

POLITECNICO DI MILANO

V Facoltà di Ingegneria (Ingegneria dell'Informazione)

Corso di Laurea Specialistica in Ingegneria Elettronica



**Modello numerico elettrotermico per analisi di sensibilità e di scaling
delle memorie a cambiamento di fase**

Relatore: prof. Ing. Daniele Ielmini

Correlatore: dott. Andrea Ghetti

Tesi di: Giovanni Ferrari

Matricola: 721388

Anno Accademico 2009/2010

Indice

Introduzione	8
1 La cella di memoria PCM	11
1.1 Principi di funzionamento	11
1.2 Proprietà del GST	13
1.3 La cella di memoria Wall	32
1.3.1 Parametri caratteristici della cella e tecniche di programmazione .	35
1.3.2 Modello analitico della cella Wall	39
1.4 Scaling della cella PCM Wall	44
2 Dati sperimentali e calibrazione del modello	50
2.1 Il modello elettrotermico	51
2.1.1 Fase cristallina	51
2.1.2 Cambiamento di fase e fase amorfa	58
2.2 Calibrazione del modello mediante i dati sperimentali	61
2.2.1 La cella standard	64
2.2.2 Resistività dell'heater	65
2.2.3 Proprietà elettriche e termiche del GST	70
2.2.4 Resistività del GST	71
2.2.5 Conducibilità termica	76
2.2.6 Calibrazione delle interfacce termiche ed elettriche	78
2.3 Validazione del modello numerico	84

<i>INDICE</i>	2
3 Applicazioni del modello	89
3.1 Analisi di sensibilità	89
3.2 Stima delle prestazioni di scaling	96
3.3 Studio di geometrie alternative	99
3.4 Analisi della ricristallizzazione parziale	104
Conclusioni	112
Bibliografia	115

Elenco delle figure

1.1	Principio di funzionamento della cella PCM.	12
1.2	Curve I-V di una tipica cella PCM.	16
1.3	Struttura atomica e a bande del GST amorfo.	17
1.4	Caratteristiche I-V sottosoglia di celle PCM nello stato amorfo al variare della temperatura.	18
1.5	Rappresentazione degli stati trappola.	19
1.6	Curve I-V di sottosoglia del GST amorfo.	23
1.7	Andamento del rapporto trappole cariche/totali e portatori minoritari/totali.	26
1.8	Raffigurazione schematica delle barriere energetiche coinvolte nel meccanismo di crescita di un cluster x all'interno della fase l . Il cluster x aumenta di dimensioni non appena un monomero acquisisce energia sufficiente per superare la barriera E_{ag} . L'elevata barriera tra x e l rende il cammino inverso molto meno probabile.	31
1.9	Immagine TEM di un cella realizzata con architettura Wall (a) e uno schema semplificato della struttura (b).	33
1.10	Struttura a bande della cella Wall.	34
1.11	Immagine TEM di un cella realizzata in struttura Wall (a) e relativo schema semplificato (b) in cui si può notare la struttura che connette la cella all'array di memoria: il BJT verticale e il contatto di base.	35
1.12	Curva I-V e R-I di una tipica cella 45nm.	36
1.13	Forma degli impulsi di programmazione: RESET, SET minimo e SET.	38
1.14	Curve di programmazione ottenute con sequenze di impulsi differenti.	40
1.15	Modello resistivo della cella PCM.	41
1.16	Metodologia di ottimizzazione della corrente di programmazione.	44
1.17	Corrente di fusione in funzione di L_h , per diversi valori di R_{SET} fissata.	45

1.18	Prestazioni di scaling delle resistenze di ON e di SET (sinistra) e della corrente di fusione (destra) nei casi di scaling isotropo e anisotropo.	47
1.19	Andamento delle cifre di merito nei casi di scaling isotropo e anisotropo.	49
2.1	Capacità termica molare del $Ge_2S B_2Te_5$	53
2.2	Struttura della cella PCM usata per la simulazione. I materiali isolanti sono trasparenti.	55
2.3	Struttura usata nella simulazione.	56
2.4	Profilo di temperatura nella struttura simulata.	57
2.5	Temperatura massima raggiunta dall'interfaccia.	59
2.6	Dipendenza dello switching nel modello dalla ionizzazione a impatto.	62
2.7	Caratteristiche IV e RI di circa 200 celle standard, per i nodi 90nm e 45nm.	64
2.8	Caratteristiche IV ed RI dello split L_h	66
2.9	Valori di R_{SET} e R_{ON} al variare della lunghezza dell'heater (L_h).	67
2.10	Caratteristiche IV di circa 50 celle no-GST (blu).	70
2.11	Profilo di temperatura simulato della cella standard e no-GST lungo l'asse verticale del dispositivo.	71
2.12	Arrhenius plot della resistenza di SET di una singola cella 45nm.	73
2.13	Energia di attivazione della cella 45nm e del GST ricavate dalle simulazioni numeriche con una tensione applicata di 0.4 V.	74
2.14	Resistenza di SET in funzione dell'altezza dello strato di GST L_c per celle dei nodi 90nm e 45nm.	75
2.15	Caratteristiche IV e profilo di temperatura al variare di K_c	76
2.16	Curve RI al variare di K_c	77
2.17	Curve RI simulate in assenza e presenza di interfacce termiche.	79
2.18	Curve di programmazione RI simulate per diversi valori di conducibilità termica interfacciale. Si nota come l'interfaccia influisca sia sul valore di I_m , sia sulla pendenza della curva.	82
2.19	Illustrazione dell'andamento esponenziale della curva RI.	83
2.20	Misure di resistenza termica interfacciale prese presso Numonyx utilizzando la tecnica MPTR.	84

