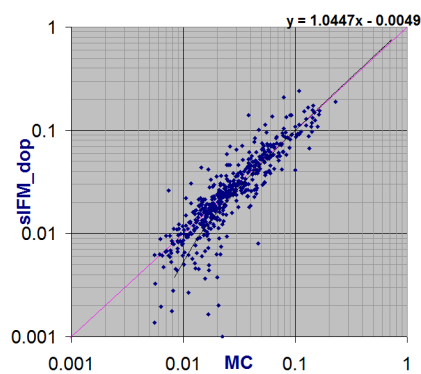
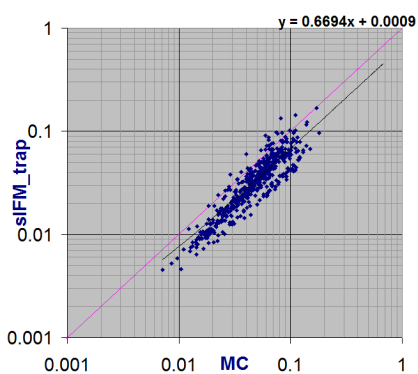
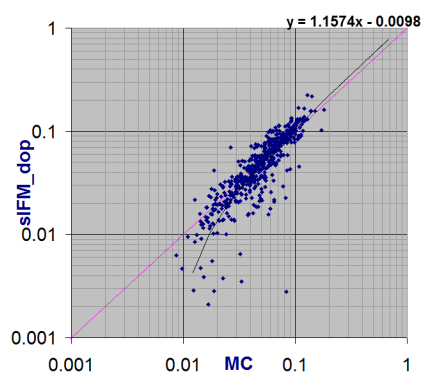
(c) sIFM\_trap-MC sotto-soglia:  $\rho = 0.916$ (d) sIFM\_doping-MC sotto-soglia:  $\rho = 0.844$ (e) sIFM\_trap-MC sopra-soglia:  $\rho = 0.833$ (f) sIFM\_doping-MC sopra-soglia:  $\rho = 0.862$ 

Figura 3.15: scatterplots su 568 valori relativi ai metodi *sIFM-trap*, *sIFM-doping* e *MC* nel caso di drogaggio uniforme, e le cumulate corrispondenti, in cui  $I_{d_{th}} = 10nA$  (sotto-soglia) e  $I_{d_{th}} = 5\mu A$  (sopra-soglia)

### 3.2.2 Drogaggi disuniformi lungo L e W

In questa sezione mostriamo e discutiamo i risultati dell'RTN per quanto riguarda i dispositivi in cui il drogaggio nominale è disuniforme lungo la lunghezza e la larghezza del canale come si vede in Fig.3.16.

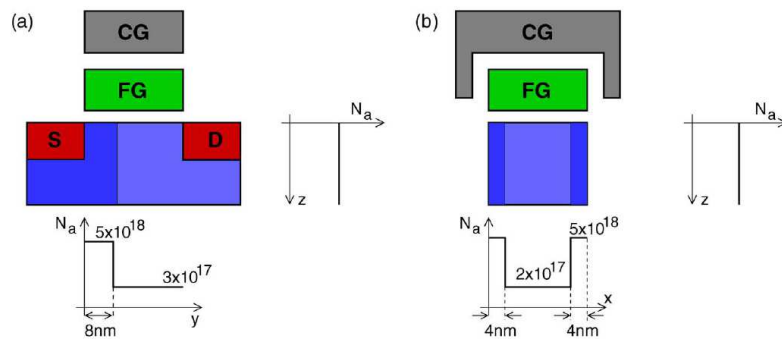
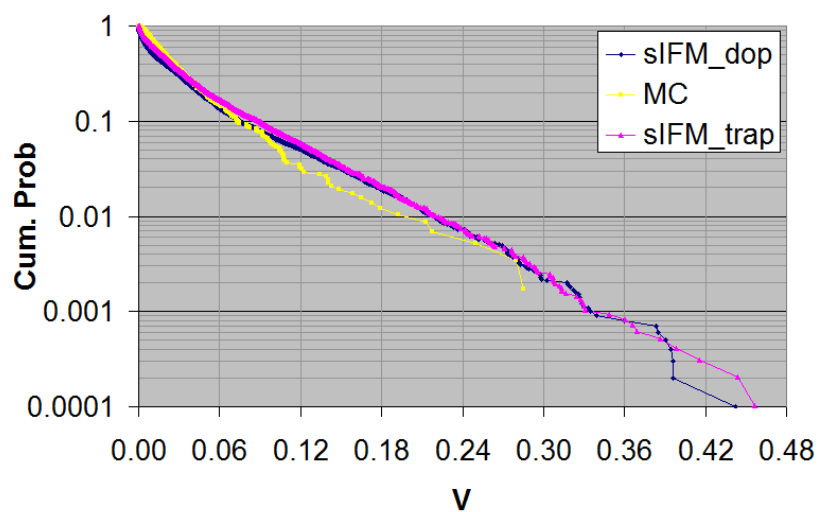


Figura 3.16: dispositivo floating-gate con profilo di drogaggio ideale disuniforme lungo (a) L e (b) W.

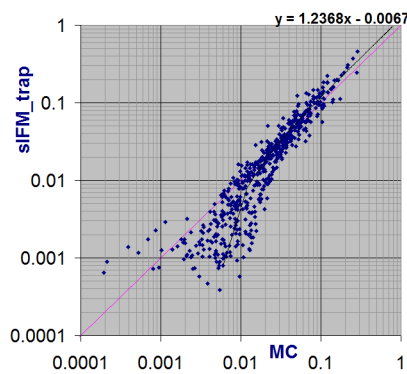
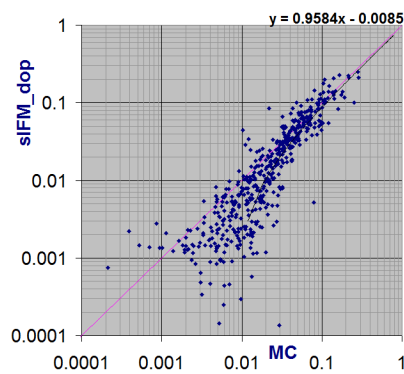
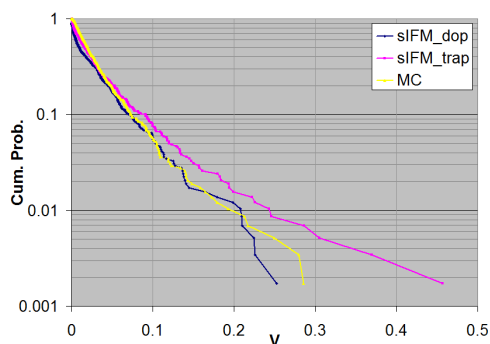
#### Drogaggio disuniforme lungo L

I risultati relativi a tal dispositivo sono visibili in Fig.3.17. Delle 10000 coppie doping-trap considerate ho costruito una statistica *Monte Carlo* selezionando un subset di 579 campioni. Per questi casi ho pertanto un confronto diretto con i metodi sIFM, come si vede dagli scatterplots di Fig.3.17(b), 3.17(c) e dalle cumulate di Fig.3.17(d).

Dalla teoria (Capitolo 1) sappiamo che adottare un profilo di drogaggio del genere significa allargare la cumulativa a basse probabilità e stringerla ad alte. A dir la verità, ciò si nota poco: il raccordo tra le due pendenze  $\lambda$  risulta essere più morbido ed avviene a probabilità molto alte. Sostanzialmente osserviamo una pendenza unica della cumulativa, che è dovuta al “dispositivo equivalente” ad alto drogaggio vicino al source. Il device, infatti, è interpretabile come due celle in serie, e dunque la statistica è dominata da quella con doping maggiore. Detto ciò, i metodi *sIFM-doping* e *sIFM-trap* dovrebbero approssimare in maniera



(a) cumulate complete

(b) sIFM\_trap-MC:  $\rho = 0.952$ (c) sIFM\_doping-MC:  $\rho = 0.807$ 

(d) cumulate costruite su 579 valori

Figura 3.17: scatterplots su 579 valori (subset dei 10000) relativi ai metodi *sIFM-trap*, *sIFM-doping* e *MC* nel caso di drogaggio disuniforme lungo la lunghezza del canale, le cumulate corrispondenti (sui 579) e quelle complete (sui 10000)

analoga a quanto visto per il caso di doping uniforme. Se come abbiamo visto questi metodi approssimano bene le distribuzioni di RTN per una cella a drogaggio accettore uniforme, lo faranno anche per due in serie. Guardando le figure 3.17(b), 3.17(c) e 3.17(a) troviamo esattamente i confronti che ci aspettavamo. Ad essere precisi, focalizzandoci sugli scatterplots ed i coefficienti di correlazione, notiamo come per il caso *sIFM-doping* ci sia un peggioramento di approssimazione rispetto al caso uniforme. Ciò è difficilmente interpretabile. Guardando la funzione di Green osserviamo che risulta avere valori nettamente più alti per il dispositivo “alto-drogato”, il che ce lo aspettavamo. Posizionando una trappola in quella regione osserviamo ampi  $\Delta V_T$  e notiamo che essa è circondata da un maggior numero di atomi droganti rispetto a prima, il che probabilmente induce effetti non lineari elevati portando il metodo a commettere errori marcati di approssimazione. In altre parole, abbiamo shift di soglia molto grandi che sono difficili da approssimare con metodi linearizzati. Guardando la retta di regressione dello scatterplot, inoltre, notiamo come per bassi valori dello shift di soglia il metodo sottostimi sempre di più tal valore, fino ad assumere valori negativi non fisici (analogamente al caso uniforme infatti). In vero, dei 10000 valori di  $\Delta V_T$  calcolati ne troviamo ben 1065 negativi, ovvero il 10.65%. Ciò non si vede dalle cumulative, in quanto incide solo ad alte probabilità (sostanzialmente la curva inizia a  $\approx 0.9$ ).

Per quanto riguarda invece lo *sIFM-trap* osserviamo viceversa una correlazione maggiore, ma su una retta di regressione, nella Fig.3.17(b), con coefficiente angolare  $\alpha = 1.24 > 1$ . In media, insomma, il metodo sovrastima lo shift di soglia. Questo, secondo me, è attribuibile al fatto che l’interazione dei droganti con la trappola ha l’effetto di stringere i percorsi conduttivi della corrente e ciò, come vedremo, induce un’inevitabile sovrastima dello shift di soglia nei casi di *Giant RTN*. Aumentare i droganti rispetto al caso precedente per il dispositivo “alto drogato” (ovvero quello che decide le basse probabilità della statistica) dovrebbe pertanto indurre in media una sovrastima della  $\Delta V_T$  per alti valori della stessa. La correlazione, invece, aumenta per il semplice fatto che la maggior parte delle trappole cade nella regione a basso drogaggio, dove l’errore è minore (ciò non sembra dalla Fig.3.17(b), ma quel grafico è in scala bilogarithmica).

Sostanzialmente per questo dispositivo abbiamo che lo *sIFM-trap* approssima meglio il MC dello *sIFM-doping*, ma costruendo le cumulative quest’ultimo approssima leggermente meglio la pendenza del MC, in quanto il primo sovrastima

le  $\Delta V_T$  nei casi di *Giant RTN*.

Mostro qui sotto le  $I_d - V_g$  per i tre metodi relative ad alcuni campioni della statistica appena descritta. Quello che si nota è che il metodo *sIFM-trap* costruisce le transcaratteristiche molto meglio dello *sIFM-doping*, il che è in linea a ciò che ci aspettavamo.

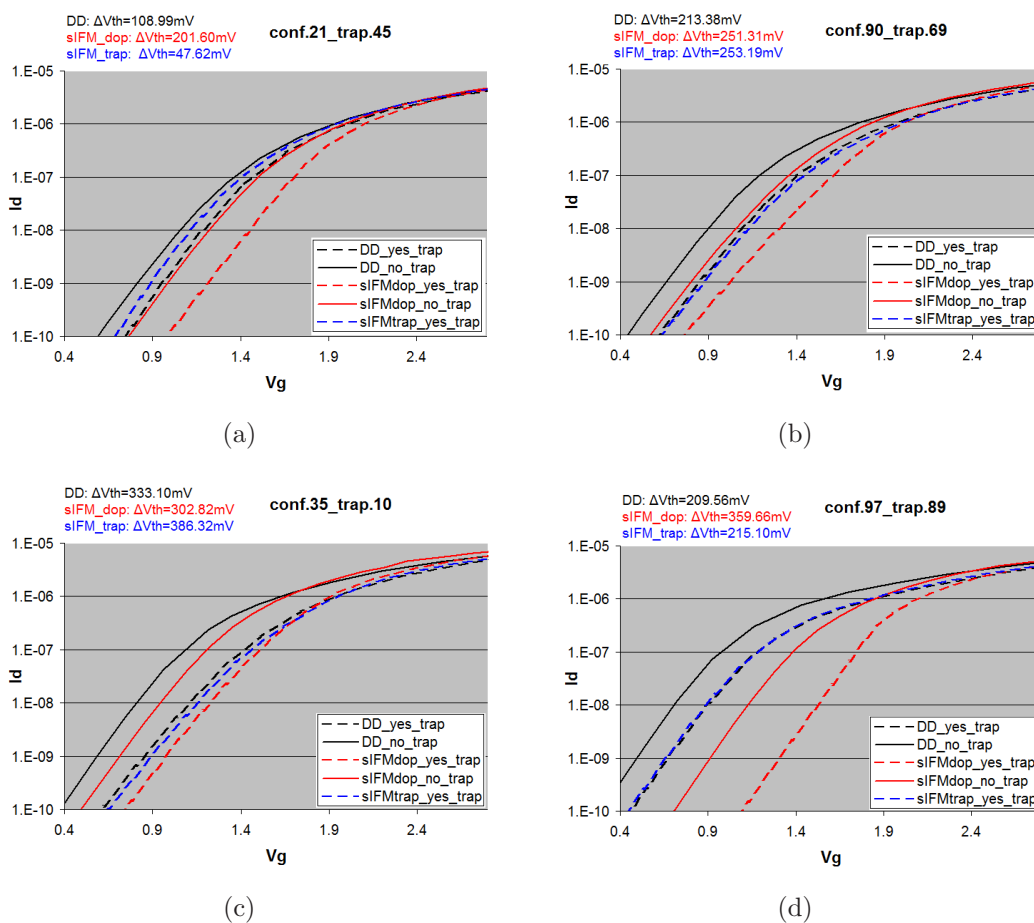


Figura 3.18: alcuni casi doping-trap particolari per il dispositivo con drogaggio disuniforme lungo L. Sono mostrate le transcaratteristiche ottenute con i diversi metodi. Le curve nere sono simulate DD, quelle ricostruite con i metodi *sIFM-doping* e *sIFM-trap* sono rispettivamente le curve rosse e blu. Le curve solide sono i casi a trappola neutra, le tratteggiate a trappola carica.

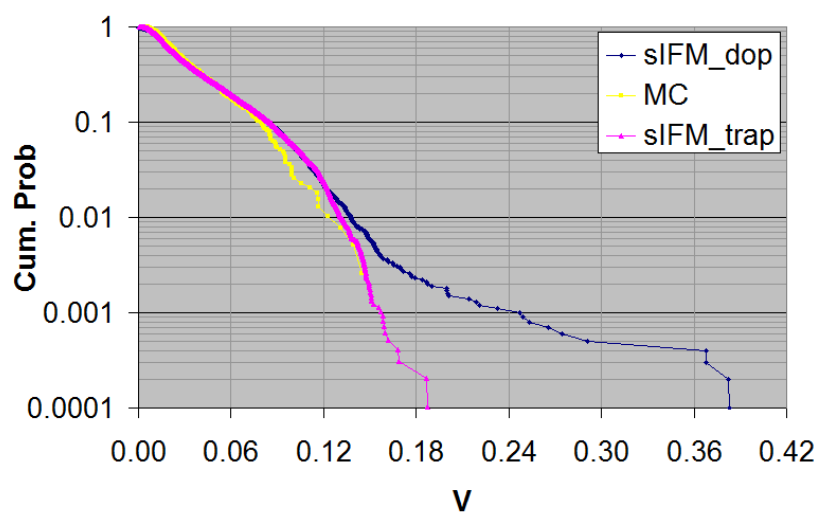
### Drogaggio disuniforme lungo W

I risultati relativi a tal dispositivo sono visibili in Fig.3.19. Delle 10000 coppie doping-trap considerate, ho costruito una statistica *Monte Carlo* selezionando un subset di 392 campioni. Per questi casi ho pertanto un confronto diretto con i metodi sIFM, come si vede dagli scatterplots di Fig.3.19(b), 3.19(c) e dalle cumulative di Fig.3.19(d).

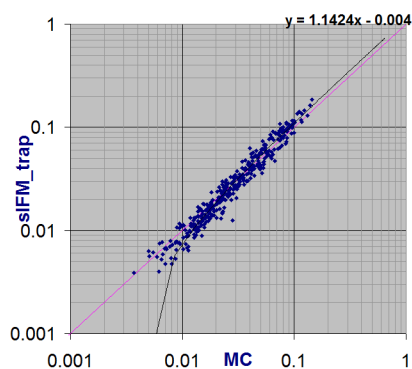
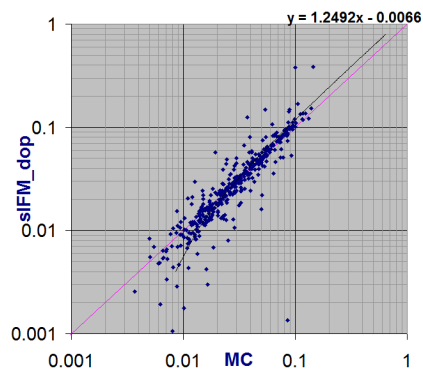
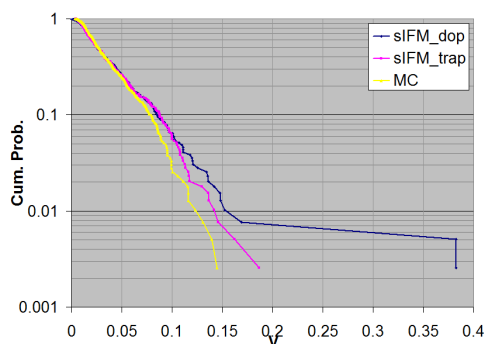
Osservando le immagini, notiamo come per il metodo *sIFM-trap* abbiamo un netto miglioramento. Questo è assolutamente in linea con le aspettative. Infatti, come avevamo anticipato, innalzare il drogaggio localmente sui bordi uniformizza la densità di corrente nel canale, pertanto la linearizzazione dell'effetto della trappola sulla conduzione risulta un' eccellente approssimazione. In vero abbiamo una correlazione con il MC quasi unitaria ( $\approx 0.98$ ) e praticamente sulla bisettrice. Una leggera sovrastima rimane, ma comunque minore rispetto a prima.

Per quanto riguarda il metodo *sIFM-doping* le cose vanno diversamente. Abbiamo visto come adottare drogaggi non costanti lungo L non cambia sostanzialmente l'approssimazione sul MC-DD, mentre adottare doping disuniformi lungo la larghezza del canale impatta in maniera significativa la cumulativa. In vero, guardando gli scatterplot, le cose non sembrano andar particolarmente peggio di prima, anzi. Se però guardiamo tal grafico in scala lineare (Fig.3.20(a)) osserviamo come ci siano 2 dispositivi "sfortunati" in cui la linearità decade terribilmente ed il metodo sovrastima enormemente lo shift di soglia. Questi casi molto estremi sono presenti anche nei dispositivi con doping uniforme e disuniforme lungo L, ma poiché le distribuzioni di probabilità cumulativa sono molto larghe, essi si "mimetizzano" impattando pochissimo sulla forma di tale curva. In questo caso, invece, dove abbiamo un forte stringimento della cumulativa a basse probabilità, essi risultano "fuori statistica" determinando le visibili code della Fig.3.19(a). Tali code risultano dunque assolutamente non fisiche e sono dettate esclusivamente da errori di linearizzazione intrinseci del metodo. Pertanto per tal dispositivo lo *sIFM-doping* approssima male il MC-DD a basse probabilità. Discorso diverso, invece, sarà per il caso retrogrado, in quanto abbiamo sì uno stringimento delle cumulative a basse probabilità ma non abbiamo queste coppie doping-trap "sfortunate" che determinano un allargamento errato di tali curve, in quanto gli ioni accettori risultano essere più in profondità e la linearizzazione dell'effetto sulla conduzione dei droganti risulta essere buona.