2.21	Validazione del modello numerico calibrato. In verde le curve IV e RI sperimentali delle celle standard 45nm e in nero le caratteristiche simulate. Ottimo fit dei dati con il set di parametri unico.	86
2.22	Curve IV per lo split L_h	86
2.23	Curve RI per lo split L_h	87
2.24	Cifre di merito in funzione di L_c	87
2.25	Validazione del modello numerico calibrato. Simulazione dei parametri R_{SET} , R_{ON} e I_m per celle con una diversa larghezza dell'area attiva (W).	88
3.1	Scatter plot delle resistenze di SET e ON in funzione della corrente di melting I_m	90
3.2	Scatter plot $R-I_m$ ricavato dal modello analitico inserendo uno spread gaussiano del 5% (heater) o del 10% (GST) su ogni grandezza.	91
3.3	Sensibilità dei parametri R_{SET} , R_{ON} e I_m alle variazioni di alcune proprietà fisiche e geometriche della cella, espresse in termini percentuali.	93
3.4	Scatter plot $R-I_m$ nei tre casi considerati.	94
3.5	Correlazione tra le resistenze di SET e di ON prima e dopo un ciclo di 10.000 operazioni di programmazione e cancellazione.	97
3.6	Sopra: caratteristiche IV ed RI simulate per le celle del nodo 32nm per alcune combinazioni di altezza di heater e GST. Sotto: riassunto dei parametri caratteristici delle celle simulate.	99
3.7	Scaling di R_{SET} e I_m come previsto dal modello analitico e dalla simulazione numerica, sovrapposto ai dati sperimentali per i nodi 90nm e 45nm.	100
3.8	Immagine TEM di una cella damascata (sinistra) e schema della struttura usato nella simulazione.	101
3.9	Sopra: caratteristiche IV ed RI delle celle damascate simulate. Sotto: riassunto dei parametri caratteristici.	102
3.10	Confronto tra le caratteristiche salienti della cella damascata e della cella Wall 32nm.	103
3.11	Immagini TEM della cella PCM 90nm in prossimità dell'interfaccia, per diverse tensioni di programmazione. È chiaramente visibile la calotta di materiale amorfo.	105
3.12	Misure sperimentali tratte dalle immagini TEM della dimensione della zona di GST fusa e della calotta di materiale amorfo. I dati si riferiscono a cella a 90nm.	106
3.13	Sinistra: schema del circuito usato per la simulazione in transitorio. Destra: andamento della velocità di cristallizzazione in funzione della temperatura.	107

3.14	Profilo di temperatura in vari istanti di tempo.	108
3.15	Risultati della simulazione della ricristallizzazione.	110
3.16	Profilo di fase alla fine del transitorio.	111

Abstract

Il lavoro di tesi verte sul modeling numerico delle memorie a cambiamento di fase (PCM). Queste sfruttano l'elevata differenza di resistività tra la fase cristallina e la fase amorfa dei materiali calcogenuri (in particolare, della lega Germanio-Antimonio-Tellurio GST) per immagazzinare il dato, letto con un semplice sense di corrente. Il cambiamento di fase è ottenuto mediante riscaldamento, fino a fusione, del GST, seguito da un raffreddamento, più o meno rapido a seconda del dato che si vuole scrivere. I fenomeni in gioco sono di natura sia elettrica, sia termica, fortemente interagenti tra loro. Per questa ragione un approccio numerico è necessario per ottenere una descrizione accurata del dispositivo. A questo scopo viene introdotto un modello numerico basato sulla descrizione del calcogenuro come un semiconduttore dotato di un elevato numero di stati trappola all'interno del gap energetico. Il modello viene calibrato estraendo i parametri descrittivi dei materiali da una grande quantità di dati sperimentali, e si dimostra che esso è in grado di riprodurre correttamente le curve caratteristiche di celle appartenenti a due nodi tecnologici differenti, 90nm e 45nm. Il modello è quindi utilizzato come strumento per l'analisi di alcune tematiche relative alle PCM: la sensibilità delle cifre di merito della cella alle variazioni statistiche della geometria e delle proprietà fisiche dei materiali; le prestazioni di scaling dei nodi successivi al 45nm; lo studio di una geometria alternativa a quella attualmente in produzione presso Numonyx, che mette in luce il contributo fondamentale del confinamento termico alle prestazioni della cella; uno studio preliminare del fenomeno di ricristallizzazione parziale del GST durante l'operazione di quenching.

Introduzione

Il presente lavoro di tesi specialistica è stato svolto presso il centro di ricerca e sviluppo R2 di Numonyx di Agrate Brianza, e ha riguardato lo studio e il perfezionamento di un modello numerico elettrotermico per le memorie a cambiamento di fase (PCM), e le sue successive applicazioni. Queste memorie sfruttano la proprietà delle leghe *calcogenure* (cioè in cui almeno uno degli elementi che la compongono è un calcogene, come il Tellurio) di passare da uno stato solido cristallino a uno stato amorfo mediante fusione seguita da un rapido *quenching*. Poiché le due fasi sono caratterizzate da resistenze elettriche diverse tra loro di svariati ordini di grandezza, la lettura del dato a temperatura ambiente avviene attraverso una semplice lettura della corrente. Le due configurazioni sono entrambe stabili, per cui la memoria è intrinsecamente non volatile. Lo sviluppo di queste memorie ha portato all'utilizzo del $\text{Ge}_2\text{Sb}_2\text{Te}_5$ come materiale attivo, caratterizzato da tempi di cristallizzazione rapidi e temperatura di fusione relativamente ridotta. Queste caratteristiche, unite a un'eccellente scalabilità, permettono alle memorie PCM di mettersi in competizione con gli esistenti dispositivi FLASH e DRAM.

A differenza dei dispositivi elettronici tradizionali, le memorie PCM per il loro funzionamento intrinseco sfruttano fenomeni non solo strettamente elettrici, ma anche termici. La necessità di fondere il materiale attivo per cambiare la programmazione della cella fa sì che l'operazione della memoria avvenga in un intervallo di temperature molto ampio, per cui gli effetti dovuti alla temperatura (variazione di resistività, conducibilità elettrica, ecc) hanno un'influenza molto pesante sulle prestazioni del dispositivo. È anche necessario lo studio di fenomeni fisici normalmente estranei all'universo dei dispositivi elettronici: la fusione, la cinetica di cristallizzazione e lo *switching ovonico*, un fenomeno tipico dei calcogenuri che consente l'iniezione di alte correnti anche in un materiale amorfo

(e quindi caratterizzato da un'elevatissima resistenza elettrica) evitando l'applicazione di campi elettrici eccessivamente elevati. La loro analisi è tuttavia molto complicata: poiché sostanzialmente ogni grandezza fisica in gioco ha una sua dipendenza dalla temperatura, un loro studio quantitativo è pressoché impossibile per via analitica, rendendo necessario un approccio numerico al problema. Per questa ragione è stata sviluppata un'estensione del simulatore di dispositivi SDevice della Synopsys per considerare tutti i fenomeni che avvengono durante il funzionamento, compreso lo switching e il cambiamento di fase [1]. Perché questo modello sia in grado di riprodurre quantitativamente le caratteristiche delle PCM è necessario effettuare una calibrazione delle numerose costanti che compaiono nelle equazioni in quanto non direttamente accessibili. Questa calibrazione sarà il fulcro del presente lavoro di tesi. Per effettuarla saranno analizzati numerosi dati sperimentali, ottenuti sia dalla letteratura, sia dai laboratori Numonyx. Lo scopo ultimo sarà ricavare un unico insieme di parametri che riproduca fedelmente il maggior numero di dati sperimentali possibile, per diversi nodi tecnologici, geometrie e materiali. Una volta dimostrato che il modello numerico è sufficientemente predittivo del comportamento fisico reale della cella, esso può essere utilizzato come utile strumento per l'analisi di geometrie alternative, celle scalate e altre problematiche.

Nel Capitolo 1 si descriveranno in dettaglio le proprietà dei materiali calcogenuri, il funzionamento di base della cella di memoria, si definirà la terminologia usata in tutto il resto del lavoro e si illustrerà il modello analitico comunemente utilizzato per descrivere in maniera sintetica il dispositivo PCM. Saranno messe in evidenza le discrepanze tra il modello analitico e i dati sperimentali, dovute all'interazione tra fenomeni termici ed elettrici non facilmente prevedibili.

Nel Capitolo 2 si illustrerà il modello numerico a elementi finiti che sta alla base delle simulazioni della cella. Successivamente si mostrerà il lavoro di calibrazione, basato sull'estrapolazione di parametri chiave da una serie di dati sperimentali e dalla letteratura. Una volta ottenuto il set di parametri che riesce a descrivere al meglio il comportamento sperimentali si provvederà a validare il modello confrontandolo con altri esperimenti, non utilizzati nel corso della calibrazione.

Nel Capitolo 3 si mostreranno alcuni esempi di utili applicazioni del modello calibrato: l'analisi di una cella scalata al nodo 32nm, utile per prevedere le prestazioni future delle

memorie PCM; un'analisi di sensibilità dei parametri chiave della cella alle variazioni delle proprietà geometriche e fisiche dei materiali, al fine di individuarne i contributi più importanti e riprodurre le distribuzioni statistiche sperimentali; uno studio di una cella con una geometria differente da quella prodotta da Numonyx, per evidenziare l'importanza di alcuni fattori, come il confinamento del calore e il self-heating, durante il funzionamento della cella.

Capitolo 1

La cella di memoria PCM

1.1 Principi di funzionamento

Nelle memorie a cambiamento di fase, l'informazione è contenuta nello stato fisico del materiale attivo, che può presentarsi in una fase amorfa o cristallina. La fase amorfa è caratterizzata da una resistività di ordini di grandezza più grande rispetto a quella del cristallino, per cui la lettura del dato memorizzato può essere effettuata mediante un sensing della resistenza del dispositivo. Le due fasi sono stabili a temperatura ambiente, ovvero mantengono il loro stato per un tempo indefinito in assenza di sollecitazioni esterne, per cui la memoria è intrinsecamente non volatile. Allo stato cristallino viene associato lo stato "SET" o "1" mentre allo stato amorfo si associa lo stato "RESET" o "0" dove, in analogia con gli switch elettronici, nominalmente lo stato "1" presenta una bassa resistenza mentre lo stato "0" impedisce il passaggio di corrente.