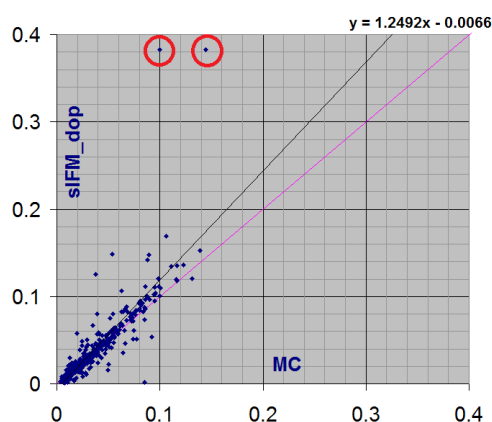


(a) cumulate complete

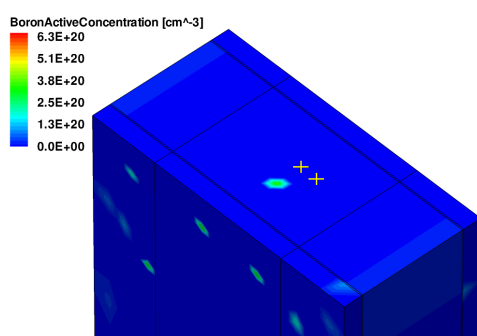
(b) sIFM\_trap-MC:  $\rho = 0.976$ (c) sIFM\_doping-MC:  $\rho = 0.839$ 

(d) cumulate costruite su 392 valori

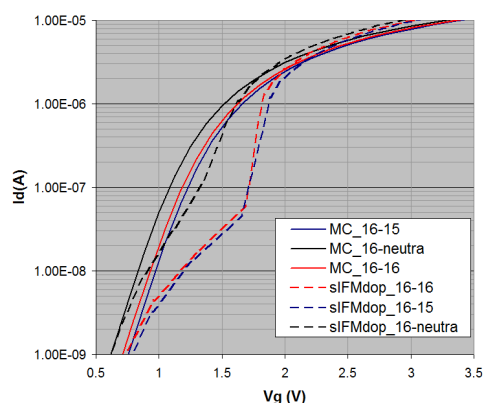
Figura 3.19: scatterplots su 392 valori (subset dei 10000) relativi ai metodi *sIFM-trap*, *sIFM-doping* e *MC* nel caso di drogaggio disuniforme lungo la larghezza del canale, le cumulate corrispondenti (sui 392) e quelle complete (sui 10000)



(a) Fig.3.19(d) vista in scala lineare. cerchiati in rosso i due casi “sfortunati” esaminati: caso 16-15 e 16-16



(b) drogaggio atomistico “16”. In giallo la posizione delle trappole 15 e 16



(c) transcaratteristiche relative ai casi 16-neutra, 16-15 e 16-16 simulate (DD) e ricostruite (sIFM-doping)

Figura 3.20: due casi particolari della statistica relativa alla Fig.3.19(d)

Consideriamo adesso i due casi più estremi per il metodo *sIFM-doping* (Fig.3.20(b) e 3.20(c)). Dalla figura 3.20(b) osserviamo come ci sia uno ione accettore sul centro-basso del dispositivo. Essendo che per il device la corrente scorre soprattutto centralmente (avendo alzato il doping sui bordi), in questo caso avrò che la densità di corrente devierà nel centro-alto della cella. Posizionando la trappola proprio in quella zona, indurrò *Giant RTN*. Per questi casi estremi lo *sIFM-doping* sovrastima terribilmente la  $\Delta V_T$  costruendo transcaratteristiche ad “S” (Fig. 3.20(c)), il che non è fisico, come discusso più avanti al par.3.2.4.

### 3.2.3 Drogaggi disuniformi verticalmente

In questa sezione mostriamo e discutiamo i risultati dell'RTN per quanto riguarda i dispositivi in cui il drogaggio nominale è disuniforme verticalmente, ovvero drogaggi retrogradi e a  $\delta$ -shape come si vede in Fig.3.21.

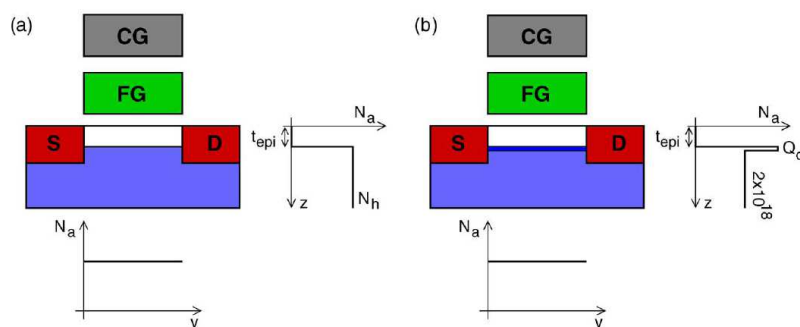
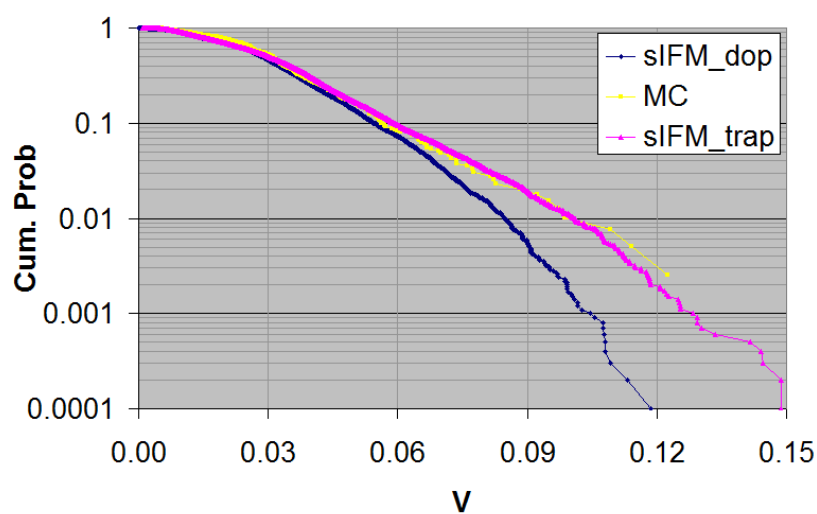


Figura 3.21: dispositivo floating-gate con profilo di drogaggio ideale (a) retrogrado e (b) a  $\delta$ -shape.

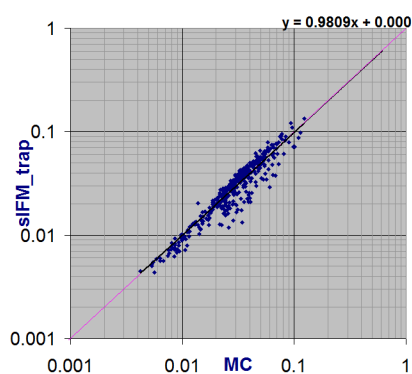
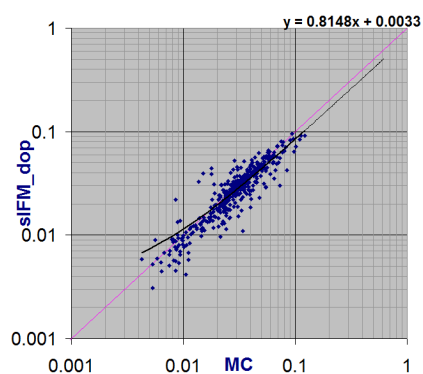
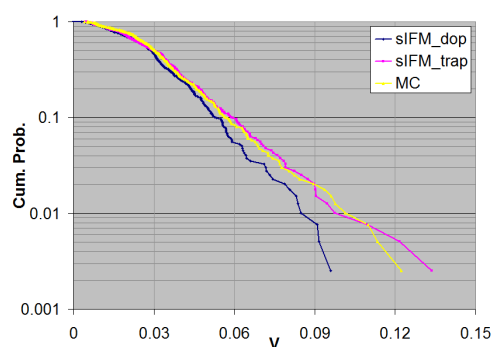
#### Drogaggio retrogrado: $t_{epi} = 16nm$

Iniziamo dapprima a mostrare e discutere i risultati per quanto riguarda il drogaggio retrogrado con svuotamento del canale fino a  $t_{epi} = 16nm$  di profondità. Essi sono riportati in Fig.3.22.

Dalla Fig.3.22(d) osserviamo come facendo un confronto diretto tra i tre metodi per gli stessi dispositivi, quello che approssima meglio il MC risulta essere nettamente lo *sIFM-trap*. Lo *sIFM-doping* è molto correlato con i risultati DD ( $\rho = 0.904$ ), ma su una linea di tendenza sub-lineare  $y = 0.8148x + 0.0033$  (facendo riferimento alla Fig.3.22(c)), il che porta inevitabilmente a ricostruire una distribuzione di probabilità cumulativa leggermente schiacciata a  $\Delta V_T$  più basse. L'approssimazione sul MC di tale curva per il metodo *sIFM-doping* non risulta migliorare rispetto al caso di drogaggio uniforme e disuniforme lungo L, al contrario delle aspettative. Allontanando gli ioni accettori dal canale, ci saremmo aspettati un'approssimazione migliore, in quanto i droganti risultano avere un



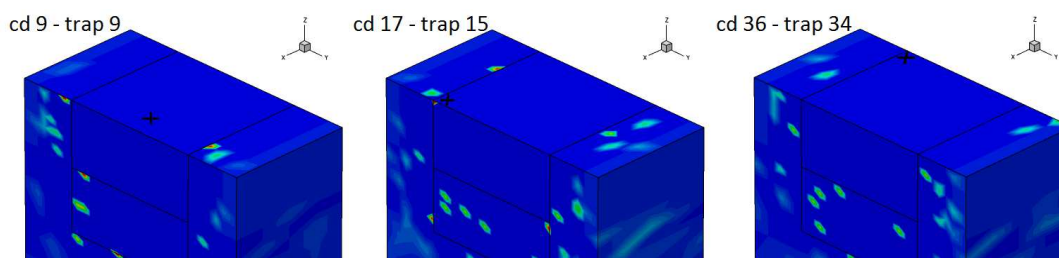
(a) cumulate complete

(b) sIFM\_trap-MC:  $\rho = 0.928$ (c) sIFM\_doping-MC:  $\rho = 0.904$ 

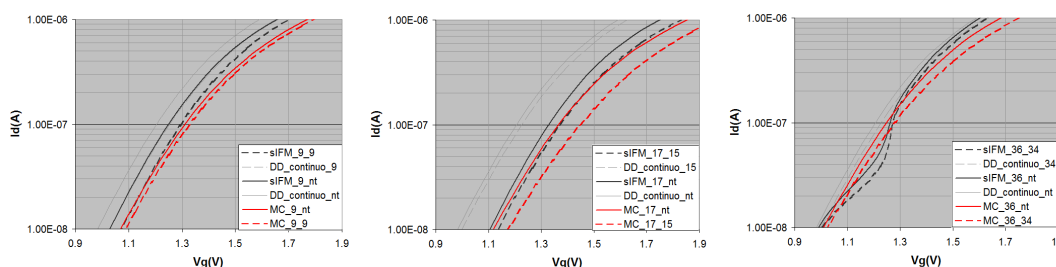
(d) cumulate costruite su 396 valori

Figura 3.22: scatterplots su 396 valori (subset dei 10000) relativi ai metodi *sIFM-trap*, *sIFM-doping* e *MC* nel caso di drogaggio retrogrado con  $t_{epi} = 16nm$ , le cumulate corrispondenti (sui 396) e quelle complete (sui 10000)

effetto minore sulla conduzione e dunque la linearizzazione dovrebbe essere corretta. Questo sembra non accadere. In realtà, però, non è vero che i droganti sono spariti dalla superficie, rimangono quelli ai lati del canale, ovvero quelli dentro le zone  $n^+$  per intenderci, che influenzano non poco la conduzione. Essi, infatti, inducono effetti non lineari sulla corrente di entità molto maggiore di quanto facciano gli ioni a profondità  $\geq 16nm$ . Inoltre essi risultano essere numerosi, in quanto la concentrazione nominale del substrato è stata alzata ad  $N_A = 10^{19}cm^{-3}$  per mantener la stessa soglia media. Tutto ciò è molto evidente osservando i tre casi particolari di doping atomistico / posizione trappola della Fig.3.23, in cui la linearizzazione del metodo *sIFM-doping* risulta approssimare molto male il DD. Nei casi 17-15 e 36-34, in particolare, un singolo atomo di boro ionizzato risulta essere a distanza di  $2nm$  dalla trappola (per il caso 36-34 non si vede in quanto è in profondità di  $1.37nm$ ), il che impatta violentemente sulla transcaratteristica facendo decadere l'ipotesi di linearità (Fig.3.23(b)). Ciò accade analogamente inserendo un atomo sul bordo della cella (caso 9-9). Questi tre casi dimostrano come affinché lo *sIFM-doping* risulti una buona approssimazione è necessario svuotare completamente il canale dal drogaggio accettore, anche nelle zone laterali  $n^+$ .



(a) densità di corrente a soglia per tre campioni della statistica a trappola scarica; la posizione di quest'ultima è evidenziata nell'immagine



(b) transcaratteristiche risultanti per questi 3 casi

Figura 3.23: 3 casi particolari. Confronto *sIFM-doping*-MC

Nella Fig.3.24 mostro le distribuzioni di probabilità cumulativa per i tre metodi per la stessa statistica precedente ma nel sotto-soglia ( $I_{d_{th}} = 10nA$ ). Quello che si vede è che lo *sIFM-doping* risulta approssimare meglio il MC. Questo ci fa capire come passare da una regione di funzionamento all'altra del dispositivo possa far cambiare la bontà delle approssimazioni in gioco: a  $I_{d_{th}} = 100nA$  sembrava molto meglio lo *sIFM-trap*, ora lo *sIFM-doping*.

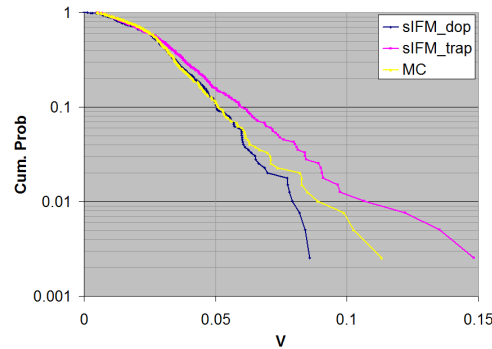
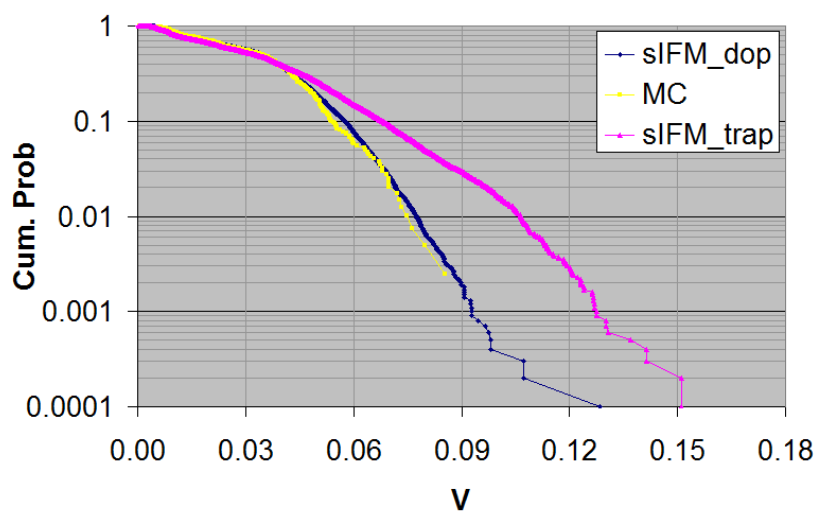
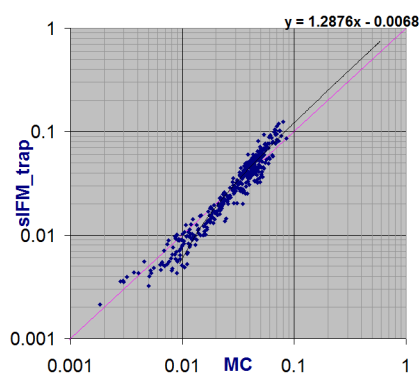
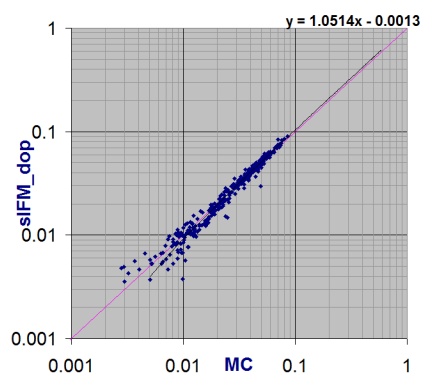
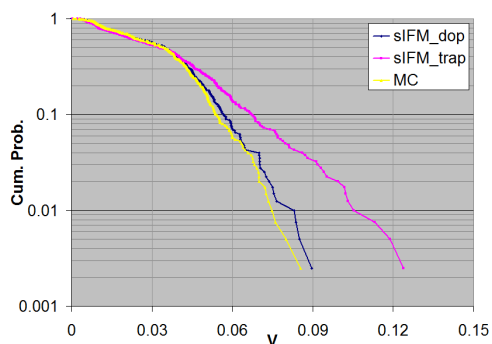


Figura 3.24: cumulate costruite sulle stesse 396 coppie doping-trap della Fig.3.22(d), ma con  $I_{d_{th}} = 10nA$ .

Ho simulato, inoltre, nuove coppie doping-trap sempre per il dispositivo a drogaggio retrogrado  $t_{epi} = 16nm$ , ma in cui lo svuotamento del canale dagli atomi accettori è totale, ovvero anche nelle regioni laterali  $n^+$  (il che è anche più rappresentativo della realtà perché la zona superficiale undoped sarà presente su tutto il wafer). In figura 3.25 ho riportato i risultati. In perfetta linea con le aspettative, si osserva una perfetta correlazione 1:1 con  $\rho > 0.99$  tra lo *sIFM-doping* e il MC-DD. Da questi risultati (e da quelli relativi al  $\delta$ -shape) consiglio fortemente di adottare simulazioni DD attivando il metodo *sIFM-doping* qualora il dispositivo avesse il canale completamente svuotato dai droganti ionizzati. A differenza dello *sIFM-doping*, il metodo *sIFM-trap* risente fortemente degli effetti di bordo. Mentre il primo simula in maniera DD l'effetto che ha una trappola qualora fosse posizionata sullo spigolo, il secondo linearizza ciò, portando l'approssimazione a sovrastimare molto i *Giant RTN*. Per questi tipi di dispositivi, dunque, lo *sIFM-doping* funziona molto meglio dello *sIFM-trap*. Nei grafici Fig.3.25(b) e 3.25(c) non ci si rende conto di cosa succede ad alti valori della  $\Delta V_T$ , in quanto bilogarithmici. Pertanto li ho riportati in scala lineare: 3.26(a), 3.26(b).