A differenza di altre memorie elettroniche, in cui i fenomeni in gioco sono principalmente di natura elettrica, i meccanismi di funzionamento della PCM vanno ricercati nell'interazione di effetti elettrici e termici. Infatti, il cambiamento di fase viene ottenuto portando il materiale attivo a una temperatura superiore a quella di fusione, e quindi allo stato liquido. La velocità con cui viene raffreddato (*quenching*) è cruciale per il funzionamento della cella: infatti se il raffreddamento avviene in tempi brevi il materiale non ha il tempo di formare un reticolo cristallino, e si presenterà nello stato amorfo a temperatura ambiente.

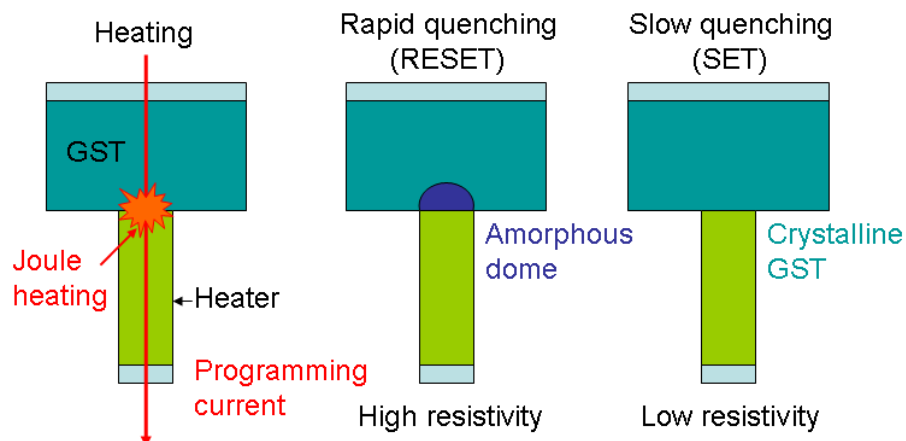


Figura 1.1: Principio di funzionamento della cella di memoria PCM. a) Si fa scorrere una corrente nella cella, che provoca dissipazione di potenza e riscaldamento per effetto Joule, fino alla fusione del materiale attivo. b) In caso di raffreddamento rapido, il materiale attivo non ha il tempo di formare i legami cristallini e si ritrova in uno stato amorfo stabile (RESET). c) In caso di raffreddamento lento, si viene a riformare un reticolo, in modo da portare la cella a temperatura ambiente nello stato cristallino (SET).

A questa temperatura gli atomi del calcogenuro non possiedono abbastanza energia termica per attivare i fenomeni che portano alla cristallizzazione, per cui il materiale rimane stabilmente nella fase amorfa. Se invece il quenching avviene più lentamente gli atomi dispongono di un tempo sufficiente a creare i legami reticolari, e alla fine del transitorio la cella si troverà nello stato cristallino. Una volta programmata, si può andare a leggere il dato applicando una tensione ai suoi capi e leggendone la corrente di uscita, che dipenderà dal dato memorizzato. Il riscaldamento della cella avviene per effetto Joule, o dissipando potenza su una resistenza in serie al materiale attivo, detta riscaldatore o *heater*, oppure direttamente per self-heating.

I materiali attivi più utilizzati in questo campo sono i calcogenuri, formati da leghe di elementi di cui almeno uno è un calcogene, ovvero appartenente al VI gruppo della tavola periodica. Queste leghe godono di importanti proprietà che ne rendono possibile l'uso nelle PCM, come un'elevata velocità di cristallizzazione e una buona stabilità termica. In essi inoltre si manifesta il fenomeno dello *switching ovonico*, già documentato nelle prime ricerche in materia da parte di Ovshinsky negli anni '60: un calcogenuro che si trova nella

fase amorfa è caratterizzato, come già detto, da un'elevata resistività; tuttavia, se la tensione applicata è tale da superare una certa soglia V_{th} , improvvisamente la sua resistenza crolla, tanto da essere trascurabile rispetto a quella dei contatti esterni, e la caratteristica I-V presenta un tratto a pendenza negativa. Questo fenomeno consente di portare una cella che si trova inizialmente nello stato amorfo alla temperatura di fusione senza dover applicare campi elettrici estremamente elevati[2].

I calcogenuri più diffusi in questo ambito sono formati da leghe di Germanio (Ge), Antimonio (Sb) e Tellurio (Te), l'elemento calcogene. La loro composizione può essere ingegnerizzata per ottenere le caratteristiche desiderate, in particolare la composizione chimica ha effetto sulla temperatura di fusione, la temperatura di cristallizzazione, l'energia di attivazione e la resistività, tutti parametri che entrano in gioco durante il funzionamento del dispositivo. La lega più utilizzata al momento è il $\text{Ge}_2\text{Sb}_2\text{Te}_5$, abbreviato in GST. Questo materiale è già da tempo impiegato nella fabbricazione di memorie ottiche a cambiamento di fase (ad esempio la tecnologia CD-ROM), in cui si sfruttano le differenti proprietà ottiche delle due fasi (come l'indice di rifrazione) per leggere il dato mediante un laser. Il suo utilizzo nel campo delle memorie elettriche è invece ancora in una fase sperimentale avanzata. Nei paragrafi successivi si descriveranno le principali proprietà del GST.

1.2 Proprietà del GST

Fase cristallina

Lo stato cristallino può essere raggiunto a partire sia dallo stato liquido, raffreddando lentamente la cella, in modo da consentire il formarsi dei legami reticolari, sia dallo stato amorfo, portando il GST sopra la sua temperatura di cristallizzazione. Nello stato cristallino il calcogenuro può presentarsi in una fase esagonale (hex) o in una fase cubica a facce centrate (fcc) [22]. Il tempo di cristallizzazione, per temperature nell'intorno di quella di glass, della fase esagonale (decine di nanosecondi) è superiore a quello della fase fcc (nanosecondi), per cui nello stato di SET il materiale si troverà principalmente nella fase cubica, che sarà pertanto presa maggiormente in considerazione. La struttura della fase fcc è di tipo policristallino, tuttavia, vista l'elevata dimensione dei cristalli (μm), è possibile

parlare di ordine di lungo raggio, e applicare il teorema di Bloch. Sotto questa ipotesi, il GST si comporta come un semiconduttore, e le sue caratteristiche elettriche sono definite dalla sua struttura a bande: presenterà quindi una banda di valenza, una banda di conduzione e un gap energetico tra le due bande, quest'ultimo pari a 0.5eV, valore ricavato da misure di assorbimento ottico. Tuttavia, il reticolo presenta un numero elevato di difetti cristallografici, in particolare la mancanza di un atomo di antimonio o germanio nel legame con il calcogene. Questo difetto prende il nome di vacanza strutturale, ed è molto importante per comprendere il funzionamento del composto, in quanto è presente in circa il 20% dei legami. Poiché gli atomi di tellurio presentano dei legami non saturati, questi ultimi si comportano come centri di cattura per gli elettroni, e possono essere descritti nell'ambito della struttura a bande come uno stato trappola di tipo accettore, posto nei pressi della banda di valenza, la cui densità è stata calcolata in base al numero di legami tra gli atomi, e vale $N_t = 5 \cdot 10^{20} \text{cm}^{-3}$. Questa configurazione energetica fa sì che il GST abbia una conduzione intrinsecamente di tipo p, come documentato in letteratura. L'esatta distanza del livello trappole dalla banda (indicata con E_{traps}) determina la posizione del livello di Fermi, che si collocherà o tra la banda e il livello accettore o all'interno della banda di valenza, a seconda del valore di E_{traps} . Il calcogenuro avrà quindi un comportamento degenere o quasi-degenere, caratterizzato da una bassa energia di attivazione (nell'ordine della decina di meV). Infatti, la conducibilità è data dalla relazione

$$\sigma = \sigma_0 e^{-\frac{E_a}{kT}} \quad (1.1)$$

in cui E_a è l'energia di attivazione della conducibilità, e dipende appunto dalla posizione del livello di Fermi rispetto al massimo della banda di valenza. La conducibilità in questa fase è molto più alta rispetto alla fase amorfa, ed è nell'ordine di qualche decina di $\Omega^{-1} \text{cm}^{-1}$. Nella figura 1.2, che riporta la caratteristica IV di una cella PCM, si osserva un tratto a bassi campi caratterizzato da una resistenza costante propria del GST, e un tratto a campi elevati in cui la resistenza si abbassa. Questo comportamento è dovuto appunto all'attivazione della conducibilità provocato dalle alte temperature che si raggiungono all'interno del dispositivo, che provoca la saturazione della resistenza a un valore imposto dai contatti esterni.

Per quanto riguarda le proprietà termiche, il GST cristallino è caratterizzato da una conducibilità termica (K_c) nell'ordine del mW/cmK e una capacità termica (C_{th}) specifica di circa 1.2 J/Kcm³. Queste specifiche portano a un tempo di risposta termico assai ridotto, caratteristica fondamentale per la realizzabilità della memoria, in quanto determina la durata minima dell'impulso di programmazione. La conducibilità termica durante l'operazione tende a essere più elevata di quella riportata, in quanto essa deriva da due contributi distinti: quello fononico (K_c^{ph}) e quello elettronico (K_c^e). Il contributo fononico tiene conto delle caratteristiche vibrazionali degli atomi del reticolo ed è tipica del materiale. Il contributo elettronico invece tiene conto del calore trasmesso dagli elettroni nel corso degli urti che subiscono tra di loro e con il reticolo, e dipende dalla concentrazione di portatori liberi secondo la legge empirica di Wiedemann-Franz, esprimibile nella forma

$$K_c^e = \Lambda \sigma T \quad (1.2)$$

in cui Λ è una costante. Da questa relazione deriva l'aumento della conducibilità termica: infatti nell'area attiva la temperatura aumenta notevolmente, così come la concentrazione di portatori generati termicamente.

Fase amorfa

La fase amorfa del GST presenta, per sua natura, una composizione atomica disordinata. La mancanza di un ordine di lungo raggio impedirebbe la rappresentazione mediante un diagramma a bande, tuttavia è stato sperimentalmente osservato che è presente un ordine di corto raggio. In virtù di questa caratteristica, è comunque possibile applicare il teorema di Bloch e tracciare la struttura energetica del calcogenuro amorfo, trattandolo a tutti gli effetti come un semiconduttore molto difettoso. La figura 1.3.a rappresenta la struttura atomica del GST amorfo. Il numero di coordinazione degli atomi è pari alla loro valenza: 4 per il Germanio, 3 per l'Antimonio e 2 per il Tellurio. Gli atomi di Tellurio formano delle catene, e in questa configurazione fondamentale rimangono due elettroni nell'orbita più esterna: essi prendono il nome di lone-pairs. In questa situazione l'atomo calcogene non è ionizzato e ha un numero di coordinazione pari a 2, e viene perciò identificato con