(a) cumulate complete

(b) sIFM\_trap-MC:  $\rho = 0.945$ (c) sIFM\_doping-MC:  $\rho = 0.991$ 

(d) cumulate costruite su 400 valori

Figura 3.25: scatterplots su 400 valori (subset dei 10000) relativi ai metodi *sIFM-trap*, *sIFM-doping* e *MC* nel caso di drogaggio retrogrado con  $t_{epi} = 16nm$  nel caso di completo svuotamento del canale, le cumulate corrispondenti (sui 400) e quelle complete (sui 10000)

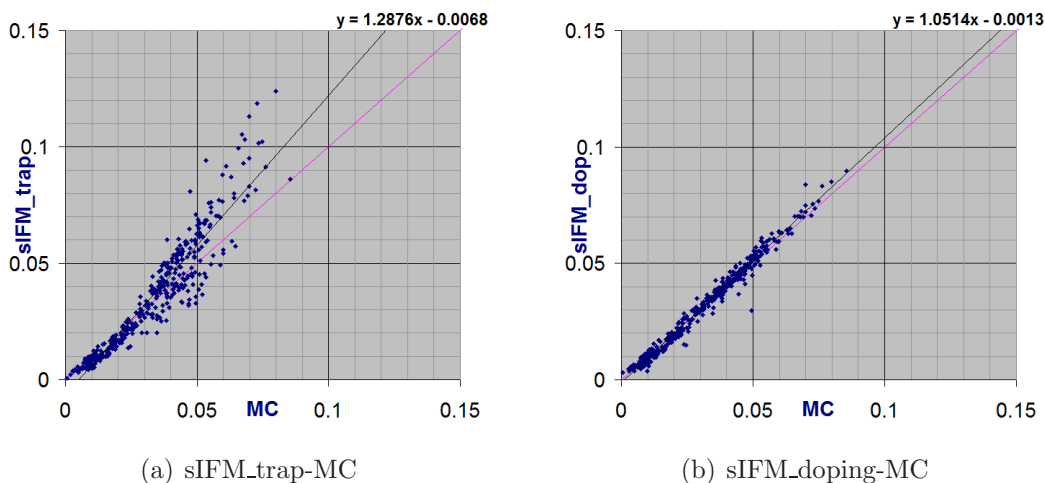
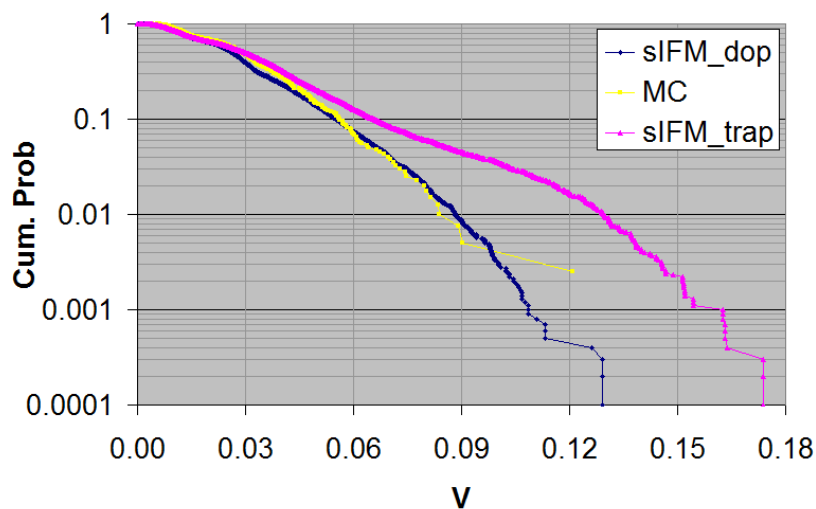


Figura 3.26: Fig.3.25(b) e 3.25(c) in scala lineare

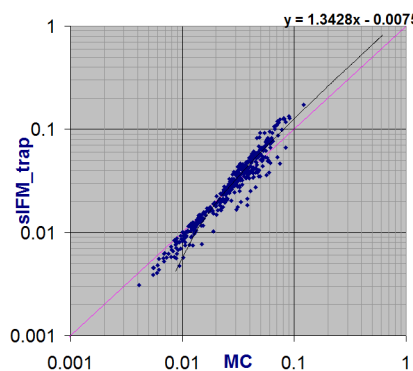
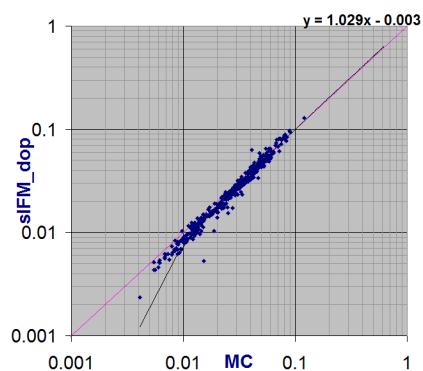
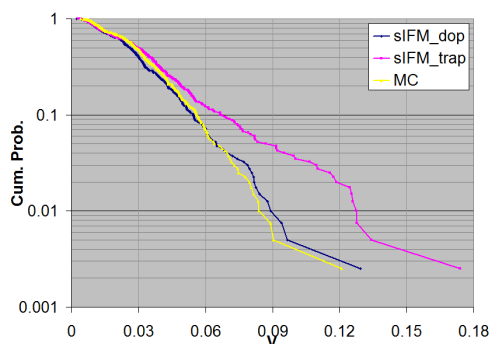
### Drogaggio a delta: $t_{epi} = 16nm$

Iniziamo dapprima a mostrare e discutere i risultati per quanto riguarda il drogaggio  $\delta$ -shape con svuotamento del canale fino a  $t_{epi} = 16nm$  di profondità. Essi sono riportati in Fig.3.27. Quello che si osserva banalmente è che questi risultati assomigliano molto più all'ultimo caso di retrogrado  $t_{epi} = 16nm$  completamente “svuotato” che al caso retrogrado in cui abbiamo lasciato i droganti nelle regioni  $n^+$ . Come mai? Il motivo è semplicissimo: adottando il drogaggio a delta abbiamo imposto un doping nel substrato più basso rispetto al caso retrogrado equivalente (per mantenere la stessa soglia media) e quindi nelle regioni laterali ci cadono davvero pochi droganti, che pertanto influiscono poco sulla conduzione. Questo implica che la linearizzazione dell'effetto degli atomi droganti sulla conduzione risulta un'approssimazione sicuramente migliore rispetto al caso “equivalente” retrogrado. I risultati ottenuti assomigliano a quelli di prima: *sIFM-doping* perfettamente correlato con il MC ( $\rho$  quasi unitario e linea tendenza vicina alla bisettrice del primo e terzo quadrante), mentre il metodo *sIFM-trap* anche in questo caso “soffre” le trappole posizionate vicine agli spigoli, ovvero sovrastima gli alti valori di  $\Delta V_T$ .





(a) cumulate complete

(b) sIFM\_trap-MC:  $\rho = 0.936$ (c) sIFM\_doping-MC:  $\rho = 0.985$ 

(d) cumulate costruite su 400 valori

Figura 3.27: scatterplots su 400 valori (subset dei 10000) relativi ai metodi *sIFM-trap*, *sIFM-doping* e *MC* nel caso di drogaggio a delta con  $t_{epi} = 16nm$ , le cumulate corrispondenti (sui 400) e quelle complete (sui 10000)

## Altri esempi

Mostro qui altri grafici (Fig.3.28(a), 3.28(b), 3.28(c) e 3.28(d)) che confrontano i metodi *sIFM-doping* e MC per drogaggi retrogradi e a delta a diversi  $t_{epi}$ . Questa volta i dispositivi simulati MC non sono un subset dei 10000 casi dello *sIFM-doping*, ma sono stati presi da [19]. Cambiano la mesh e le coppie doping-trap, ma non la fisica, la geometria e i drogaggi. Pertanto il confronto si può fare e quello che si nota è un ottimo accordo con il metodo sIFM sulle cumulative.

Nelle Fig.3.29 e 3.30 ho riportato tutti i risultati *sIFM-doping* per i profili verticali alle diverse profondità in cui metto il drogaggio. Quello che si nota, in perfetta linea con la teoria, è che aumentando  $t_{epi}$  le cumulative si stringono a bassi valori di shift di soglia.

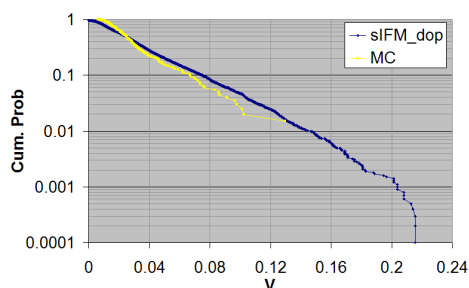
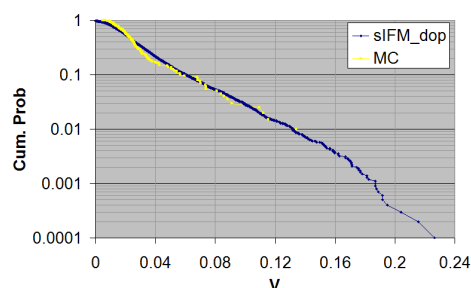
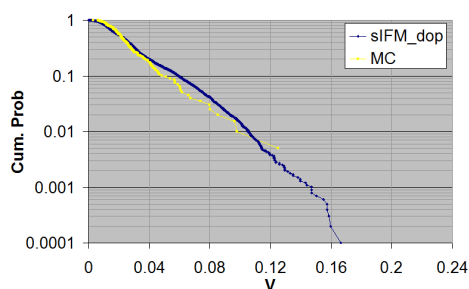
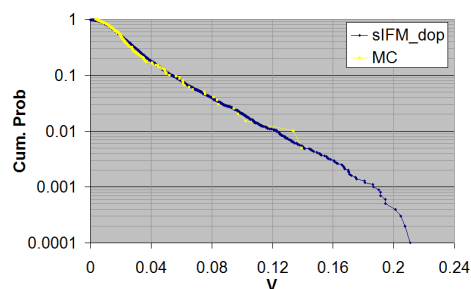
(a) drogaggio retrogrado  $t_{epi} = 2.5nm$ (b) drogaggio retrogrado  $t_{epi} = 5nm$ (c) drogaggio retrogrado  $t_{epi} = 10nm$ (d) drogaggio  $\delta$ -shape  $t_{epi} = 10nm$ 

Figura 3.28: distribuzioni di probabilità cumulative per altri drogaggi disuniformi verticali. Confronto sIFM\_doping-MC

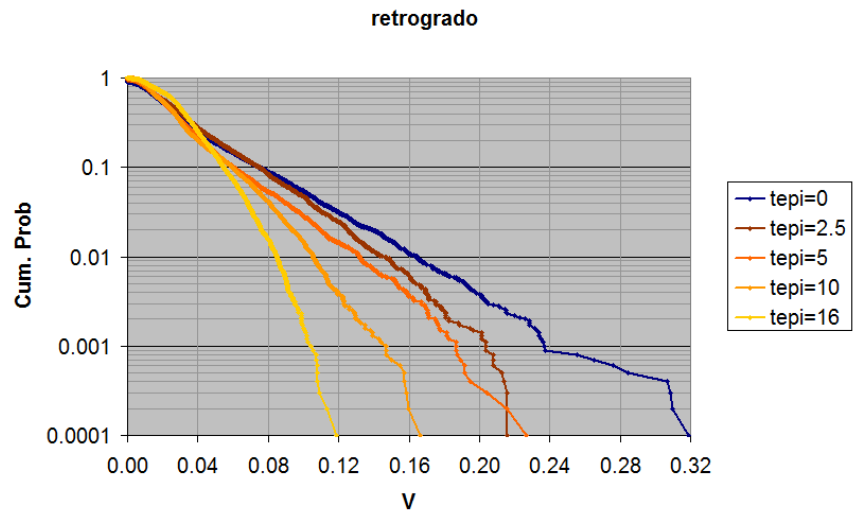


Figura 3.29: cumulative *sIFM-doping* per drogaggio retrogrado alle diverse  $t_{epi}$  ( $t_{epi} = 0$  vuol dire drogaggio uniforme).

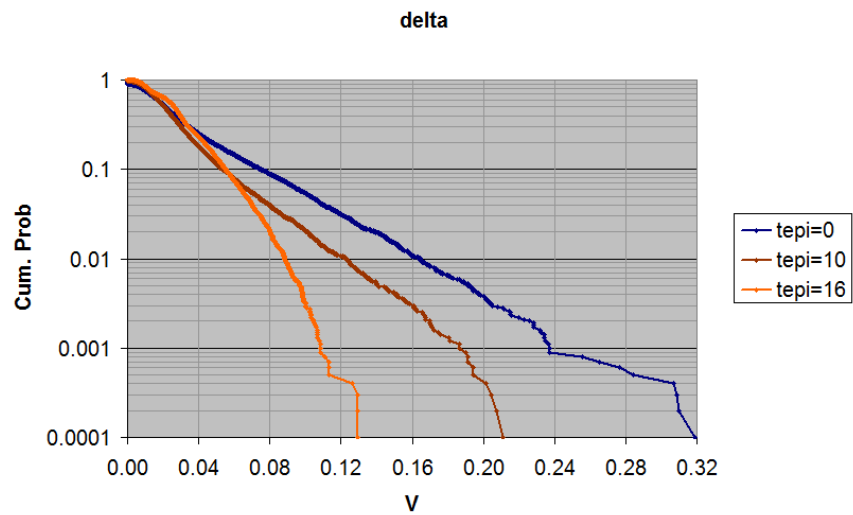


Figura 3.30: cumulative *sIFM-doping* per drogaggio  $\delta$ -shape alle diverse  $t_{epi}$  ( $t_{epi} = 0$  vuol dire drogaggio uniforme).

### 3.2.4 Errori di linearizzazione dei metodi sIFM

Nei paragrafi precedenti ci siamo limitati a mostrare solo alcuni esempi di ricostruzione delle curve con i metodi sIFM. Come abbiamo sottolineato, quello che ci aspettiamo è che il metodo *sIFM-trap* funzioni bene quando la conduzione nel canale è particolarmente uniforme (drogaggio spinto sui bordi o celle con spigoli arrotondati), in quanto esso linearizza l'effetto dell'intrappolamento di un elettrone nel suo cammino da source a drain, e che lo *sIFM-doping* risulti essere una buona approssimazione nel caso in cui i droganti abbiano un effetto lieve sulla conduzione e dunque linearizzabile (drogaggi retrogradi o  $\delta$ -shape). Questo è più o meno quanto ci aspettiamo in modo intuitivo e qualitativo, ovvero quanto prevediamo che accada mediamente nel calcolo delle  $\Delta V_T$ . Diventa assai difficile, invece, prevedere cosa possa succedere per i singoli dispositivi, per le singole coppie configurazione di drogaggio - trappola. In questo paragrafo tenteremo pertanto di studiare per alcuni casi gli effetti non lineari sulla conduzione del drogaggio atomistico e della trappola. Mi limiterò a mostrare alcuni esempi cercando di interpretarli, senza però dare alcuna soluzione di previsione data la complessità degli effetti in questione.

#### sIFM-doping: effetti non lineari del drogaggio

Il metodo di ricostruzione *sIFM-doping* linearizza l'effetto di ogni atomo drogante ai contatti. Poiché gli atomi presenti sono decine/centinaia, diventa impossibile prevedere quanto sia l'entità dell'errore per tal linearizzazione, perché dovremmo considerare l'effetto (non lineare) congiunto di ognuno di essi. In altre parole, dati un profilo atomistico di doping e una trappola all'interfaccia non si può assolutamente dire a priori quanto sarà l'errore dovuto alla linearizzazione.

Ho qui riportato due esempi molto significativi, sottolineati anche nella pubblicazione [25] e mostrati in Fig.3.31 e Fig.3.32. Possiamo trovare due casi estremi in cui la linearizzazione sull'effetto dei droganti decade completamente: nella Fig.3.31b) abbiamo il caso (doping uniforme) in cui gli atomi droganti sono posizionati soprattutto nello spigolo inferiore del dispositivo mentre la trappola in quello superiore; nella Fig.3.31a) (caso drogaggio disuniforme lungo L) abbiamo

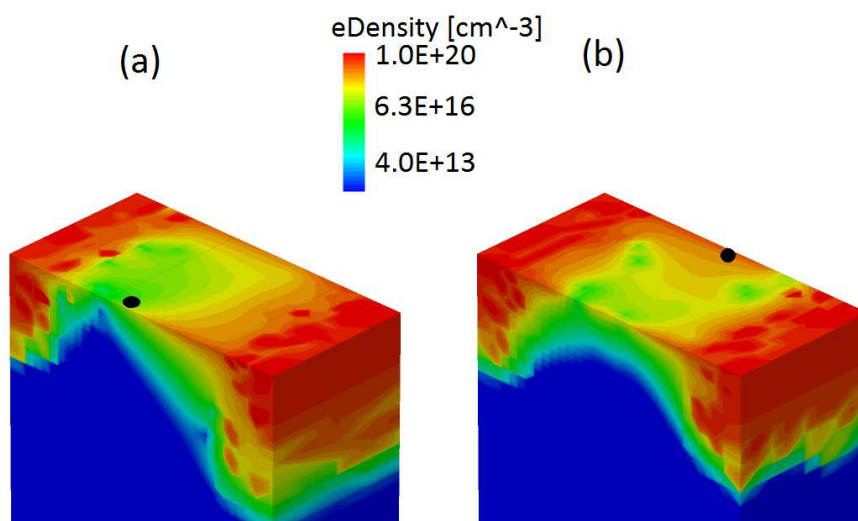


Figura 3.31: concentrazione elettronica simulata (DD) a soglia in dispositivi danti un largo errore nel calcolo dello shift di soglia nel caso di drogaggio disuniforme lungo L (a) ed uniforme (b) per il metodo *sIFM-doping*. La posizione della trappola è evidenziata nell'immagine.

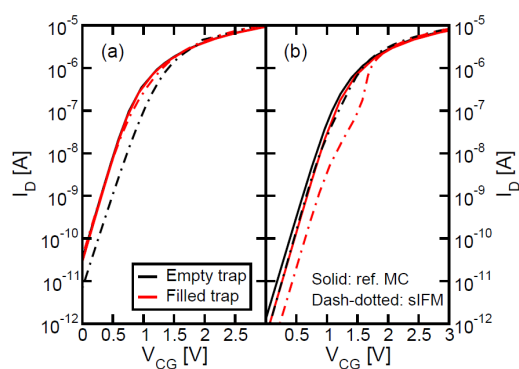


Figura 3.32: transcaratteristiche  $I_d - V_{cg}$  per i dispositivi della Fig.3.31 nel caso di trappola neutra e carica. Confronto tra il metodo di ricostruzione *sIFM-doping* e il riferimento *Monte Carlo* (DD)

la stessa situazione per i droganti, ma la trappola stavolta è ubicata nel lato inferiore anch'essa.

Nel primo di questi due esempi (caso b) ), la corrente a trappola scarica scorre tutta nello spigolo superiore del dispositivo, e posizionando la trappola lungo quel bordo ottengo *Giant RTN*. È facile capire come in questo caso estremo la linearizzazione risulta una pessima approssimazione, in particolare accade come lo *sIFM-doping* sottostimi di molto la corrente nel caso di trappola carica ricostruendo una transcaratteristica a forma di "S" (Fig.3.32b) ), ovvero la cui derivata prima non è monotona, il che è assolutamente non fisico.

Osserviamo ora il più interessante caso a). La situazione a trappola scarica è analoga prima, ma stavolta la trappola è ubicata nel lato basso del dispositivo. Poiché la corrente scorre pochissimo nel lato inferiore, intrappolando un elettrone in quella regione si induce uno shift di soglia molto piccolo. Ricostruendo le transcaratteristiche con il metodo *sIFM-doping* calcoliamo però una  $\Delta V_T < 0$ , il che è assolutamente non fisico. In particolare il metodo ricostruisce molto male il caso a trappola neutra, sovrastimando di molto la soglia  $V_T^0$ . Per capire il perché di ciò dobbiamo fare riferimento alle equazioni su cui si basa il metodo 2.7-2.8 applicate al dispositivo in questione, la cui funzione di Green e profilo atomistico sono raffigurati in Fig.3.33. Dopo aver simulato DD il dispositivo, ci si calcola la deviazione di corrente  $\delta I_d$  indotta dal drogaggio atomistico come l'integrale di  $\delta N$  ( $N_{atomistico}$  di Fig.3.33a)  $-N_{ref}$ ) moltiplicata per la  $G_d$  ( $G_d^0(r)$  o  $G_d^-(r)$  di Fig.3.33b) e c). Ora, semplifichiamo il problema schematizzandolo come in Fig.3.34:  $\delta N > 0$  in basso e  $\delta N < 0$  in alto, le funzioni di Green risultano  $< 0$  sul bordo superiore, mentre su quello inferiore  $G_d^0(r) < 0$  e  $G_d^-(r) \approx 0$  (come è giusto in seguito all'intrappolamento in quel punto di un elettrone). Da notare come nella Fig.3.34 ho usato le convenzioni di Sdevice, ovvero considerare il segno della Green Function negativo e lo shift di corrente positivo. Capito ciò, nel caso di trappola carica calcolerò una certa  $\delta I_d > 0$ , mentre nel caso di trappola scarica l'effetto del bordo inferiore  $\delta N G_d^0 < 0$  annulla quello del bordo superiore  $\delta N G_d^0 > 0$  facendo sì che la deviazione di corrente dal dispositivo con doping continuo sia molto bassa  $\delta I_d \approx 0$ . Questo vuol dire che nei casi in cui gli atomi sono casualmente più presenti in un lato rispetto all'altro, l'effetto complessivo calcolato è nullo ai contatti, perché da una parte avrò un innalzamento di soglia locale dovuto alla presenza di molti ioni accettori ( $V_T^h > V_T^{ref}$ ), dall'altra un abbassamento della soglia locale per la mancanza di atomi droganti ( $V_T^l < V_T^{ref}$ )

e l'effetto complessivo approssimato è sostanzialmente nullo rispetto al dispositivo con doping continuo. Questo avviene in quanto facendo l'integrale 2.7 i due effetti si annullano, che vuol dire che la soglia totale calcolata come media delle soglie locali risulta essere quella del riferimento  $V_T^{tot} = \frac{V_T^l + V_T^h}{2} \approx V_T^{ref}$ . In realtà tale approssimazione è falsa. Per capire ciò possiamo vedere il dispositivo di Fig.3.34 come due dispositivi in parallelo, uno con soglia alta  $V_T^h$ , l'altro con soglia bassa  $V_T^l$ . La soglia del dispositivo totale non è la soglia media, ma ha un valore molto più vicino alla  $V_T^l$ . Infatti se il dispositivo con soglia bassa si accende prima, si accende prima anche il dispositivo complessivo. Per questo motivo lo shift di corrente  $\delta I_d$  calcolato dal dispositivo di riferimento risulta fortemente sottostimato. Nel calcolo del bias point  $i$ -esimo per il caso a trappola scarica  $V_g^i = V_{g,ref}^i - \delta I_d^i / y_{dg}^i$  non solo lo shift di corrente risulta essere ben minore, ma accade anche che la transconduttanza  $y_{dg}$  risulta maggiore. Tali due effetti portano ad una netta sovrastima della soglia  $V_T^0$  tanto da indurre un calcolo di  $\Delta V_T$  negativo.