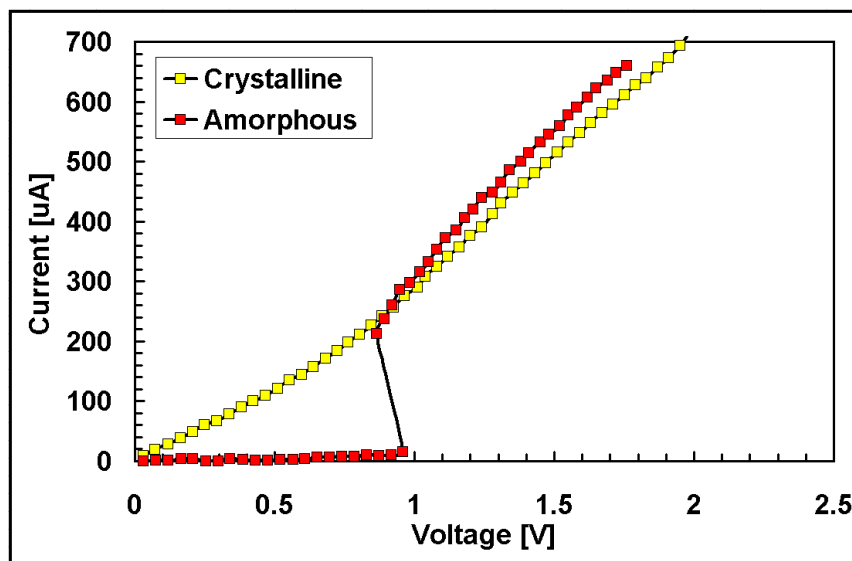


Figura 1.2: Curve I-V di una tipica cella PCM 90nm nelle due fasi operative. Nello stato cristallino la resistenza ad alti campi è dominata dai contatti. Nello stato amorfo si può notare il tipico snap-back della caratteristica dovuto allo switching di soglia.

C_2^0 . A temperatura finita il GST amorfo, come tutti i vetri, presenta dei difetti. Essi si presentano come atomi di Tellurio con numero di coordinazione 1 carichi negativamente, e con coordinazione 3 carichi positivamente, indicati rispettivamente con C_1^- e C_3^+ . Questi difetti hanno una concentrazione che varia da 10^{17} a 10^{20} cm^{-3} . In base a queste considerazioni è possibile tracciare un diagramma a bande anche per il materiale in fase amorfa. Misure ottiche hanno mostrato che il gap energetico è pari a 0.7eV, e che l'assorbimento presenta una coda alle basse energie, dovuta alla presenza dei lone-pairs. Questa viene rappresentata come una coda energetica vicino al massimo della banda di valenza, detta coda di Urbach, i cui stati elettronici sono caratterizzati da una bassa mobilità, dato che sono fortemente localizzati. I difetti nella struttura sono invece descritti con stati trappola di tipo donore per gli atomi C_3^+ e di tipo accettore per gli atomi C_1^- : infatti, essendo elettricamente carichi, hanno la possibilità di catturare gli elettroni e le lacune. La struttura a bande delle due fasi è riportata in figura 1.3.b[3].

In una lega calcogenura nello stato amorfo avviene un fenomeno elettrico di primaria importanza, che consente l'utilizzo pratico delle memorie a cambiamento di fase: lo *switch-*

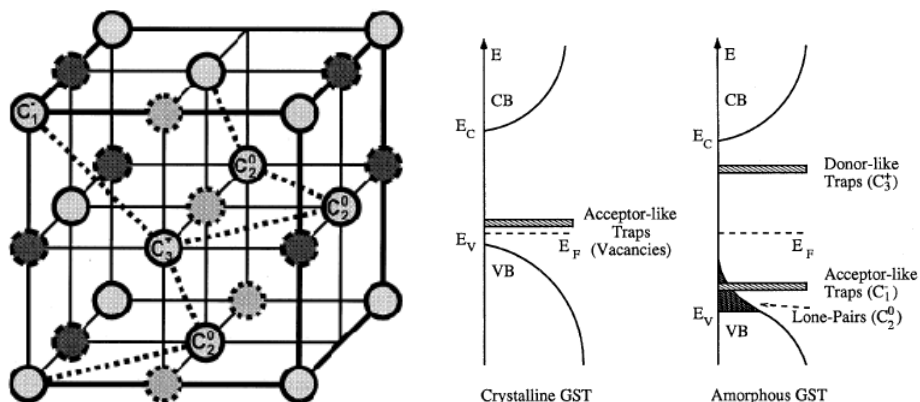


Figura 1.3: Sinistra: struttura atomica del GST amorfo disegnata su una geometrica cubica. In realtà la struttura è deformata. La linea tratteggiata raffigura la catena di atomi di Tellurio, che possono presentarsi o nello stato fondamentale (C_2^0) o come difetti con diversa coordinazione (C_1^- o C_3^+). Destra: struttura a bande del GST nello stato cristallino (sinistra) e amorfo (destra). Le caratteristiche principali del materiale sono: conducibilità di tipo p, presenza di stati trappola accettori nel cristallino dovuti alle vacanze strutturali, presenza di stati trappola accettori e donori nell'amorfo per descriverne i difetti strutturali, presenza di una coda di stati localizzati nella banda di valenza dovuta agli elettroni lone-pairs.