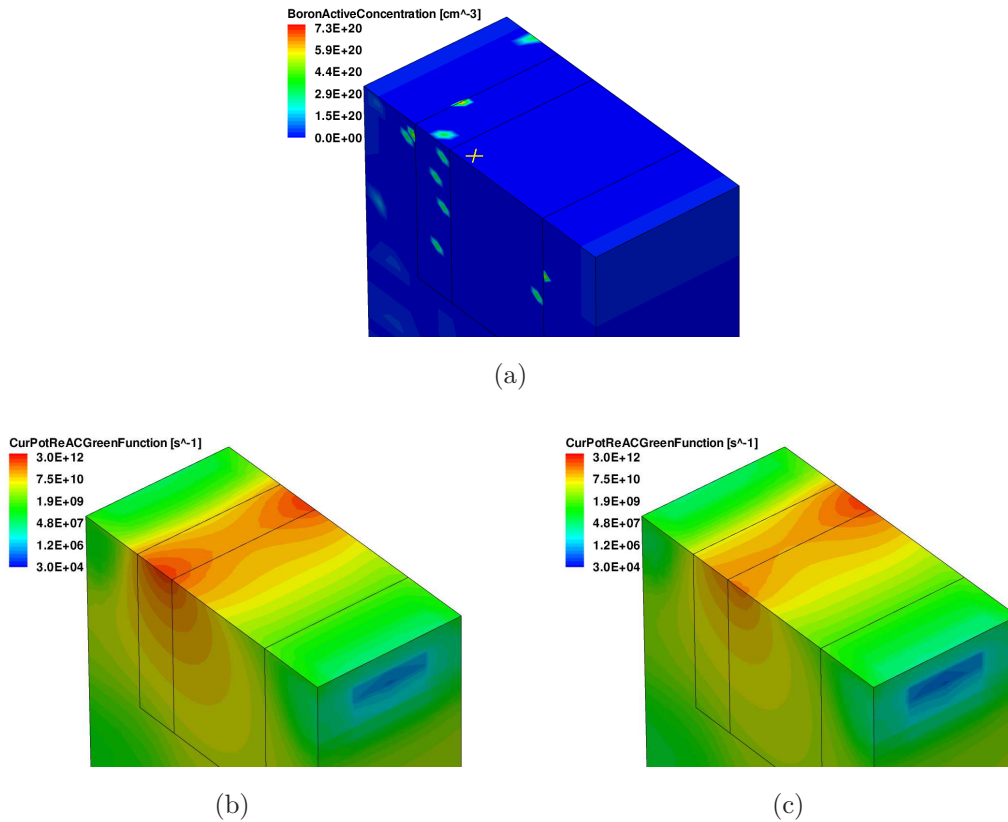


Figura 3.33: a)  $N_A(r)$ , b)  $G_d^0(r)$  e c)  $G_d^-(r)$  per il caso Fig.3.31a)

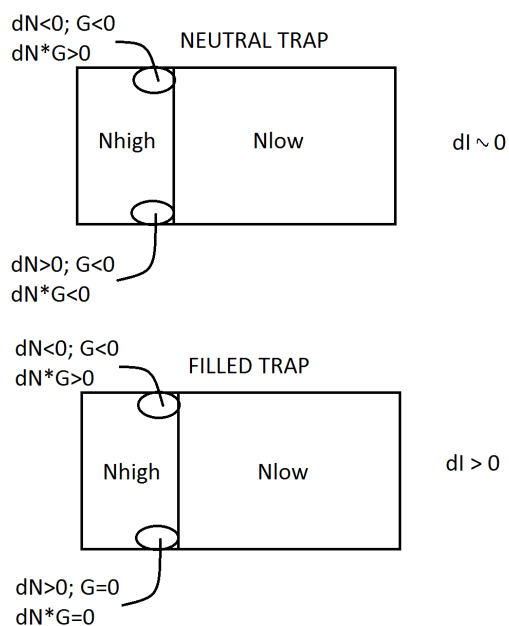


Figura 3.34: descrizione qualitativa dell'errore di linearizzazione per il caso Fig.3.31a)

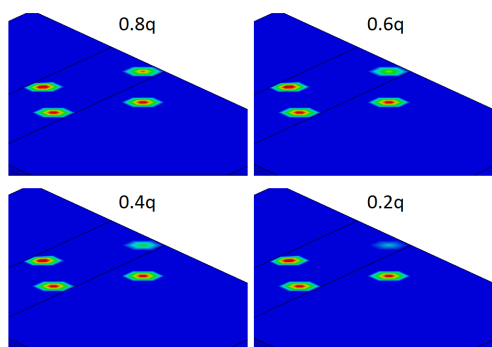
### sIFM-trap: carica frazionaria

Come già detto più volte, il metodo *sIFM-trap* linearizza l'effetto della trappola sulla conduzione ai contatti. Il metodo vede la trappola come “piccola perturbazione” dal dispositivo atomistico di riferimento. Pertanto, in linea teorica, considerando l'elettrone intrappolato con una carica  $Q \rightarrow 0$  il metodo ricostruisce sempre meglio la transcaratteristica, in quanto la perturbazione è minore. In altre parole, considerando la trappola piena come carica frazionaria piccola, per esempio  $dq = 0.2 \times e$ , dovremmo trovare un errore inferiore nel calcolo dello shift di soglia. Questo è ciò che accade, come mostrato in Fig.3.35. In relazione al dispositivo con drogaggio disuniforme lungo la lunghezza del canale, ho focalizzato l'attenzione sulla coppia doping-trap più critica, in cui l'errore nel calcolo della  $\Delta V_T$  è maggiore ( $\Delta V_T^{sIFM-trap} = 456.72mV$  vs  $\Delta V_T^{DD} = 287.82mV$ ). Si vede bene come anche nel caso più critico troviamo che l'errore di linearizzazione diminuisce al diminuire della carica:

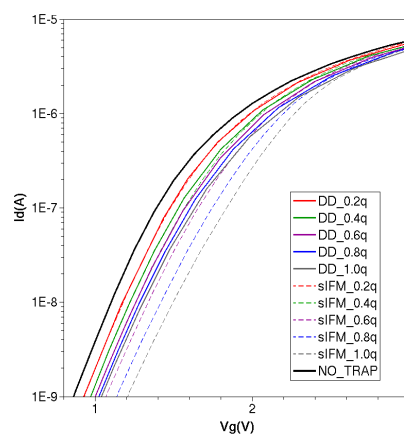


$$\left| \Delta V_T^{sIFM-trap} - \Delta V_T^{DD} \right| \xrightarrow{dq \rightarrow 0} 0 \quad (3.2)$$

Si vede bene, dalla Fig.3.35(b), come passando da una carica frazionaria all'altra, il metodo *sIFM-trap*, essendo lineare, ricostruisce la transcaratteristica allontanandosi dalla curva a trappola scarica in maniera proporzionale alla carica inserita. Per esempio, passando da  $0.2q$  a  $0.4q$  lo shift di soglia semplicemente raddoppia. Simulando DD i dispositivi a carica frazionaria, vediamo che non è così! A  $0.2q$  le transcaratteristiche dei due metodi coincidono, poi man mano che inseriamo carica l'allontanamento dal caso "no trap" diventa sub-lineare per il DD; ciò porta il metodo *sIFM-trap* a sovrastimare lo shift di soglia. Faccio notare che questo è un caso molto sfortunato (e proprio per questo lo analizziamo), nella realtà la maggior parte delle trappole subiscono shift di soglia proporzionali alla carica inserita, almeno finché la conduzione è sostanzialmente uniforme (non è il caso in questione poiché siamo sul bordo, come visibile dalla Fig.3.35(a)).



(a) carica frazionaria inserita come doping frazionario



(b) transcaratteristiche risultanti per il metodo DD-MC e lo *sIFM-trap*

Figura 3.35: confronto tra *sIFM-trap* e MC-DD con carica frazionaria per una particolare coppia configurazione doping - trappola del caso di drogaggio disuniforme lungo L

Vorrei ora spendere due parole sulla costruzione di un progetto del genere. Sdevice non permette di inserire una trappola la cui carica non è unitaria. Per far ciò ho usato i tool Snmesh e Tdx. Quest'ultimo permette di accedere ai dataset di concentrazione del drogaggio nodo per nodo, ma non permette, purtroppo, di accedere alle coordinate del nodo stesso, che è pertanto semplicemente indicizzato per regione. Inoltre, anche se conoscessi l'indice esatto non saprei che concentrazione di carica mettere nel nodo, in quanto non so quale sia il Voronoi Box corrispondente. Per superare questi ostacoli, procedo nella seguente maniera: inserisco la carica nel dispositivo come atomo drogante ionizzato attraverso il tool Snmesh dandogli le coordinate della trappola, dopodiché carico con Tdx il dataset "BoronActiveConcentration" sia per il dispositivo con trappola scarica sia per quello con trappola carica, e creo un nuovo file .tdr in cui nel nodo nel quale trovo la differenza di concentrazione moltiplico il valore per 0.2 (per esempio). Fatto ciò, purtroppo, abbiamo altri due problemi. Innanzitutto Snmesh all'interfaccia di due regioni inserisce il drogante in ambo le parte, raddoppiando di fatto la carica inserita. Per tal motivo un atomo drogante verrà posizionato anche nell'ossido essendo che siamo all'interfaccia, ma ciò non è un problema in quanto Sdevice lo ignora. È invece problematico quando la carica sta tra due regioni nel silicio, come per esempio accade per il caso di drogaggio disuniforme lungo L per le regioni "DDR\_high" e "DDR\_low": in quel caso dovrò dimezzare la carica inserita. L'altro, e più importante, problema di adottare tal procedimento, sta nel fatto che per alcuni nodi, i quali sono difficili da individuare a priori, Snmesh considera un Voronoi Box differente da quello che fa Sdevice, dunque non mi inserisce esattamente la carica di 1 atomo, ma un valore diverso e variabile da nodo a nodo. Questo è soprattutto vero (ma non solo!) vicino ai bordi del dispositivo, dove sono presenti nodi aggiuntivi a pochissimi angstrom di profondità. Per evitare errori grossolani l'unico modo è quello di confrontare la concentrazione che Snmesh e Sdevice inseriscono nodo per nodo all'interfaccia e considerare semplicemente quella di Sdevice (dove si trova la differenza) inserendola con Tdx.

Ho affermato come simulando il dispositivo a "trappola frazionaria" ci aspettiamo un errore molto minore rispetto al MC-DD. Ciò l'abbiamo visto per un caso piuttosto estremo e critico (Fig.3.35(a)), ma è generalizzabile per qualsiasi coppia doping-trap e per qualsiasi dispositivo? La risposta è banalmente sì, in quanto il concetto che ci sta dietro non cambia, ovvero il fatto che lo *sIFM-trap* linearizza l'effetto della trappola sulla conduzione ai contatti. Per tal motivo ho

simulato una gran quantità di casi in maniera DD con carica frazionaria  $0.2q$  per diversi tipi di drogaggio e ho confrontato tali risultati con quelli ottenuti dallo *sIFM-trap*, i cui shift di soglia sono quelli trovati a trappola “unitaria” ridotti di un fattore 5. Ciò che ottengo è visibile dalle figure successive: Fig.3.36(a) per il drogaggio uniforme, Fig.3.36(b) per il drogaggio disuniforme lungo L, Fig.3.36(c) per il drogaggio disuniforme lungo W e Fig.3.36(d) per il drogaggio retrogrado con svuotamento a  $t_{epi} = 16nm$ . Le correlazioni tra MC e *sIFM-trap* sono quasi unitarie e le tendenze giacenti sulla bisettrice. Da notare come non ho simulato tutti i casi che avevo a disposizione per il conflitto Sdevice-Snmesb di cui ho discusso precedentemente. A parte per il caso Fig.3.36(b) dove le cose funzionano piuttosto bene, per gli altri casi i nodi più critici per tal problematica sono soprattutto, come già detto, vicino al bordo. Pertanto negli scatterplots non troveremo i casi “peggiori” per le Fig.3.36(a) e 3.36(d) e i “migliori” per la Fig.3.36(c).

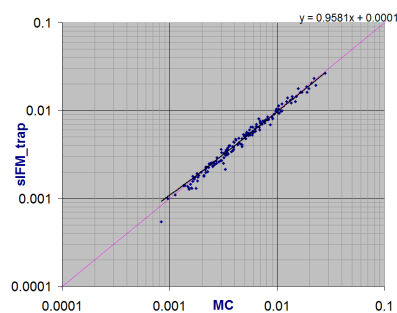
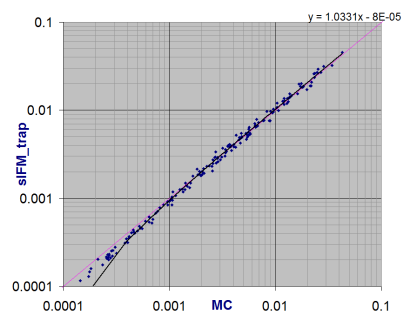
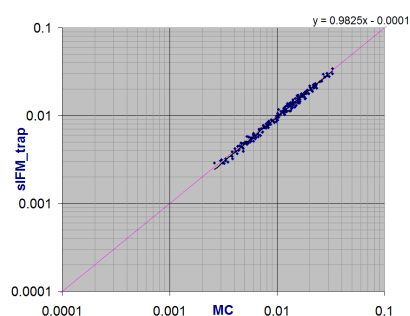
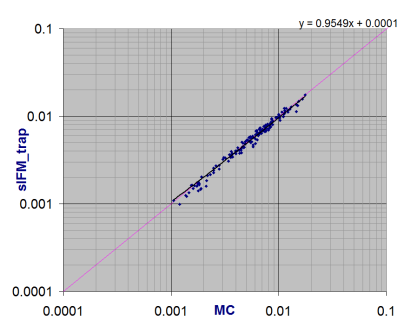
(a) drogaggio uniforme  $\rho = 0.991$ (b) drogaggio disuniforme lungo L  
 $\rho = 0.995$ (c) drogaggio disuniforme lungo W  
 $\rho = 0.991$ (d) drogaggio retrogrado ( $t_{epi} = 16nm$ )  $\rho = 0.989$ 

Figura 3.36: scatterplots *sIFM\_trap*-*MC* nel caso di trappola a carica frazionaria ( $0.2q$ ) per i diversi dispositivi

### sIFM-trap: effetti non lineari della trappola

Relativamente al caso uniforme e trappola “intera”, le figure 3.37(a) e Fig.3.37(b) mostrano l’errore (assoluto e relativo) commesso dal metodo *sIFM-trap* a metà canale al variare della posizione della trappola lungo la larghezza del dispositivo per tutti i 568 casi simulati (Fig.3.41(a)). Come vediamo, non solo l’errore assoluto ( $\Delta V_T^{DD} - \Delta V_T^{sIFM-trap}$ ) risulta essere più elevato ai bordi, ovvero dove la linearità viene meno, ma anche quello relativo. Ci aspettiamo, dunque, come potevamo immaginare, che i casi più critici da approssimare siano proprio quelli relativi a trappole ubicate vicino agli spigoli della cella.

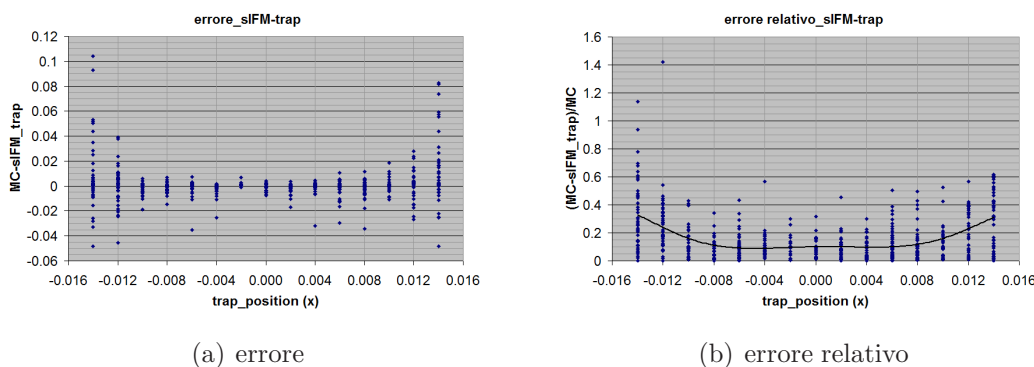


Figura 3.37: errori commessi dallo *sIFM-trap* nel calcolo dei  $\Delta V_T$  per il drogaggio uniforme per tutti i 568 casi della Fig.3.8(b) al variare della posizione della trappola lungo la  $W$  del dispositivo con  $y$  fissata a metà canale ( $= 32nm$ )

Analizziamo ora due configurazioni di drogaggio particolari delle 100 simulate: il “Device 1” (Fig.3.38(a)) ed il “Device 9” (Fig.3.38(b)). Nel dispositivo 9 la densità di corrente scorre in entrambi i bordi, mentre nel device 1 solo sul bordo inferiore per la presenza nel lato opposto di ioni accettori che bloccano la conduzione in quella zona. Faccio notare come per quest’ultimo dispositivo la corrente tende a scorrere in un percorso conduttivo largo circa  $2nm$ , come si vede in Fig.3.38(b), e per  $y \approx 42nm$  tal cammino si stringe ( $\approx 1nm$ ) perché a  $1.37nm$  di profondità e  $2nm$  di distanza dallo spigolo è presente uno ione accettore che blocca la conduzione più all’interno della cella schiacciandola di fatto sul bordo.

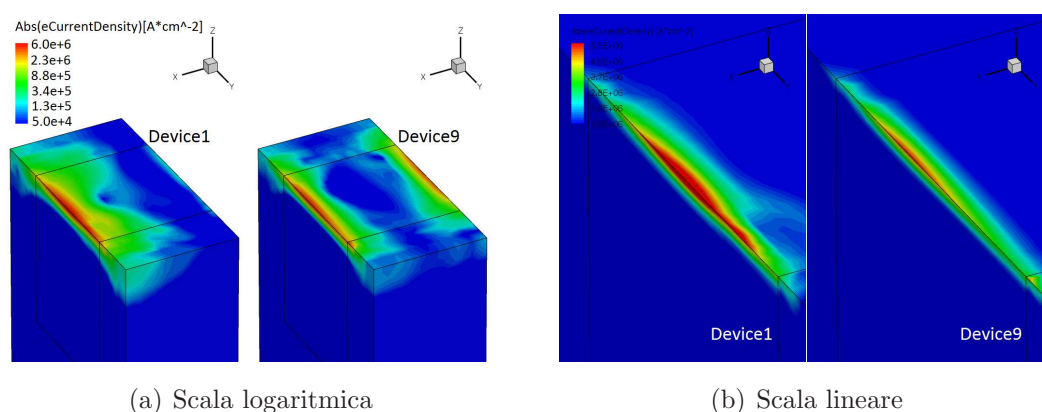


Figura 3.38: Densità di corrente a soglia per i dispositivi “1” e “3” a trappola scarica

Le Fig.3.39(a) e 3.39(b) mostrano bene come tali dispositivi sono in linea esattamente con la tendenza vista per l’immagine precedente, ovvero che il metodo *sIFM-trap* approssima bene lo shift di soglia finché la trappola è all’interno del canale, quando invece sul bordo ho una forte conduzione e la trappola è posizionata in quella zona gli effetti non lineari si sentono ed il metodo *sIFM* approssima male la  $\Delta V_T$ .

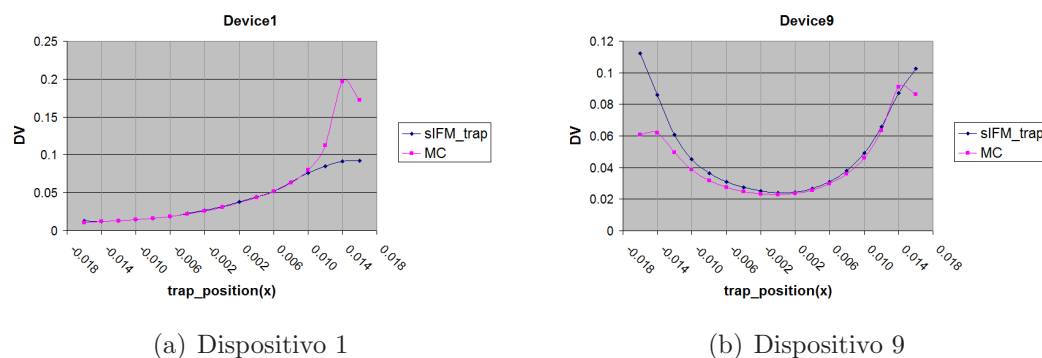


Figura 3.39:  $\Delta V_T$  calcolati da *sIFM-trap* e dal MC al variare della posizione della trappola lungo x a metà canale (Fig.3.41(a))

Ora, andiamo a vedere cosa succede se posizioniamo la trappola sul bordo inferiore (quello “critico” per la linearità) a passo di  $2nm$  per i due dispositivi (Fig.3.41(b)). I risultati ottenuti sono visibili in Fig.3.40(a) e 3.40(b).