ing di soglia. Lo stato amorfo presenta una resistività di svariati ordini di grandezza più elevata rispetto a quello cristallino. Se si vuole scrivere un dato in una cella inizialmente cancellata, è necessario scaldarla mediante l'iniezione di corrente, che provoca l'emissione di calore per via dell'effetto Joule. Per ottenere una dissipazione di potenza tale da far superare al GST la temperatura di cristallizzazione, sarebbe necessario applicare una tensione elevatissima, a causa dell'enorme resistenza dell'amorfo, e di conseguenza si avrebbero campi elettrici molto alti, che provocano problemi nella circuiteria di selezione della cella di memoria (effetti di canale corto, iniezione di carica nell'ossido, ecc.). Invece, non appena la tensione che cade sul GST supera un certo valore, detto tensione di soglia V_{th} , la corrente aumenta improvvisamente con uno snap-back di tensione, pur restando il materiale nella fase amorfa. Dopo che è avvenuto lo switching è possibile raggiungere la potenza necessaria alla programmazione evitando l'utilizzo di tensioni eccessive: infatti la resistenza del GST amorfo switchato è molto bassa, e viene dominata dai contatti esterni in serie alla cella. Questo fondamentale fenomeno verrà analizzato dettagliatamente in

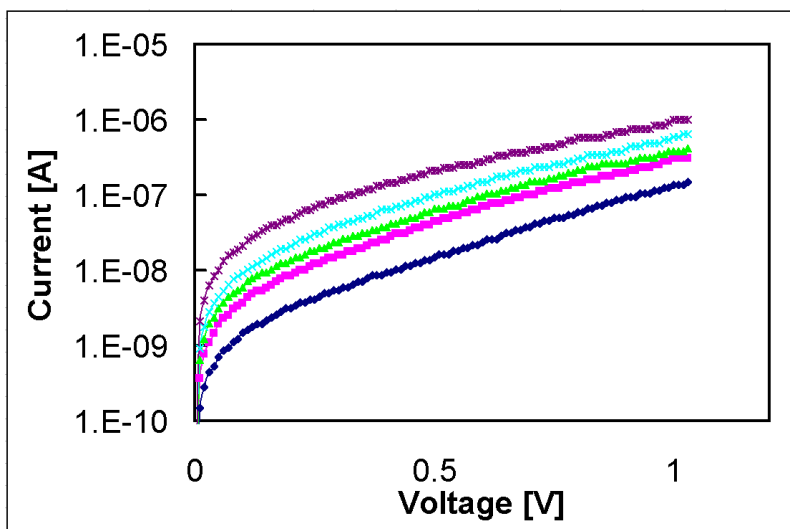


Figura 1.4: Caratteristiche I-V sottosoglia di celle PCM nello stato amorfo al variare della temperatura.

seguito.

Il modello appena descritto prende il nome di **modello di Adler**[4], dal nome del suo primo propositore. Partendo da questa struttura a bande è possibile realizzare un modello numerico che tenga conto delle principali caratteristiche di conduzione del GST nello stato vetroso, compreso lo switching: verrà descritto dettagliatamente nel capitolo 2. Tuttavia è anche necessaria una descrizione di tipo analitico che renda immediatamente comprensibili i meccanismi fisici che rendono possibile il trasporto elettronico in un materiale isolante.

La conduzione nell'amorfo

La figura 1.4 mostra in dettaglio la caratteristica I-V del calcogenuro amorfo per tensioni tali da non provocare lo switching, per diversi valori di temperatura. Questa regione è detta *regione di sottosoglia*. Si può osservare che per tensioni relativamente basse (inferiori a 0.4 V) la corrente aumenta linearmente con la tensione, mentre per tensioni più elevate si ha un andamento esponenziale. Inoltre la corrente aumenta sensibilmente con la temperatura; è possibile estrarre l'energia di attivazione della conduzione da un grafico di tipo

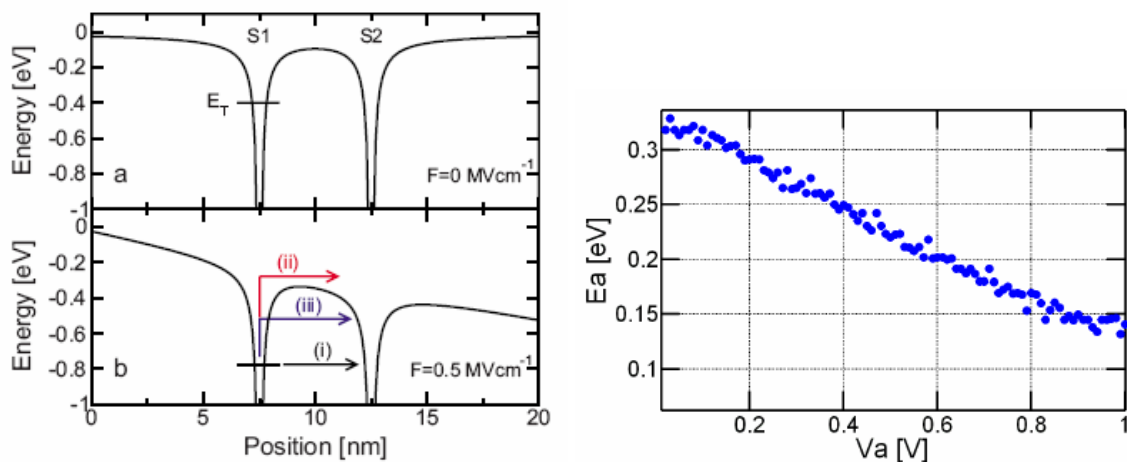


Figura 1.5: Sinistra: profilo energetico di due stati trappola S1 e S2 caratterizzati da un profilo coulombiano senza campo elettrico applicato (a) e con campo elettrico applicato (b). Vengono anche mostrati i meccanismi di trasporto possibili: tunneling (i), emissione termica (ii) e tunneling assistito termicamente (iii). Destra: andamento sperimentale dell'energia di attivazione in funzione della tensione applicata.