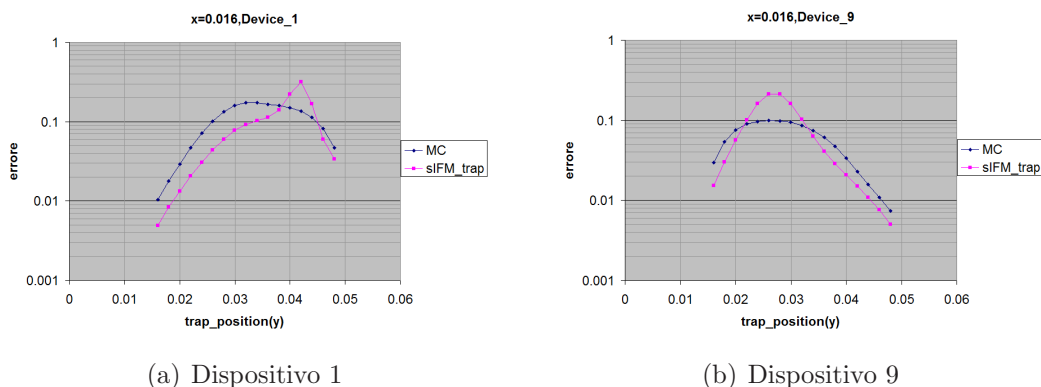


Figura 3.40:  $\Delta V_T$  calcolati da *sIFM-trap* e dal MC al variare della posizione della trappola lungo  $y$  sul bordo inferiore (Fig.3.41(b))

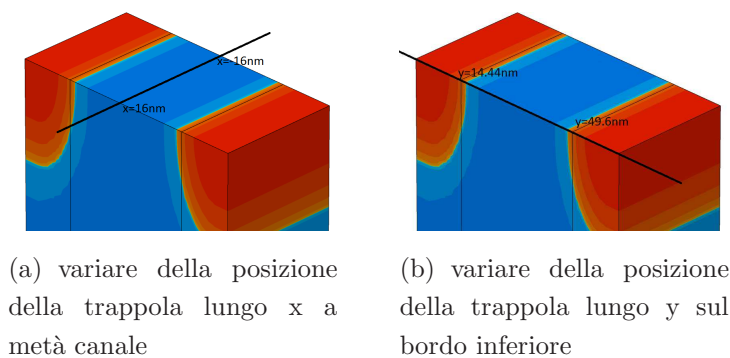


Figura 3.41: spiegazione figure precedenti

Osservando tali figure vediamo come, focalizzando l'attenzione sul Device 1, lo *sIFM-trap* sottostimi lo shift di soglia quando il percorso conduttivo è abbastanza largo, quando invece si stringe la  $\Delta V_T^{sIFM-trap}$  cresce vertiginosamente fino a sovrastimare quella calcolata DD. Perché avviene ciò? Analizziamo la situazione in due punti con riferimento alla Fig.3.42:  $y = 26nm$  (centro del canale “largo”) ed  $y = 42nm$  (centro del canale “stretto”). Come abbiamo visto nel paragrafo precedente, lo *sIFM-trap* approssima perfettamente il MC-DD nel caso di trappola frazionaria  $dq = 0.2 \times e$ . Nel caso la trappola sia carica  $0.2e$  e sia posizionata ad  $y = 26nm$  ottengo una certa  $\Delta V_T^{sIFM-trap} \approx \Delta V_T^{DD}$ . In questo caso, essendo che la carica è piuttosto piccola ed il cammino conduttivo largo, il canale non si è chiuso, non ho per intenderci *Giant RTN* in cui si chiude un filamento di corrente, ed infatti prevediamo che la corrente ci scorra attorno a tale carica,

facendo sì diminuire la  $I_d$  totale ma non eccessivamente. Passando da  $0.2e$  ad  $1e$  lo *sIFM-trap*, essendo lineare, moltiplica semplicemente per un fattore 5 lo shift di soglia, mentre nel DD la linearità non c'è. Quello che succede, infatti, è che aumentando la carica frazionaria il canale tende a chiudersi causando una diminuzione di corrente sempre più grande di quella prevista dal metodo sIFM, pertanto ci aspettiamo che lo shift di soglia sia sovra-lineare con la carica, ovvero che la funzione  $\Delta V_T(dq)$  abbia derivata seconda positiva  $\Delta V_T''(dq) > 0$  (anziché nulla!). Ad  $1e$  il cammino conduttivo è chiuso totalmente causando *Giant RTN*, che fa sì che lo *sIFM-trap* sottostimi la  $\Delta V_T$ .

Ad  $y = 42nm$  ci aspettiamo accada la situazione opposta. Essendo che il cammino conduttivo della corrente è molto stretto, posizionando anche una carica molto piccola ( $0.2q$ ) il canale è sostanzialmente chiuso, causando uno shift di soglia molto elevato  $\Delta V_T^{sIFM-trap} \approx \Delta V_T^{DD}$ . In altre parole: allo ione a  $2nm$  di distanza dallo spigolo si somma l'effetto della carica posizionata sul bordo chiudendo praticamente il canale già quando tal carica è ben più piccola dell'unità. Passando ora da  $0.2e$  ad  $1e$ , lo *sIFM-trap* moltiplica sempre per 5 lo shift di soglia, mentre nel DD ci sarà solo un fattore  $\approx 2$  per esempio, in quanto aggiungendo carica ottengo sì una diminuzione della corrente totale, ma inferiore di quella calcolata dallo sIFM, poiché il canale è già chiuso praticamente a  $0.2e$  ed aumentando la carica ottengo un effetto sempre minore al contatto di drain. Pertanto quello che prevediamo è una funzione  $\Delta V_T(dq)$  sub-lineare, ovvero con  $\Delta V_T''(dq) < 0$ .

Tutti questi ragionamenti sono visualizzabili dalla schematizzazione della Fig.3.42.

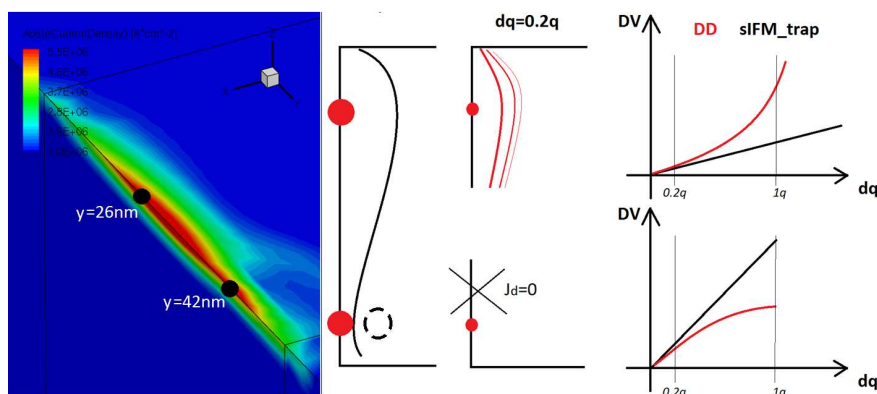


Figura 3.42: schematizzazione degli effetti non lineari della trappola ad  $y = 26nm$  ed  $y = 42nm$  per il Device 1

Per quanto riguarda invece il dispositivo 9, abbiamo la situazione duale. Ad  $y = 26nm$  troviamo uno ione accettore a  $8nm$  dal bordo ed  $1.37nm$  di profondità, che fa sì che la densità di corrente sia piccata e piuttosto localizzata sullo spigolo a quella  $x$ . Ad  $y = 42nm$ , invece, la corrente è più bassa ma comunque scorre in un filamento conduttivo piuttosto largo. Ci aspettiamo pertanto uno shift di soglia con derivata seconda negativa posizionando la trappola ad  $y = 26nm$  e positiva ad  $y = 42nm$  (seppur lievemente).

Andiamo ora a verificare questi ragionamenti simulando DD i dispositivi a trappole “frazionarie” ed andando a confrontare i risultati con quelli approssimati del metodo *sIFM-trap*. Qui di seguito ho riportato le densità di corrente alla tensione di  $1V$  ( $\approx V_T^0$  per entrambi i dispositivi) per i dispositivi 1 (Fig.3.43) e 9 (3.44). Vediamo esattamente quello che ci aspettavamo per il dispositivo 1 a carica  $q = 0.2e$ : se posizioniamo tale trappola ad  $y = 26nm$  il canale conduttivo non si è chiuso e la corrente scorre attorno alla carica perché non trova “ulteriori ostacoli”, mentre se la posizioniamo ad  $y = 42nm$  accade che la corrente, trovando anche lo ione accettore vicino, deve scorrere molto più internamente, inducendo una forte diminuzione di conduttività e quindi la quasi totale chiusura del canale. Dopodiché passando da  $0.2e$  ad  $1e$  il filamento conduttivo si chiude per entrambi i casi provocando uno shift di soglia pertanto analogo. Discorso contrario, come ci aspettavamo, per il dispositivo 9. Per esso notiamo come posizionando una carica  $0.2q$  in  $y = 26nm$  la corrente cerca di scorrere tra i due ostacoli trappola-drogante ma comunque ha subito una fortissima riduzione, ovvero la quasi totale chiusura del canale conduttivo; discorso diverso per il caso di trappola posizionata ad  $y = 0.42nm$ : il canale è largo e la tendenza è pertanto quella di avere una  $\Delta V_T''(dq) > 0$ , ma comunque la densità di corrente è molto bassa, quindi prevediamo che la curva sarà di poco sovra-lineare, ovvero che il canale si dovrebbe chiudere effettivamente a  $q \gg 1e$ .

Tali osservazioni trovano conferma guardando i risultati di tali simulazioni, ovvero la Fig.3.45. In vero passando da  $0.2q$  a  $1q$  per il dispositivo 1 si ha un incremento dello shift di soglia di 7.374 caso  $y = 26nm$  (ovvero  $\Delta V_T(dq)$  sovra-lineare) e di 2.654 caso  $y = 42nm$  (ovvero  $\Delta V_T(dq)$  sotto-lineare) e non di 5 come lo *sIFM-trap* fa. Al contrario per il dispositivo 9 troviamo un incremento dello shift di soglia di un fattore 2.200 caso  $y = 26nm$  e di 6.652 caso  $y = 42nm$ .



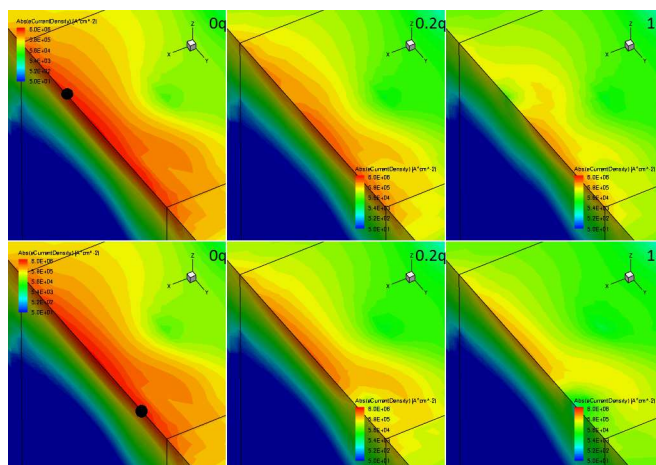


Figura 3.43: Densità di corrente ad 1V per il Device 1 ( $V_T^0 = 1V$ ) a trappola nulla, frazionaria  $0.2q$  e unitaria nei due casi studiati ( $y = 26\text{nm}$  ed  $y = 42\text{nm}$ )

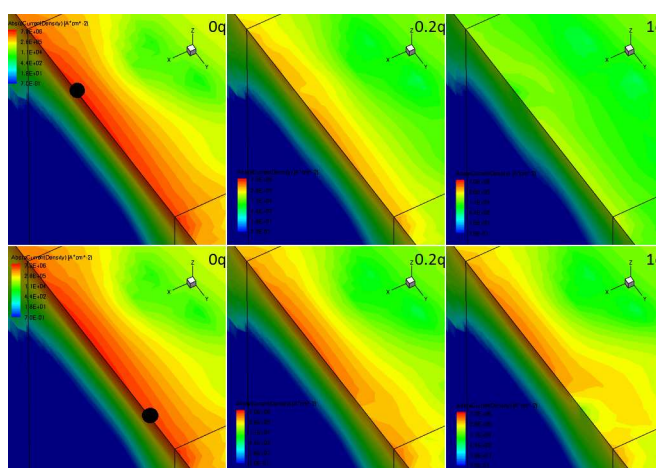


Figura 3.44: Densità di corrente ad 1V per il Device 9 ( $V_T^0 = 1V$ ) a trappola nulla, frazionaria  $0.2q$  e unitaria nei due casi studiati ( $y = 26\text{nm}$  ed  $y = 42\text{nm}$ )

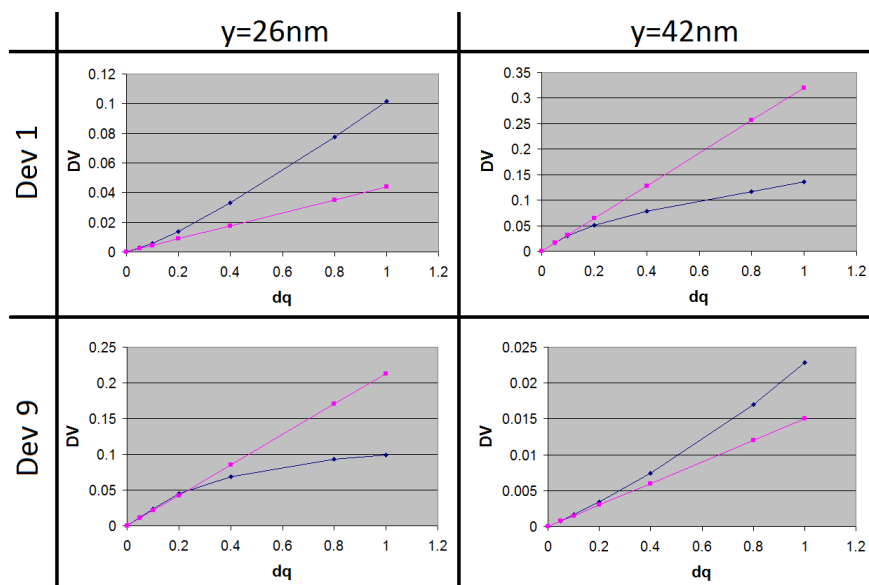
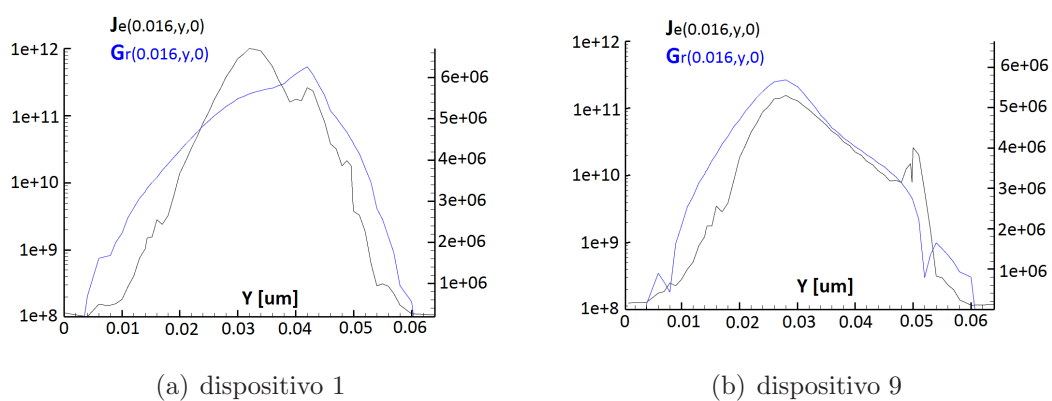


Figura 3.45:  $\Delta V_T^{DD}$  e  $\Delta V_T^{sIFM-trap}$  per i dispositivi 1 e 9 per trappole posizionate alle coordinate  $(16, 26, 0)nm$  e  $(16, 42, 0)nm$  alle varie cariche frazionarie



(a) dispositivo 1

(b) dispositivo 9

Figura 3.46: funzioni di Green e densità di corrente a trappola scarica a soglia ( $\approx 1V$ ) lungo lo spigolo basso del dispositivo

A confermare i ragionamenti appena fatti sono i grafici relativi alle funzioni di Green valutate sullo spigolo inferiore a soglia Fig.3.46(a) e 3.46(b). Da come Mueller e Shulz teorizzavano, ci aspettiamo una funzione di Green (ovvero una sensibilità ai droganti/trappole) in qualche modo proporzionale alla densità di corrente stessa. Questo infatti accade più o meno per il dispositivo 9. Per il dispositivo 1, invece, abbiamo una funzione di Green piccata non dove la densità di corrente è massima, bensì dove il percorso conduttivo si stringe, a causa dell'errore di linearizzazione discusso in questo paragrafo. Faccio inoltre notare come le funzioni di Green lungo  $y$  sono della stessa forma delle  $\Delta V_T^{sIFM-trap}(y)$  corrispondenti viste precedentemente in quanto proporzionali.

La figura 3.45 è abbastanza eloquente: simulando DD il dispositivo, in vero, avremo una certa funzione  $\Delta V_T(dq)$ . Il metodo *sIFM-trap* è l'approssimazione al prim'ordine di tal curva, ovvero la tangente all'origine, e non conta gli effetti di ordine superiore, quelli non lineari. Per lo *sIFM-trap*, infatti, la  $\Delta V_T''(dq) = 0$  mentre nella realtà abbiamo visto come essa possa essere negativa o positiva a seconda di quanto sia piccata la densità di corrente e di come la trappola interagisca con gli atomi droganti ionizzati. A posteriori questa analisi si può fare ed è sempre abbastanza intuitiva, ma a priori risulta davvero complicato stimare l'errore di linearizzazione sullo shift di soglia. Un modo molto semplice potrebbe essere quello di simulare DD il dispositivo a trappola riempita con carica frazionaria a  $0.2e$  e  $0.4e$ , in modo da capire almeno il segno della derivata seconda (dai tre punti  $0 - 0.2 - 0.4$ ). Ciò, però, risulta molto sconveniente, in quanto con solo una simulazione (invece che due) posso direttamente calcolare lo shift di soglia a trappola "unitaria" con errore nullo (e non fare solo una stima). Quello che invece si può fare, vedendo le Fig.3.40(a) e 3.40(b), è una sorta di filtro passa basso (o media mobile) dei valori trovati sugli spigoli, in quanto in quelle zone lo *sIFM-trap* ha escursioni a bassi-alti valori più piccate del MC, ma "in media" risulta una buona approssimazione.

### 3.2.5 Verifiche sulla metodologia

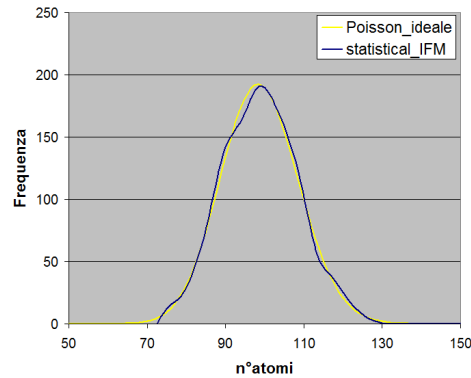
In questa breve sezione mostro alcune necessarie verifiche fatte per accertare che le metodologie discusse ed elaborate finora siano valide. In particolare faccio vedere come il metodo *statistical-IFM* produca effettivamente dei profili atomistici di drogaggio in linea con quello ideale-continuo; mostro poi come le distribuzioni di probabilità cumulativa costruite DD non dipendano in modo eccessivo dalla mesh e che siano in perfetta sintonia con il MC costruito dai dispositivi randomizzati dal metodo; infine mostro come la statistica  $(1 + N_T)N_D$  riproduca la stessa cumulativa costruita dalla statistica  $2N_TN_D$ , cosa che ho dato finora per scontato.

#### Numero dei droganti

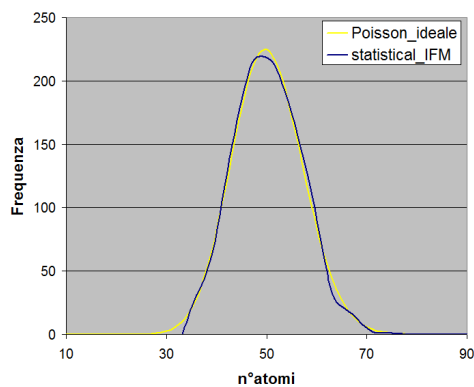
La prima verifica da fare è quella relativa al fatto che il metodo *sIFM* costruisca effettivamente i profili “giusti”. Per tal motivo ho voluto accertare che nelle regioni “di canale” il metodo inserisca effettivamente il corretto numero di atomi droganti accettori, ovvero seguente la statistica di Poisson.

Per far ciò basta calcolare per ciascun dispositivo randomizzato dal metodo l'integrale di volume della concentrazione  $N_A(r)$  nella regione del canale, che mi dà semplicemente il numero totale di ioni accettori presenti in tale zona. Dopodiché da tutti questi campioni verifico che la media sia in linea con quella del dispositivo continuo e la varianza sia uguale alla media. Le distribuzioni di probabilità relative ai dispositivi con drogaggio uniforme e disuniforme lungo L sono riportate nelle Fig.3.47(a), 3.47(b) e 3.47(c). Questi dati sono stati ottenuti contando gli atomi per 1000 dispositivi randomizzati dal metodo. Si vede bene come le distribuzioni, le medie e le varianze risultanti siano in buon accordo con quelle relative alle distribuzioni di Poisson ideali. I dispositivi utilizzati per le statistiche viste in precedenza, però, sono  $N_{doping} = 100$  e non 1000. L'accordo con la Poisson ideale rimane buono anche in questo caso e ciò si conferma anche per gli altri tipi di dispositivi analizzati. Evito di riportare tali risultati in quanto ripetitivi concettualmente.

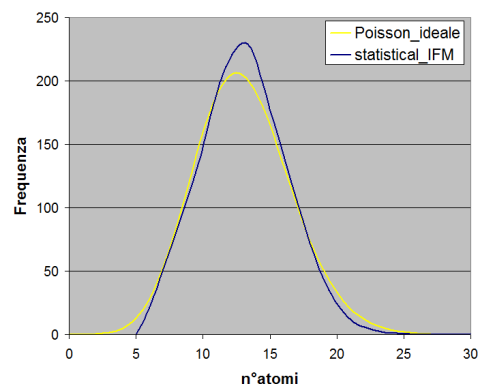
Vorrei ora spendere due parole sulla procedura fatta per arrivare ai risultati delle figure appena analizzate. In vero, il calcolo degli integrali di volume per così



(a) drogaggio uniforme.  $\bar{N}_{ideal} = \sigma_{ideal}^2 = 99.12$ ;  $\bar{N}_{sIFM} = 98.86$  e  $\sigma_{sIFM}^2 = 95.51$



(b) drogaggio disuniforme lungo L, zona alto drogata.  $\bar{N}_{ideal} = \sigma_{ideal}^2 = 50.69$ ;  $\bar{N}_{sIFM} = 50.24$  e  $\sigma_{sIFM}^2 = 45.97$



(c) drogaggio disuniforme lungo L, zona basso drogata.  $\bar{N}_{ideal} = \sigma_{ideal}^2 = 13.07$ ;  $\bar{N}_{sIFM} = 13.01$  e  $\sigma_{sIFM}^2 = 10.11$

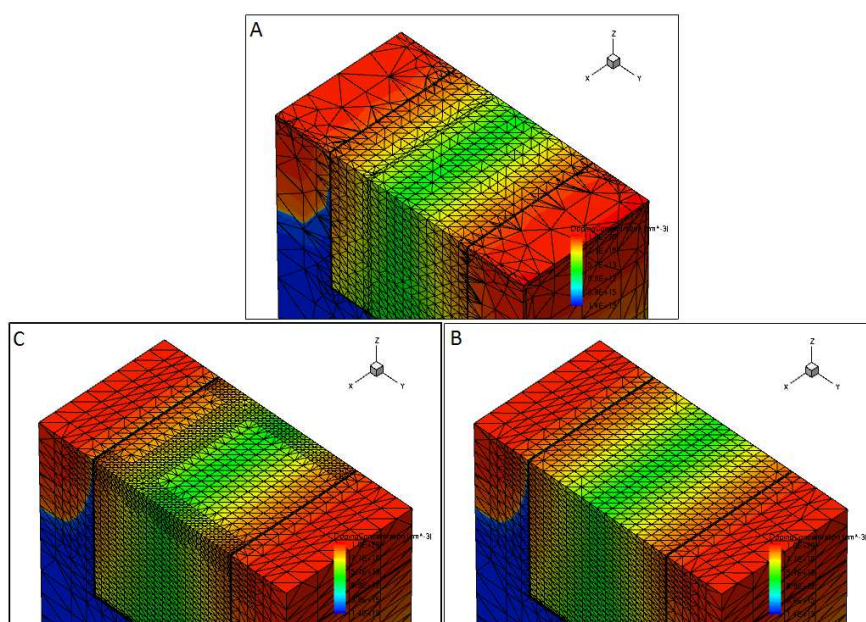
Figura 3.47: distribuzioni di probabilità degli atomi presenti nella regione svuotata costruite su 1000 dispositivi randomizzati nel caso di drogaggio uniforme e disuniforme lungo L

tanti dispositivi necessita di essere svolto in maniera automatica attraverso una macro. Il tool della Synopsys *Tecplot* permette di visualizzare i dataset  $N_{A,\nu}(r)$  e di calcolarci l'integrale per regione. Tale tool ha poi la possibilità di scrivere, nel proprio linguaggio, una macro che permette di svolgere una qualsiasi operazione in tale programma attivandola tramite linea di comando. Pertanto ho aperto il file .tdr contenente i 1000 dataset e ho registrato una macro nella quale calcolavo  $\int N_{A,\nu}(r)d^3r$  per la prima configurazione di drogaggio. Dopodiché ho cambiato lo script corrispondente attivando un ciclo e indicizzando il dataset, ovvero il profilo di doping ( $= \nu$ ) che va da 1 a 1000. Attivando tale macro ho fatto calcolare tutti questi integrali in un file di testo, che, dopo averlo modificato, ho opportunamente letto con Inspect, il quale mi ha permesso di stampare tali numeri come variabili swb.

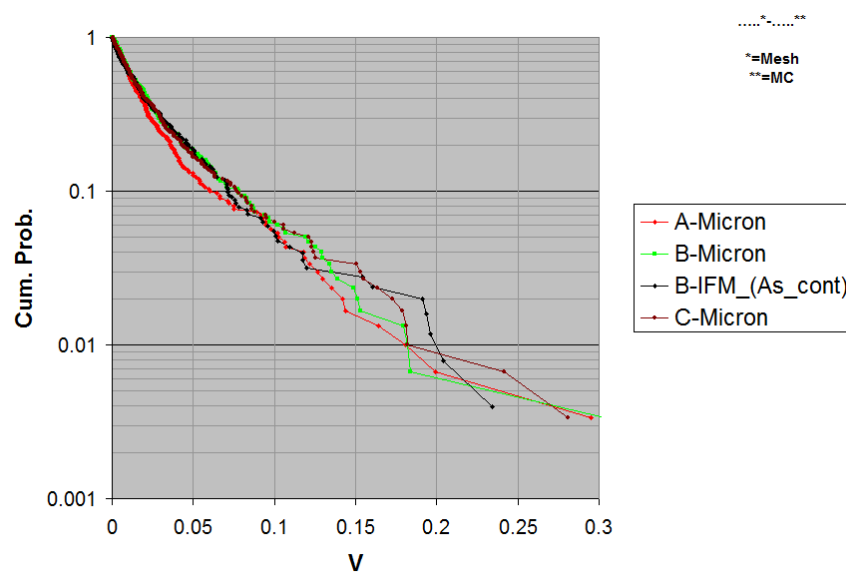
### Dipendenze delle cumulative

Nel paragrafo precedente ho mostrato come i dispositivi randomizzati dal metodo *sIFM* abbiano un numero di atomi per regione concorde con la statistica di Poisson. Ciò ovviamente non basta, infatti, come ho più volte sottolineato nel primo capitolo, quello che conta di più per le dispersioni statistiche non è il numero degli atomi, bensì la loro posizione spaziale. Questo è più difficile da verificare, anche se più facile da implementare. Pertanto ho costruito cumulative randomizzando dispositivi con il metodo MC costruito in azienda e le ho confrontate con quelle relative al MC basato su dispositivi randomizzati da Synopsys. Un'altra dipendenza può essere dovuta alla mesh adottata sul dispositivo, pertanto ho costruito per le stesse coppie doping-trap cumulate da diverse mesh. I risultati relativi a tali confronti sono visualizzabili nelle Fig.3.48(a), 3.48(b) e 3.49 per il caso di drogaggio disuniforme lungo la lunghezza del canale.

Il primo confronto è certamente da fare tra le cumulative costruite da diverse mesh (stesse coppie doping-trap). Un po' di dipendenza c'è (Fig.3.48(b) e 3.49), ma se raffiniamo abbastanza la griglia le cumulative sono sovrapposte (B e C in Fig.3.48(b)). Voglio far notare come per la mesh C ho raffinato ulteriormente le zone più "critiche", ovvero i bordi del dispositivo e la regione di transizione alto-basso del drogaggio. La curva nera di Fig.3.48(b) è stata ottenuta adottando



(a) diverse mesh adottate: (A) lasca, (B) fine, (C) molto fine



(b) cumulative risultanti dalle mesh A, B e C per le stesse 300 coppie doping-trap randomizzate con il MC “aziendale”. Esse vengono confrontate anche con il MC costruito su dispositivi randomizzati da IFM in cui ho imposto l’Arsenico continuo per il drogaggio donore

Figura 3.48: distribuzioni di probabilità cumulativa costruite variando la mesh ed il metodo di randomizzazione dei dispositivi per il caso di drogaggio disuniforme lungo L

la mesh B (Fig.3.48(a)) e i profili randomizzati dal metodo IFM. Per essi ho poi modificato tutti i profili donore imponendoli continui. Dalla figura si osserva un ottimo accordo tra la curva nera e quella verde, in cui cambia solo il metodo di randomizzazione dei dispositivi e non la mesh.

Con questi semplici risultati ho mostrato come il metodo IFM randomizza in modo corretto i dispositivi e il MC costruito su di essi risulta dare risultati analoghi a quelli relativi al MC “aziendale”. Ho mostrato poi come per le stesse coppie doping-trap ci sia una leggera dipendenza dalla mesh, che però svanisce qualora si adottassero mesh sufficientemente fini.

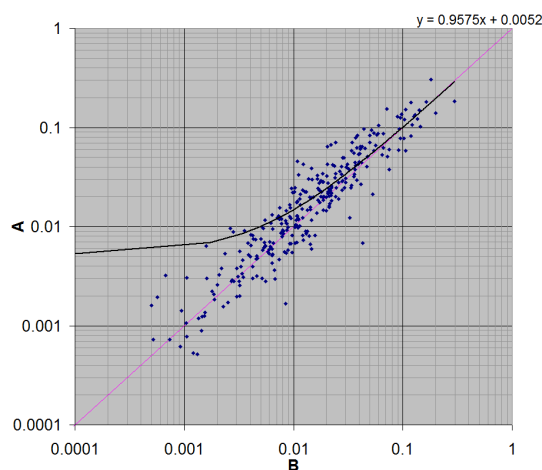


Figura 3.49: scatterplot dei 300 campioni ottenuti con le mesh A e B:  $\rho = 0.888$

### Statistica $(1 + N_T)N_D$ vs Statistica $2N_TN_D$

I metodi *sIFM-doping* e *sIFM-trap* costruiscono statistiche considerando 100 trappole differenti per una singola configurazione atomistica di drogaggio (il che forma 100 diverse coppie doping-trap). Nella realtà non abbiamo trappole o profili di drogaggio ripetuti, bensì su  $N$  dispositivi troviamo  $N$  coppie doping-trap dove i profili e la posizione delle trappole non sono ripetuti. Per verificare che la statistica adottata nella tesi sia in linea con quest’ultima è necessario fare un confronto “MC”-“MC”. In un caso costruiamo la statistica su  $2N_TN_D$  simulazioni (2 per ogni coppia doping-trap), nell’altro su  $(1 + N_T)N_D$  (in quanto non devo



risimulare i casi di trappola scarica per le coppie doping-trap aventi il profilo di drogaggio in comune). Il confronto tra queste due metodologie di costruzione della statistica relativa all'RTN è visualizzabile in Fig.3.50.

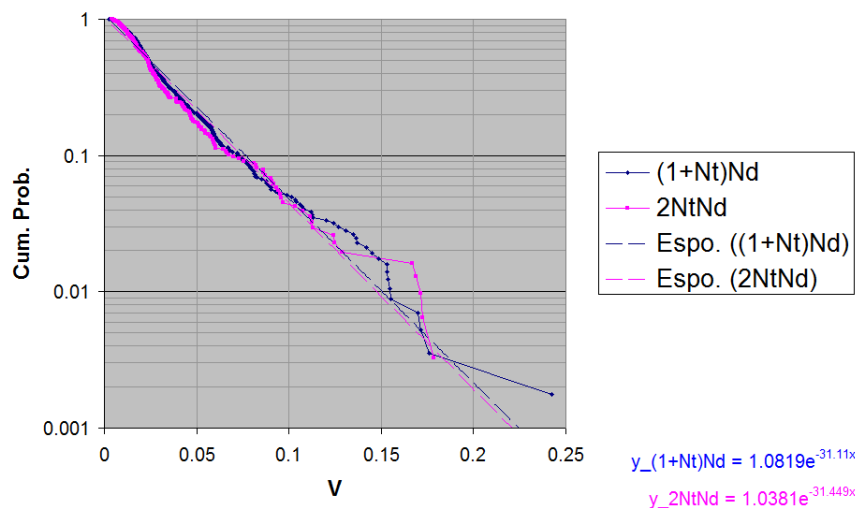


Figura 3.50: cumulative costruite con le “metodologie”  $(1 + N_T)N_D$  e  $2N_TN_D$

Notiamo dalla figura come le due cumulate relative alle due metodologie discusse sono praticamente sovrapposte e danti le stesse pendenze  $\lambda$ . Dev'essere così, poiché è vero che consideriamo per diversi dispositivi le stesse configurazioni di drogaggio, ma cambiando la posizione della trappola le coppie diventano del tutto scorrelate. Questo mostra, inoltre, come adottare la statistica  $(1 + N_T)N_D$  sia vantaggioso per il metodo MC, in quanto si guadagna un fattore  $\frac{2N_T}{1+N_T}$  in numero totale di simulazioni.

### 3.2.6 Osservazioni conclusive sui metodi

In questa tesi ho utilizzato due innovativi metodi per lo studio delle fluttuazioni statistiche dell'RTN: *sIFM-trap* e *sIFM-doping*. Il primo linearizza l'effetto della trappola sulla conduzione, il secondo l'effetto degli atomi droganti. Abbiamo visto come essi risultano essere quasi sempre delle ottime approssimazioni, in particolare lo *sIFM-trap* approssima eccellentemente il riferimento MC qualora la conduzione nel canale è abbastanza uniforme, e lo *sIFM-doping* fa lo stesso nei casi in cui il canale risulta svuotato dal drogaggio. Le cose vanno male, invece, per lo *sIFM-trap* nei casi di doping retrogradi perché “soffre” degli effetti non lineari sui bordi, e per lo *sIFM-doping* nei casi di drogaggi più alti sui bordi, in quanto i casi patologici non si mimetizzano più nella cumulativa e formano delle code non fisiche.

Per quanto detto e mostrato nell'elaborato di tesi, i metodi *sIFM-doping* e *sIFM-trap* tendono a funzionare bene per quanto riguarda la costruzione della cumulativa quasi sempre, se si vede però il singolo campione lo shift di soglia è spesso mal stimato, peggio ancora per quanto riguarda le transcaratteristiche singole, soprattutto in riferimento al metodo *sIFM-doping*. Per fare un esempio, se abbiamo un dispositivo “sbilanciato” con gli atomi soprattutto in un lato e la trappola posizionata sul lato opposto, gli shift di soglia calcolati spesso sono molto lontani da quelli del MC, e le transcaratteristiche ancora di più. Se invece la trappola viene messa lungo il percorso “debole” osserviamo che, mentre lo *sIFM-trap* stima bene la  $\Delta V_T$ , lo *sIFM-doping* calcola una  $\Delta V_T < 0$ . Questo si può dire in generale, mentre nei casi in cui i metodi funzionano bene (citati in precedenza), possiamo anche vedere il singolo campione della statistica e le transcaratteristiche ricostruite, l'accordo infatti risulta perfetto.

In linea generale direi che alcune correzioni sul metodo possono essere fatte. Per esempio dopo aver ricostruito i 10000 casi, si prendono 100 coppie doping-trap e si simulano DD (200 simulazioni per correggere lo *sIFM-doping* e 100 simulazioni per lo *sIFM-trap*, in quanto il caso di trappola neutra è già stato simulato). Dopodiché si confrontano i risultati e si leggono gli scatterplots in modo tale da capire quanta correlazione ci sia e su che linea di tendenza. In vero, se abbiamo che il metodo tende a sovra-stimare o sotto-stimare troppo il MC, ci si applica una correzione empirica correggendo ogni shift di soglia applicando la formula della linea di tendenza stessa.

I metodi descritti risultano approssimazioni del MC, ma estremamente più veloci di esso. Infatti per una statistica costruita su  $N_T \times N_D$  valori dello shift di soglia sono necessarie  $(1 + N_T)N_D$  simulazioni MC (vedi paragrafo precedente),  $(1 + N_T)$  simulazioni per il metodo *sIFM-doping* e  $N_D$  simulazioni per lo *sIFM-trap*. Essendo che però le simulazioni per i metodi *sIFM* necessitano per ogni bias point, oltre alla risoluzione DD, anche il calcolo delle funzioni di Green, le simulazioni per essi durano un tempo maggiore. Tra una simulazione DD ed una DD in cui è stato attivato il metodo IFM c'è pertanto un fattore  $k > 1$  per quanto riguarda il tempo di calcolo. Tal fattore dipende esponenzialmente dalla mesh; per le mesh utilizzate nel mio lavoro di tesi esso risulta essere  $k \approx 2$ . Pertanto i guadagni temporali per i metodi *sIFM-doping* e *sIFM-trap* risultano essere  $N_D/k$  e  $(1 + N_T)/k$  rispettivamente ( $\approx 50$  nel mio caso). Per il caso *sIFM-trap* ho aggiunto (nella tabella sotto) anche un fattore  $\alpha$  in quanto è necessario, concluse le simulazioni, valutare la funzione di Green all'interfaccia nodo per nodo associandola alle coordinate spaziali corrispondenti. Diciamo che nel mio caso il tempo per far ciò è  $\approx 1sec$  per ogni funzione di Green calcolata (cioè per ognuno dei 41 bias point e per ognuno dei 100 dispositivi), ovvero  $\alpha \approx 4100sec$ .

Metodo	$N_{simulazioni}$	tempo(1=1simMC)
MC	$(1 + N_T)N_D$	$(1 + N_T)N_D$
sIFM-doping	$(1 + N_T)$	$(1 + N_T)k$
sIFM-trap	$N_D$	$(N_D)k + \alpha$

Detto tutto questo, potrei consigliare qualche approfondimento sul metodo. Abbiamo visto come per il drogaggio retrogrado con  $t_{epi} = 16nm$ , lo *sIFM-doping* stima perfettamente gli shift di soglia dei singoli campioni della statistica, mentre nel caso uniforme no. Un'idea può essere quella di capire quale sia la  $t_{epi}^{minima}$  per la quale i campioni singoli vengono ben stimati dal metodo e come essa dipenda dalla concentrazione. Abbiamo poi visto come uniformizzando la conduzione nel canale alzando il drogaggio sui bordi lo *sIFM-trap* sia un'approssimazione eccellente al MC. Dovrebbe accadere lo stesso nel caso di drogaggio uniforme arrotondando gli spigoli della cella. Un'idea può essere quella di studiare questo caso e verificare che ciò accada. Infine, un ultimo spunto può essere quello di studiare in maniera più approfondita gli effetti non lineari della trappola e pertanto i segni della  $\Delta V_T''(dq)$ , capendo soprattutto cosa succede nel caso di drogaggio retrogrado in relazione alla forte sovrastima del metodo *sIFM-trap*.

## Capitolo 4

# Applicazione di modelli analitici su un MOS

In questo breve capitolo mostro un esempio di applicazione di alcuni modelli analitici per quanto riguarda transistori MOS non particolarmente scalati, i cui dati sono stati presi da [27] e [28]. Per tali dispositivi si è osservato che essi non seguono l'andamento previsto dai modelli di Asenov, Takeuchi e Stolk per l'RDF, ma presentano dei picchi di dispersioni della tensione di soglia a circa  $1/\sqrt{WL} = 1\mu m^{-1}$ , il che non è assolutamente previsto dalla teoria. Cercheremo pertanto la causa di tali dispersioni in diversi tipi di fluttuazioni. Alla fine di questa analisi la spiegazione più plausibile ritengo essere il fatto che tal dispositivo nella regione di  $1/\sqrt{WL} = 1\mu m^{-1}$  entri nella zona di roll-off e pertanto le marcate fluttuazioni di soglia si possono spiegare con fluttuazioni geometriche.

Tal lavoro esula quasi del tutto dall'argomento principale della tesi, ovvero ciò che riguarda il metodo IFM, ma è comunque nell'ambito della *Variability*.

### 4.1 Il dispositivo e le fluttuazioni “anomale”

In figura Fig.4.1 troviamo i dati relativi alla  $\sigma_{V_T}$  per un MOS al variare della sua area, o meglio al variare della  $1/\sqrt{WL}$ . Quello che dovremmo vedere in questo grafico è una retta che si sovrappone ad una costante. La costante è relativa al “fondo”, fluttuazioni statistiche generiche, mentre la retta è relativa alla maggior causa di fluttuazione della soglia: l'RDF. Infatti, come ricordo dai modelli analitici 1.21, 1.18 e 1.20, la dispersione della tensione di soglia va come

l'inverso della radice quadrata dell'area  $WL$ . Dall'immagine vediamo che ciò non accade, infatti la curva si biforca a seconda di quale larghezza consideriamo. In [27] si è pertanto fittata la curva secondo il modello empirico:

$$\sigma_{V_T} = \sqrt{\left(\frac{\partial V_T}{\partial t_{ox}}\right)^2 \sigma_{t_{ox}}^2 + \left(\frac{C_{ms}}{\sqrt{WL}}\right)^2 + \left(\frac{C_{ss}}{L^3}\right)^2} \quad (4.1)$$

dove  $\partial V_T/\partial t_{ox} = 1.8V/\mu m$ ,  $C_{ms} = 4 \times 10^{-3}V\mu m$  e  $C_{ss} = 1.2 \times 10^{-5}V\mu m^{-3}$ .

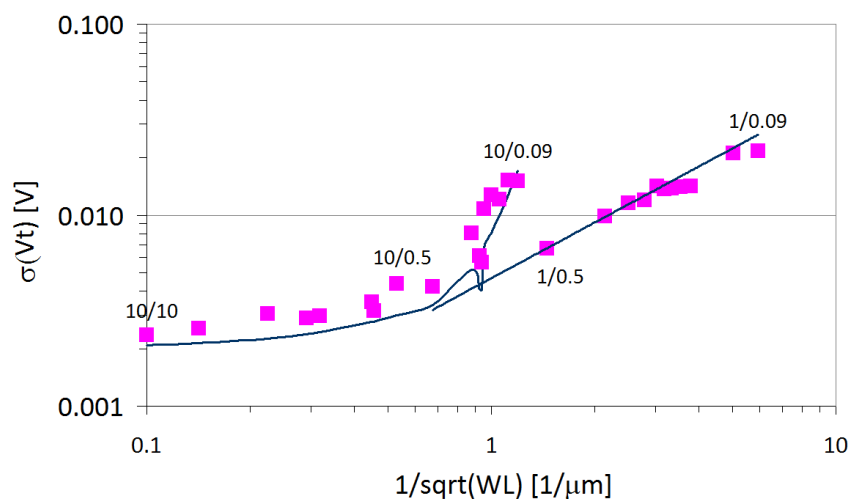


Figura 4.1: dispersione della tensione di soglia al variare dell'area. Dati sperimentali ([27]).

Il primo termine dell'equazione 4.1 riguarda il "fondo", dovuto alla fluttuazione dello spessore dell'ossido, che risulta essere l'effetto di dispersione dominante per aree sufficientemente grandi, ovvero  $1/\sqrt{WL}$  piccole. Il secondo termine è relativo all'RDF, mentre il terzo è il termine di "correzione" empirica aggiunto fittando i dati.

In realtà in questo capitolo non faremo riferimento ai dispositivi relativi ai dati della figura 4.1, ma faremo riferimento ad altri. Questo perché ho potuto accedere solo a quest'ultimi; il concetto però non cambia in quanto tali "picchi" di dispersione della  $\sigma_{V_T}$  si ripetono anche in questo caso, seppur con intensità minore. I dati relativi ai dispositivi in questione sono visualizzabili in Fig.4.2 (scala bilogartitmica) e 4.3 (scala lineare). Si nota meno la biforcazione rispetto

ai dati precedenti, ma essa è presente: i dati fuori tendenza sono stati evidenziati in Fig.4.2 e sono ben visibili in scala lineare (Fig.4.3).

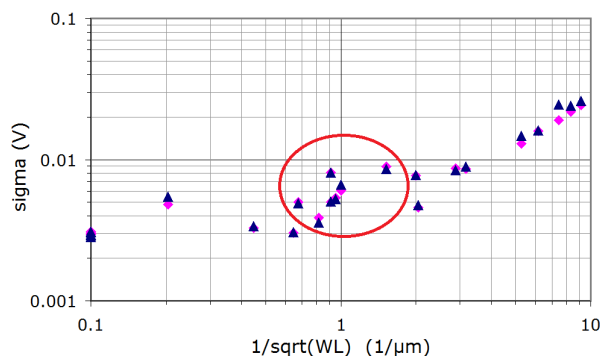


Figura 4.2: dispersione della tensione di soglia al variare dell'area. I colori diversi sono relativi ai due MOS tra di loro "matched" disegnati su silicio

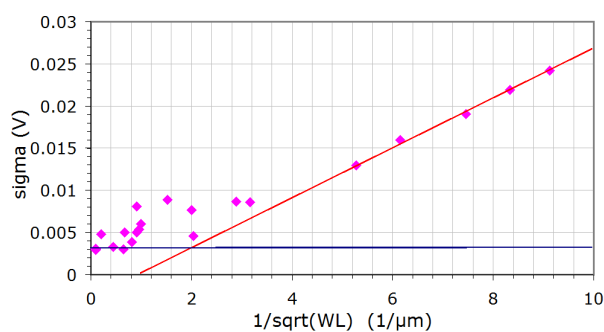


Figura 4.3: dispersione della tensione di soglia al variare dell'area.

Per l'applicazione dei modelli analitici è necessario semplificare la struttura, per esempio considerando la  $N_A^{eq}$  tale per cui si ha la stessa caduta di potenziale nel bulk avendo quel determinato profilo non analitico  $N_A(r)$ . Per far ciò ho utilizzato il dispositivo risultato delle simulazioni di processo (fatte con il tool *Sprocess*), per farne una caratterizzazione elettrica (con *Sdevice*). Da essa estrapolo a soglia la  $N_A^{eq}$ , la  $\phi_B^{eq}$  e la  $W_d^{eq}$ . I dispositivi simulati sono dispositivi 2D, il primo (Fig.4.4(a)) ha lunghezza di canale  $100nm$ , il secondo (Fig.4.4(b))  $10\mu m$ . Ovviamente il dispositivo uscente dalla simulazione di processo così com'è non risulta adatto a quella elettrica, poiché per esempio avente una mesh implementata

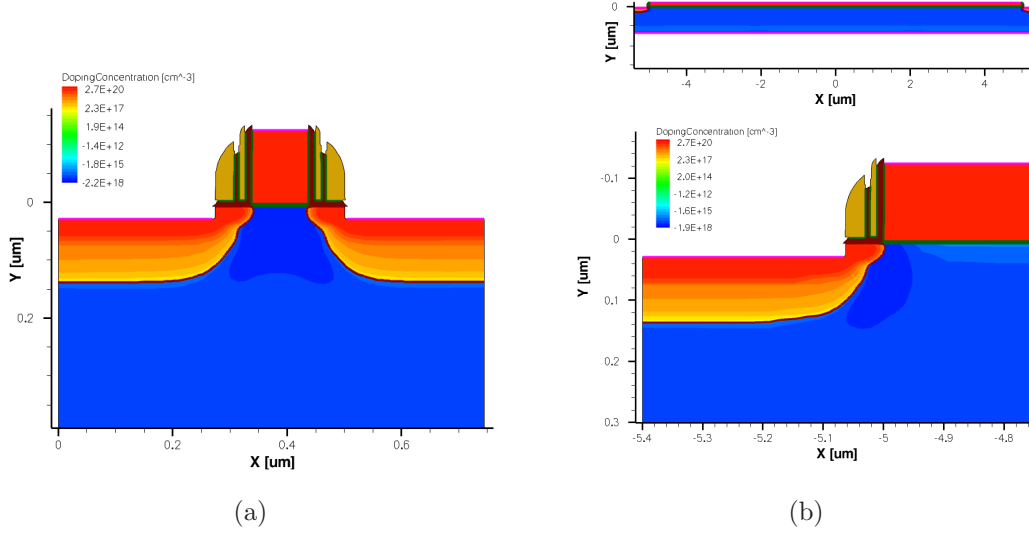


Figura 4.4: Dispositivo con (a)  $L_{ch} = 100nm$  e (b)  $L_{ch} = 10\mu m$

a diverse esigenze. Pertanto risulta necessario fare alcune modifiche preliminari: togliere ed attaccare ossidi, nitruri e siliciuri “inutili”, specchiare la struttura in quanto simulata solo metà, settare i contatti, implementare una mesh più adatta alla simulazione elettrica (per esempio raffinare la regione interfacciale tra ossido e silicio dove la conduzione è maggiore) e caricare i dataset delle concentrazioni di drogaggio interpolandoli sulla nuova mesh. Questi (ed altri) passaggi vengono fatti con Sprocces, Sde, Snmesh: si passa dalla Fig.4.5(a) alla Fig.4.5(b). Ora, plottando il profilo del campo elettrico a soglia come in Fig.4.6 ottengo  $N_A^{eq}$ ,  $\phi_B^{eq}$  e la  $W_d^{eq}$  iterando le tre equazioni sottostanti a partire dal campo elettrico massimo:

$$N_A = \frac{E_{max}\epsilon_{Si}}{qW_d} \quad (4.2)$$

$$\phi_B = \frac{kT}{q} \ln \left( \frac{N_A}{n_i} \right) \quad (4.3)$$

$$W_d = \frac{4\phi_B}{E_{max}} \quad (4.4)$$

La tensione di soglia, invece, la calcoliamo da Sdevice direttamente. A questo punto consideriamo i due dispositivi con area uguale, ovvero con  $1/\sqrt{WL} = 1\mu m^{-1}$ : quello a canale lungo risulta essere stretto ( $W = 0.1\mu m$ ), il dispositivo corto in-

vece largo ( $W = 10\mu m$ ). Dai modelli analitici per l'RDF vediamo che:

$\sigma_{V_T}$ (mV)	Stolk	Takeuchi	Asenov
$L = 0.1\mu m$	1.36	1.58	1.15
$L = 10\mu m$	0.84	1.26	0.56

Applicando il metodo nIFM otteniamo dispersioni analoghe:  $\sigma_{V_T}^{0.1\mu m} = 1.724mV$  e  $\sigma_{V_T}^{10\mu m} = 0.986mV$ . Ciò che osserviamo da tutti questi risultati è che il dispositivo corto risulta avere poca dispersione in più della tensione di soglia, ma non tanto di più in modo da giustificare i “picchi” della Fig.4.2. Dai dati sperimentali risulta una  $\sigma_{V_T}^{0.1\mu m} > 6mV$ ! Stiamo sbagliando di un fattore 4 insomma.

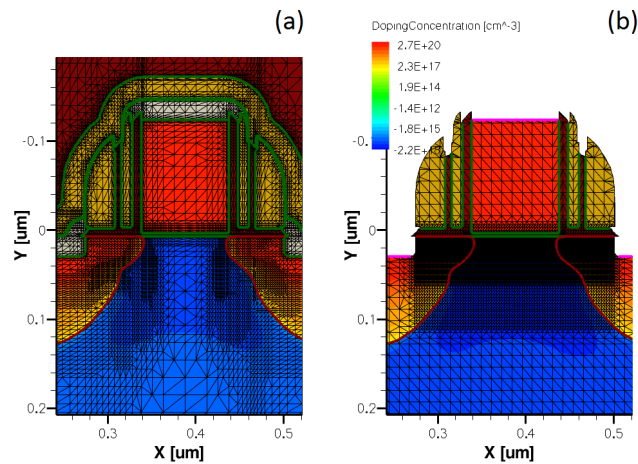


Figura 4.5: (a) dispositivo uscente da Sprocess; (b) dispositivo usato come input per Sdevice.

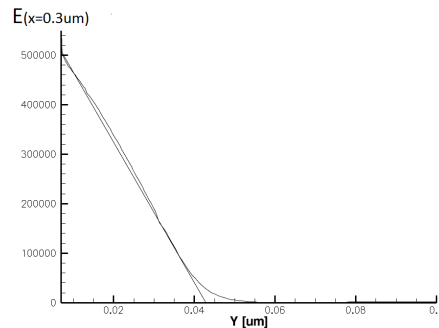


Figura 4.6: campo elettrico a soglia per il dispositivo  $100nm$



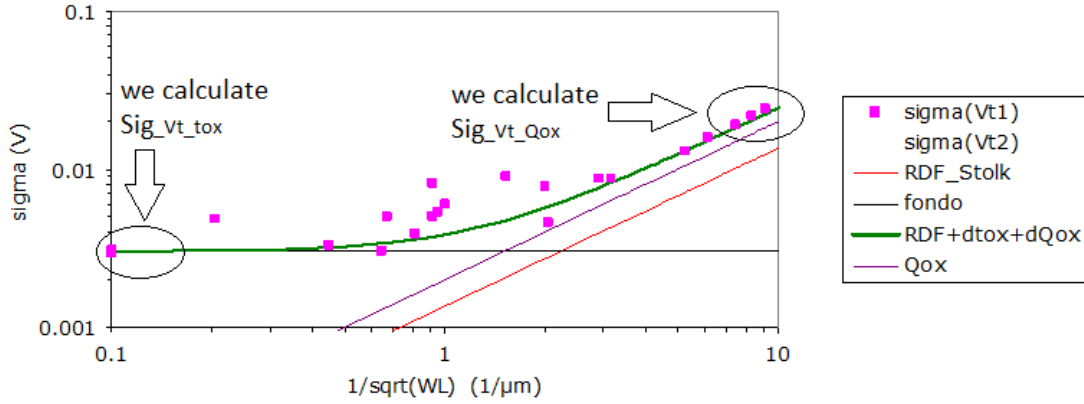


Figura 4.7: dispersione della tensione di soglia al variare dell'area. Evidenziati i modelli analitici

Dai dati equivalenti trovati possiamo ipotizzare diverse cause di dispersioni statistiche. Ad aree “grandi” ( $x$  basse in Fig.4.7) domina la fluttuazione dello spessore dell'ossido, ad aree piccole ( $x$  alte in Fig.4.7) ipotizziamo che domina la fluttuazione delle cariche nell'ossido secondo la statistica di Poisson. Pertanto troviamo:

$$\sigma_{V_T}^{tox} \approx 3mV \quad \sigma_{V_T}^{tox} \approx \sigma_{tox} \frac{\partial V_T}{\partial tox} \rightarrow \sigma_{tox} \approx 0.051nm = 2.17\% \quad (4.5)$$

$$\sigma_{V_T}^{Qox} \approx \frac{2mV \mu m}{\sqrt{WL}} \quad \sigma_{V_T}^{Qox} \approx \frac{q\sqrt{N'_{eq}}}{C'_{ox}\sqrt{WL}} \rightarrow N'_{eq} \approx 3.5 \times 10^{12} cm^{-2} \quad (4.6)$$

Tali dati risultano assolutamente ragionevoli. L'unico modo per spiegare i campioni che rimangono fuori da questi modelli sono le fluttuazioni geometriche. Probabilmente il transistor si trova nella regione di roll-off e una piccola variazione sulla lunghezza del canale si tramuta in una variazione di tensione di soglia che domina a  $\sqrt{WL} = 1\mu m$  aventi  $L$  molto corte. Per dimostrare ciò focalizzo l'attenzione solo sul dispositivo “corto e largo” avente  $L = 100nm$  utilizzando un modello analitico per il roll-off del transistor (mostrato dettagliatamente nella prossima sezione).

## 4.2 Modello analitico per il roll-off del MOS

Il transistor MOS a canale lungo ha la tensione di soglia idealmente indipendente dalla lunghezza del canale. Ciò è vero fin quando l'approssimazione di canale lungo risulta valida; in vero se l'effetto delle giunzioni è rilevante riscontriamo i fenomeni di canale corto (SCE), come per esempio l'effetto DIBL, ovvero l'induzione da parte del drain dell'abbassamento dell'energia potenziale nel canale, il che provoca un'abbassamento della tensione di soglia dal momento in cui le giunzioni aiutano lo svuotamento del canale. Un'altro effetto, molto correlato al DIBL, è il cosiddetto roll-off del transistor, ovvero l'abbassamento della soglia con il diminuire della lunghezza del MOS.

### 4.2.1 Il modello

In questa sezione mostro un modello analitico in cui sono presenti un gran numero di approssimazioni in modo da considerare solo gli aspetti base per quanto riguarda l'effetto di canale corto. Una geometria 2D semplificata è visualizzabile in Fig.4.8. Consideriamo il potenziale intrinseco  $\psi(x, y)$  in un substrato uniformemente drogato  $N_A$  in cui assumiamo completo svuotamento in  $ABEF$ . L'equazione di Poisson si trasforma nella risoluzione del seguente problema differenziale per il potenziale:

$$AFGH \quad \frac{\partial^2 \psi(x, y)}{\partial x^2} + \frac{\partial^2 \psi(x, y)}{\partial y^2} = 0 \quad (4.7)$$

$$ABEF \quad \frac{\partial^2 \psi(x, y)}{\partial x^2} + \frac{\partial^2 \psi(x, y)}{\partial y^2} = \frac{qN_A}{\epsilon_{Si}} \quad (4.8)$$

Facciamo ora un'ulteriore approssimazione: volendo eliminare la condizione al contorno per la quale il campo elettrico ha un salto di circa  $\epsilon_{Si}/\epsilon_{ox} \approx 3$  all'interfaccia Si/Ox, consideriamo l'ossido spesso il triplo ( $3t_{ox}$ ) e costante dielettrica costante ovunque =  $\epsilon_{Si}$ , in modo da preservare la capacità dell'ossido ed avere una geometria equivalente con materiale omogeneo. A questo punto imponiamo le condizioni al contorno:

$$GH \quad \psi(-3t_{ox}, y) = V_{gs} - V_{fb} \quad (4.9)$$

$$AB \quad \psi(x, 0) = \psi_{bi} \quad (4.10)$$

$$EF \quad \psi(x, L) = \psi_{bi} + V_{ds} \quad (4.11)$$

$$CD \quad \psi(W_d, y) = 0 \quad (4.12)$$

dove la distanza  $BC \approx W_S = \sqrt{\frac{2\epsilon_{Si}\psi_{bi}}{qN_A}}$  e la distanza  $DE \approx W_D = \sqrt{\frac{2\epsilon_{Si}(\psi_{bi}+V_{ds})}{qN_A}}$ . Inoltre in  $FG$  e  $HA$  imponiamo una variazione lineare del potenziale, e in  $BC$  e  $DE$  parabolica.

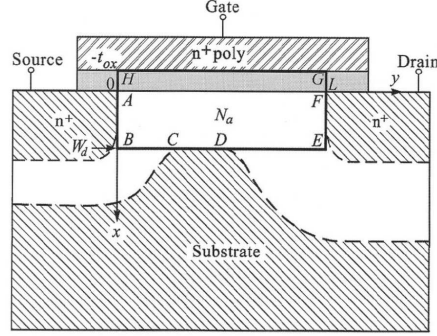


Figura 4.8: geometria 2D semplificata del dispositivo

### 4.2.2 Tecniche di soluzione

Una tecnica di soluzione fa uso del principio di sovrapposizione degli effetti rompendo il potenziale elettrico nei seguenti termini:

$$\psi(x, y) = v(x, y) + u_L(x, y) + u_R(x, y) + u_B(x, y) \quad (4.13)$$

in cui  $v(x, y)$  risolve la Poisson non omogenea considerando la condizione al contorno superiore, mentre  $u_L(x, y)$ ,  $u_R(x, y)$  e  $u_B(x, y)$  risolvono la Poisson omogenea (equazione di Laplace sul rettangolo) facendo in modo che la  $\psi(x, y)$  soddisfi rispettivamente la condizione al contorno sinistra, destra e sul bordo inferiore. In altre parole  $v(x, y) + u_L(x, y)$  soddisfa la condizione al contorno su  $HB$ ,  $v(x, y) + u_R(x, y)$  su  $GE$  e  $v(x, y) + u_B(x, y)$  su  $BE$ .

Una soluzione naturale per la  $v(x, y)$  è l'espressione del potenziale 1D del MOS a canale lungo:

$$v(x, y) = \psi_s - \frac{V_{gs} - V_{fb} - \psi_s}{3t_{ox}} x \quad x \in [-3t_{ox}, 0] \quad (4.14)$$

$$v(x, y) = \psi_s \left(1 - \frac{x}{W_d}\right)^2 \quad x \in [0, W_d] \quad (4.15)$$

Qui  $\psi_s$  è il potenziale superficiale di canale lungo, ed è relazionato alla larghezza della regione di svuotamento  $W_d$  dall'espressione  $\psi_s = \frac{qN_A W_d^2}{2\epsilon_{Si}}$ . Da notare come

tale soluzione non solo soddisfa la condizione al contorno su  $HG$ , ma anche su  $CD$ . La richiesta poi che essa abbia derivata continua all'interfaccia si tramuta nella ben nota relazione tra  $\psi_s$  e  $V_{gs}$ :

$$\frac{V_{gs} - V_{fb} - \psi_s}{3t_{ox}} = \frac{2\psi_s}{W_d} = \sqrt{\frac{2qN_A\psi_s}{\epsilon_{Si}}} \quad (4.16)$$

Per gli altri termini bisogna trovare la soluzione per l'equazione differenziale di Laplace sul rettangolo imponendo una condizione al contorno solo in un lato. Consideriamo il caso  $u_L(x, y)$ : l'equazione da risolvere è la 4.17 con condizione al contorno raffigurata dall'immagine 4.9.

$$\frac{\partial^2 u_L(x, y)}{\partial x^2} + \frac{\partial^2 u_L(x, y)}{\partial y^2} = 0 \quad (4.17)$$

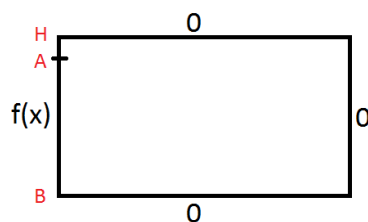


Figura 4.9: condizioni al contorno per  $u_L(x, y)$ , in cui  $f(x) = \psi_{bi} - v(x, 0)$  in AB e  $f(x) = \frac{\psi_{bi} - (V_{gs} - V_{fb})}{3t_{ox}}(x + 3t_{ox}) + (V_{gs} - V_{fb}) - v(x, 0)$  in AH

Per risolvere ciò applichiamo il metodo di separazione delle variabili 4.18-4.19: l'equazione 4.17 si trasforma nel sistema di equazioni indipendenti 4.20.

$$u_L(x, y) = X(x)Y(y) \quad (4.18)$$

$$\frac{X''(x)}{X(x)} = -\frac{Y''(y)}{Y(y)} = -\lambda \quad (4.19)$$

$$\begin{cases} X''(x) + \lambda X(x) = 0 \\ Y''(y) - \lambda Y(y) = 0 \end{cases} \quad (4.20)$$

Per quanto riguarda la funzione  $X(x)$ , in particolare, dobbiamo risolvere il seguente problema agli autovalori:

$$\begin{cases} X''(x) + \lambda X(x) = 0 \\ X(-3t_{ox}) = X(W_d) = 0 \end{cases} \quad (4.21)$$

Essendo che la 4.21 risulta un problema di Sturm Liouville regolare, le autosoluzioni  $\{X_n(x)\}_{n=1}^{\infty}$  rappresentano un sistema ortonormale completo in  $L^2(-3t_{ox}, W_d)$ :

$$X_n(x) = b_n \sin\left(\frac{n\pi(x + 3t_{ox})}{W_d + 3t_{ox}}\right) \quad (4.22)$$

Gli autovalori risultano  $\lambda_n = \left(\frac{n\pi}{W_d + 3t_{ox}}\right)^2$ , pertanto per la  $Y(y)$  dobbiamo risolvere il seguente problema:

$$\begin{cases} Y''(y) - \lambda_n Y(y) = 0 \\ Y(L) = 0 \end{cases} \quad (4.23)$$

La soluzione è del tipo  $Y(y) = c_1 e^{\sqrt{\lambda_n} y} + c_2 e^{-\sqrt{\lambda_n} y}$ . Imponendo la condizione al contorno essa diventa:

$$Y_n(y) = c_n \operatorname{Sh}\left(\frac{n\pi(L - y)}{3t_{ox} + W_d}\right) \quad (4.24)$$

Infine, essendo che la somma pesata di soluzioni indipendenti di un' equazione omogenea è ancora soluzione, la  $u_L(x, y)$  diventa:

$$u_L(x, y) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi(x + 3t_{ox})}{W_d + 3t_{ox}}\right) \operatorname{Sh}\left(\frac{n\pi(L - y)}{3t_{ox} + W_d}\right) \quad (4.25)$$

Imponiamo ora l'ultima condizione 4.26, che diventa la 4.27 e semplificata la 4.28:

$$\psi(x, 0) = v(x, 0) + u_L(x, 0) \quad (4.26)$$

$$\psi(x, 0) - v(x, 0) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi(x + 3t_{ox})}{W_d + 3t_{ox}}\right) \operatorname{Sh}\left(\frac{n\pi L}{3t_{ox} + W_d}\right) \quad (4.27)$$

$$f(x) = \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi(x + 3t_{ox})}{W_d + 3t_{ox}}\right) \quad (4.28)$$

Ora, dall'analisi di Fourier sugli spazi di Hilbert, sappiamo che  $\forall f(x) \in L^2(-3t_{ox}, W_d)$  i coefficienti  $B_n$  si trovano come prodotto scalare nello spazio tra la funzione e i vettori della base, dunque:

$$B_n = \frac{2}{W_d + 3t_{ox}} \int_{-3t_{ox}}^{W_d} f(x) \sin\left(\frac{n\pi(x + 3t_{ox})}{W_d + 3t_{ox}}\right) dx \quad (4.29)$$

Nel mezzo del dispositivo,  $y \approx L/2$ ,  $u_L$  varia come  $\exp\{-n\pi L/[2(W_d + 3t_{ox})]\}$ ; se il canale non è troppo corto (il caso studiato in questo capitolo per esempio)

i termini di ordine superiore possono essere trascurati. Ciò porta ad integrare l'equazione 4.29 solo per  $n = 1$ :

$$B_1 = \frac{2}{\pi} \left[ \left( 1 + \frac{\sin\theta}{\theta} \right) \psi_{bi} - \left( \frac{\sin\theta}{\theta} + \frac{2\sin\theta}{\pi - \theta} - \frac{2(1 + \cos\theta)}{(\pi - \theta)^2} \right) \psi_s \right] \quad (4.30)$$

dove  $\theta = \pi(3t_{ox})/(W_d + 3t_{ox})$ .

Analogamente, per  $u_R(x, y)$  troviamo:

$$u_R(x, y) = \sum_{n=1}^{\infty} c_n \sin \left( \frac{n\pi(x + 3t_{ox})}{W_d + 3t_{ox}} \right) Sh \left( \frac{n\pi y}{3t_{ox} + W_d} \right) \quad (4.31)$$

$$\begin{aligned} C_1 &= c_1 Sh \left( \frac{\pi L}{3t_{ox} + W_d} \right) \\ &= \frac{2}{\pi} \left[ \left( 1 + \frac{\sin\theta}{\theta} \right) (\psi_{bi} + V_{ds}) - \left( \frac{\sin\theta}{\theta} + \frac{2\sin\theta}{\pi - \theta} - \frac{2(1 + \cos\theta)}{(\pi - \theta)^2} \right) \psi_s \right] \end{aligned} \quad (4.32)$$

Il termine rimanente  $u_B(x, y)$  può tranquillamente esser trascurato, in quanto in  $CD$  la condizione al contorno è già soddisfatta, ed i contributi rimanenti in  $BC$  e  $DE$  portano ai coefficienti  $D_n$  che sono molto minori di  $B_n$  e  $C_n$ .

La soluzione approssimata per il potenziale elettrico risulta, nella regione  $x \in [0, W_d]$ :

$$\psi(x, y) = \psi_s \left( 1 - \frac{x}{W_d} \right)^2 + \frac{B_1 Sh \left( \frac{\pi(L-y)}{3t_{ox} + W_d} \right) + C_1 Sh \left( \frac{\pi y}{3t_{ox} + W_d} \right)}{Sh \left( \frac{\pi L}{3t_{ox} + W_d} \right)} \sin \left( \frac{\pi(x + 3t_{ox})}{W_d + 3t_{ox}} \right) \quad (4.33)$$

### 4.2.3 Tensione di soglia e dispersione

Per conoscere la tensione di soglia consideriamo il potenziale superficiale  $\psi(0, y)$ . Esso ha un minimo per (utilizzando le approssimazioni di [8])

$$y = y_c \approx \frac{L}{2} - \frac{W_d + 3t_{ox}}{2\pi} \ln \left( \frac{C_1}{B_1} \right) \approx \frac{L}{2} - \frac{W_d + 3t_{ox}}{2\pi} \ln \left( 1 + \frac{V_{ds}}{\psi_{bi} - 0.4\psi_s} \right) \quad (4.34)$$

che corrisponde al punto di massimo della barriera di potenziale. Esso quando il drain è basso è a centro canale  $L/2$ , quando il drain è alto si avvicina al source. Nella regione di sotto-soglia, la densità di carica di canale varia esponenzialmente con il potenziale in quel punto. Pertanto la conduzione di corrente è controllata

dal potenziale più basso nel canale, ovvero dal potenziale in  $y = y_c$ . Sostituendo otteniamo:

$$\psi(0, y_c) = \psi_s + 2\sqrt{B_1 C_1} \exp\left(-\frac{\pi L/2}{W_d + 3t_{ox}}\right) \sin\left(\frac{\pi(3t_{ox})}{W_d + 3t_{ox}}\right) \quad (4.35)$$

Nell'equazione 4.35 il primo termine è quello relativo al canale lungo, il secondo termine rappresenta l'innalzamento indotto dalle giunzioni. Pertanto per effetto del canale corto abbiamo il  $\Delta\psi_s$ . Per tradurlo in variazione della tensione di soglia dobbiamo passare attraverso il fattore  $\Delta V_T / \Delta\psi_s = m = 1 + \frac{3t_{ox}}{W_{dm}}$ . L'abbassamento della tensione di soglia risulta pertanto:

$$\Delta V_T = 2m\sqrt{B_1 C_1} \exp\left(-\frac{\pi L/2}{W_d + 3t_{ox}}\right) \sin\left(\frac{\pi(3t_{ox})}{W_d + 3t_{ox}}\right) \quad (4.36)$$

Da tal modello analitico possiamo dire che per MOS in cui sono presenti effetti di canale corto la tensione di soglia si può scrivere come:

$$V_T = V_{T_0} - \Delta V_T(L) \quad (4.37)$$

dove il secondo termine è espresso dall'equazione 4.36 e dipende dalla lunghezza del canale. Pertanto se abbiamo fluttuazioni di processo tali per cui abbiamo una dispersione per quanto riguarda la lunghezza del canale, la tensione di soglia varierà come:

$$\begin{aligned} \sigma_{V_T} &\approx \frac{\partial V_T}{\partial L} \delta L = \frac{\partial \Delta V_T}{\partial L} \delta L \\ &\approx \frac{m\pi}{W_d + 3t_{ox}} \sqrt{B_1 C_1} \exp\left(-\frac{\pi L/2}{W_d + 3t_{ox}}\right) \sin\left(\frac{\pi(3t_{ox})}{W_d + 3t_{ox}}\right) \sigma_L \end{aligned} \quad (4.38)$$

L'equazione 4.38 ci dà un'espressione analitica per la deviazione standard della soglia  $\sigma_{V_T}$  indotta da una dispersione della lunghezza di canale  $\sigma_L$  nella regione di roll-off del transistor.

### 4.3 Risultati del modello

In riferimento al modello 4.38 e per il MOS lungo  $L = 100nm$ , ci manca ancora la  $\psi_{bi}$ , che ho stimato essere  $\psi_{bi} \approx 1V$  a soglia. Ora, assumendo che:

$$\sigma_{V_T} = \sqrt{(\sigma_{V_T}^{tox})^2 + (\sigma_{V_T}^{Q_{ox}})^2 + (\sigma_{V_T}^{RDF})^2 + (\sigma_{V_T}^{\Delta L})^2} \approx 6.6mV \quad (4.39)$$

$\sigma_{V_T}^{\Delta L}$  risulta essere  $\approx 5.27mV$ . Utilizzando il modello 4.38 e i dati che abbiamo, la dispersione della lunghezza del canale risulterebbe essere  $\sigma_L \approx 5.7nm$  ovvero  $\sigma_L^{rel} \approx 5.7\%$ , il che è un po' troppo a dire il vero.

In realtà con questo modo di ragionare abbiamo fatto due approssimazioni ulteriori: innanzitutto adottando l'espressione 4.38 consideriamo solo la variazione al prim'ordine della dispersione, e poi nell'espressione 4.36 non abbiamo solo le fluttuazioni di  $L$ , ma anche tutte le altre che vanno a sommarsi. Pertanto ho instaurato un modellino Monte Carlo in excel in cui genero 300 dispositivi diversi randomizzando nel numero i droganti e le cariche nell'ossido secondo la statistica di Poisson e randomizzando 300 lunghezze di canale e 300 spessori dell'ossido secondo una distribuzione normale; poi ho fatto calcolare analiticamente lo shift di soglia 4.36 per ognuno e dunque la  $\sigma_{V_T}^{tot}$ . Risulta che sono sufficienti in realtà solo  $\sigma_L \approx 2.6nm$  ovvero una  $\sigma_L^{rel} \approx 2.6\%$  per arrivare a dispersioni di soglia di quell'entità. Tali fluttuazioni per la lunghezza risultano assolutamente nella norma nel processo di fabbricazione. Possiamo quindi dire che la causa principale della fluttuazione della soglia per tal dispositivo risulta essere la combinazione di fluttuazioni, in particolare quella su  $L$ , nella regione di roll-off del MOS.

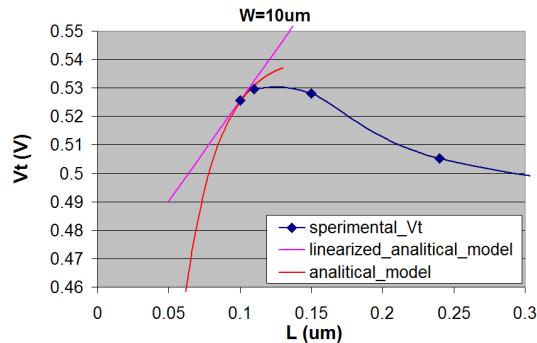


Figura 4.10:  $V_T$  medie per i dispositivi con  $W = 10\mu m$  al diminuire della lunghezza  $L$ . Focalizziamo l'attenzione per il device avente  $L = 100nm$



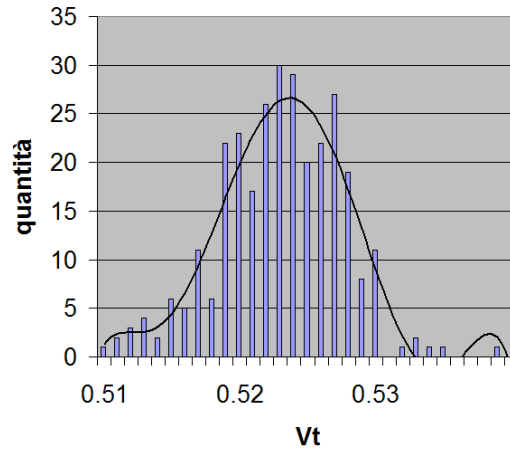


Figura 4.11: distribuzione della tensione di soglia ottenuta con il modello MC

In questo capitolo ho mostrato come gli effetti di canale corto possono indurre dispersioni significative della tensione di soglia. Ho dimostrato ciò facendo riferimento al roll-off relativo alla lunghezza del canale utilizzando i dati delle simulazioni di processo per il dispositivo avente  $L = 100nm$  ed il modello analitico corrispondente. Guardando però la figura Fig.4.2 si nota che accade la stessa cosa al diminuire della  $W$ . Infatti è anche presente il fenomeno del roll-off relativo alla larghezza del canale che pertanto induce picchi della  $\sigma_{V_T}$  analogamente al caso analizzato.

# Conclusioni

I dispositivi di ultima generazione possono essere considerati ormai “atomici”, nel senso che contengono qualche manciata di atomi. Le fluttuazioni statistiche delle caratteristiche elettriche di dispositivi nominalmente identici riscontrate sperimentalmente sono riconducibili, per la maggior parte, proprio alla natura atomistica del drogaggio, in particolare per quanto riguarda i fenomeni di RTN (*Random Telegraph Noise*) ed RDF (*Random Dopant Fluctuation*).

Il metodo IFM (acronimo di *Impedance Field Method*), implementato nei software della *Synopsys*, è un metodo approssimato per l’analisi statistica delle fluttuazioni dovute al fenomeno dell’RDF. Lo *statistical-IFM* (sIFM), in particolare, è una variante dell’IFM implementata per la ricostruzione in maniera veloce delle transcaratteristiche  $I_D - V_{CG}$  dei dispositivi “reali” (aventi il drogaggio la cui natura è discreta) a partire dal dispositivo di riferimento con doping continuo. In questo elaborato di tesi ho valutato la possibilità di investigare la statistica dell’ampiezza dell’RTN (*Random Telegraph Noise*) per dispositivi MOS di nanoscala facendo uso dello sIFM. Da tal metodo ho creato due implementazioni per l’RTN: lo *sIFM-doping* e lo *sIFM-trap*. In entrambi i casi i risultati mostrano che i metodi offrono un pratico e conveniente compromesso tra accuratezza e sforzo computazionale per la ricostruzione della distribuzione dell’ampiezza dell’RTN. Rispetto al riferimento MC (acronimo di *Monte Carlo*), infatti, le distribuzioni statistiche sono quasi sempre coincidenti ed ottenute con un vantaggio notevole in termini di tempo di simulazione. Per quanto riguarda invece il singolo campione della statistica i metodi faticano ad approssimare con sufficiente precisione l’ampiezza dello shift di soglia e le transcaratteristiche rispetto al riferimento MC. Questo non è sempre vero, ci sono casi particolari in cui i metodi, implementati in situazioni piuttosto favorevoli, costruiscono campione per campione le ampiezze  $\Delta V_T$  e le transcaratteristiche corrispondenti con errori trascurabili.

# Bibliografia

- [1] by Synopsys: “Modeling Random Dopant Fluctuation Effects in MOSFETs Using Impedance Field Methods”
- [2] by Synopsys: “Simulation of Random Dopant Fluctuation Effects in NMOSFETs ”
- [3] by Synopsys: “Modeling Random Variability Effects in Impedance Field Methods ”
- [4] Andreas Wettstein: “Modelling Random Dopant Fluctuations as Noise with Sentaurus Device ”
- [5] Peter A. Stolk, Frans P. Widdershoven, D. B. M. Klaassen: “Modeling Statistical Dopant Fluctuations in MOS Transistors”
- [6] K. Takeuchi, T. Fukai, T. Tsunomura, A. T. Putra, A. Nishida, S. Kamohara, T. Hiramoto: “Understanding Random Threshold Voltage Fluctuation by Comparing Multiple Fabs and Technologies”
- [7] Asen Asenov, Andrew R. Brown, John H. Davies, Savas Kaya, Gabriela Slavcheva: “Simulation of Intrinsic Parameter Fluctuations in Decanometer and Nanometer-Scale MOSFETs”
- [8] Yuan Taur, Tak H. Ning: “Fundamentals of Modern VLSI Devices”
- [9] Francesco Biancardi: “Analisi di Scaling delle Fluttuazioni della Tensione di Soglia in Dispositivi MOSFET”
- [10] Diego Frasson: “Sviluppo e Validazione di Modelli Discreti per lo Studio della Dispersione di Soglia e delle Correnti di Perdita nei MOSFET ”

- 
- [11] R. Gusmeroli, C. Monzio Compagnoni, A. Riva, A. S. Spinelli, A. L. Lacaita, M. Bonanomi, and A. Visconti: “Defects spectroscopy in  $SiO_2$  by statistical random telegraph noise analysis”
- [12] C. Monzio Compagnoni, R. Gusmeroli, A. S. Spinelli, A. L. Lacaita, M. Bonanomi, and A. Visconti: “Statistical model for Random Telegraph Noise in Flash Memories”
- [13] A. Asenov, R. Balasubramaniam, A. R. Brown, and J. H. Davies: “RTS amplitudes in decanometer MOSFETs: 3-D simulation study”
- [14] K. Fukuda, Y. Shimizu, K. Amemiya, M. Kamoshida, and C. Hu: “Random telegraph noise in Flash memories - model and technology scaling”
- [15] A. Ghetti, C. Monzio Compagnoni, A. S. Spinelli, A. Visconti: “Comprehensive Analysis of Random Telegraph Noise Instability and Its Scaling in Deca-Nanometer Flash Memories”
- [16] H. H. Mueller and M. Schulz: “Conductance modulation of submicrometer Metal-Oxide-Semiconductor Field-Effect Transistor by single-electro trapping”
- [17] H. H. Mueller and M. Schulz: “Random telegraph signal: an atomic probe of the local current in field-effect transistor”
- [18] N. Tega, H. Miki, T. Osabe, A. Kotabe, K. Otsuga, H. Kurata, S. Kamohara, K. Tokami, Y. Ikeda, and R. Yamada: “Anomalously large threshold voltage fluctuation by complex Random Telegraph Signal in floating gate flash memory”
- [19] Andrea Ghetti, Salvatore Maria Amoroso, Aurelio Mauri, and Christian Monzio Compagnoni: “Impact of Nonuniform Doping on Random Telegraph Noise in Flash Memory Devices”
- [20] A. Ghetti, S. M. Amoroso, A. Mauri, and C. Monzio Compagnoni: “Doping engineering for random telegraph noise suppression in decanometer Flash memories”

- 
- [21] A. Mauri, C. Monzio Compagnoni, S.M. Amoroso, A. Maconi, A. Ghetti, A. S. Spinelli, and A. L. Lacaita: “Comprehensive investigation of statistical effects in nitride memories Part I: Physics-based modeling”
- [22] A. Ghetti, M. Bonanomi, C. Monzio Compagnoni, A. S. Spinelli, A. L. Lacaita, and A. Visconti: “Physical modeling of single-trap RTS statistical distribution in Flash memories”
- [23] Salvatore Maria Amoroso, Andrea Ghetti, Andrew R. Brown, Aurelio Mauri, Christian Monzio Compagnoni, and Asen Asenov: “Impact of Cell Shape on Random Telegraph Noise in Decanometer Flash Memories”
- [24] A. Asenov: “Random dopant induced threshold voltage lowering and fluctuations in sub- $0.1\mu\text{m}$  MOSFET’s: a 3-D Atomistic simulation study”
- [25] G. Torrente, N. Castellani, A. Ghetti, C. Monzio Compagnoni, A. L. Lacaita, A. S. Spinelli, A. Benvenuti: “Assessment of the Statistical Impedance Field Method for the Analysis of the RTN Amplitude in Nanoscale MOS Devices”
- [26] G. Torrente, N. Castellani, A. Ghetti, C. Monzio Compagnoni, A. L. Lacaita, A. S. Spinelli, A. Benvenuti: “Investigation of the RTN amplitude statistics of nanoscale MOS devices by the statistical impedance field method”
- [27] L. Bortesi, L. Vendrame, P. Fantini, A. Spessot, and A. L. Lacaita: “Characterization and Modeling Methodology for the Evaluation of Statistical Variation of MOSFETs”
- [28] Grazia Laricchiuta: “Caratterizzazione E Modeling Statistico di Transistori CMOS in Tecnologia 90nm ”
- [29] Synopsys, Zurich, Switzerland: Sentaurus device user guide, G.-2012.06 edn. (2012)