Arrhenius, che mostra la corrente letta in funzione di $1/kT$ per diverse tensioni di lettura. La pendenza di queste curve rappresenta l'energia di attivazione E_a . Dai dati sperimentali si osserva che essa vale circa 0.3eV e che diminuisce all'aumentare della tensione applicata. Il valore numerico di E_a , che riflette la distanza del livello di Fermi dalla banda di conduzione, conferma che il GST vetroso ha un comportamento di tipo intrinseco, infatti risulta essere circa metà del gap ottico misurato (0.7 eV). L'andamento della caratteristica I-V può essere riprodotto mediante il modello di Adler, inserendo nel calcolo del livello di Fermi e del potenziale gli stati trappola descritti in precedenza. Operativamente, questo approccio verrà utilizzato nel corso di questo lavoro per simulare il comportamento elettrico della cella nello stato di RESET e ricavarne la corrente di lettura per bassi campi elettrici applicati (circa 0.4V). Tuttavia la presenza degli stati trappola nel gap energetico rende pressoché impossibile ricavare una forma analitica chiusa per la corrente che scorre nella cella, ed è quindi difficile avere un'idea dei meccanismi fisici della conduzione.

Il calare dell'energia di attivazione all'aumentare della tensione applicata suggerisce anche un'interpretazione alternativa al modello di Adler, in cui la conduzione è controllata

dai portatori intrappolati[5]. Si prenda come riferimento la figura 1.5, che mostra il profilo di potenziale, che si ipotizza essere coulombiano, associato a due trappole di tipo donore (stati possibili: 0/+) situate nel gap energetico del materiale amorfo. Per passare dallo stato S1 allo stato S2 l'elettrone deve superare una barriera di potenziale di altezza $E_C - E_T$, dove E_T è l'energia dello stato trappola. E_C è stato preso come il massimo della barriera tra S1 e S2. L'elettrone può essere trasmesso oltre la barriera grazie a un fenomeno di tunneling (i), di emissione termica (ii) o di tunneling assistito termicamente (iii). L'andamento dell'energia di attivazione fa pensare che il trasporto sia dominato dai fenomeni di tipo (ii): infatti quando viene applicato un campo elettrico F tale da non permettere lo switching del dispositivo ($F > 0.4 \text{ MVcm}^{-1}$) la forma del profilo di potenziale varia, e l'altezza della barriera diminuisce. Dato che l'energia di attivazione della conduzione nel caso di emissione termica coincide proprio con l'altezza della barriera, l'andamento sperimentale di E_a può essere correttamente descritto da un modello di questo tipo. La conduzione di corrente avviene quindi per emissioni successive di portatori da una trappola a quella seguente. Questo meccanismo prende il nome di *conduzione per hopping*.

Il trasporto per emissione termica viene tipicamente descritto dal modello Poole-Frenkel, che prevede una dipendenza esponenziale della corrente dalla tensione nella forma

$$I = I_{PF} e^{\beta_{PF} \sqrt{V_a}} \quad (1.3)$$

in cui I_{PF} e β_{PF} sono costanti. Questo modello tuttavia assume come ipotesi iniziale che l'emissione avvenga da una trappola isolata alla banda di conduzione, e trascura l'influenza delle trappole adiacenti, poiché sono sufficientemente lontane da poter essere considerate indipendenti ai fini del calcolo del profilo di potenziale. Questa condizione potrebbe non essere verificata nel caso del GST, in cui la densità di trappole è molto elevata. Grazie a delle simulazioni numeriche si può osservare come l'abbassamento di barriera tenda a essere linearmente dipendente dalla tensione applicata quando le trappole sono molto vicine tra loro ($\Delta z < 5 \text{ nm}$), e non più dipendente dalla radice di V . Il modello di trasporto diventa quindi

$$I = I_P e^{\beta_P V_a} \quad (1.4)$$

detto *modello Poole*. L'abbassamento lineare della barriera di potenziale può essere visto come la conseguenza del fatto che la posizione del massimo di potenziale tra le due trappole è circa costante. Con questa approssimazione l'abbassamento di barriera può essere scritto nella forma

$$\Delta U \approx -qV_a \frac{\Delta z}{2u_a} \quad (1.5)$$

Si può quindi ora enunciare un semplice modello analitico monodimensionale per la corrente di sottosoglia. La velocità media di un elettrone si può scrivere come la distanza media tra due trappole adiacenti divisa per il tempo medio di trasferimento tra i due stati, modificato dalla presenza del campo elettrico. Il tempo di trasferimento lungo la direzione del campo elettrico (*tempo di forward*), si scrive

$$\tau_{\rightarrow} = \tau_0 e^{(E_C - E_T - \Delta z)/kT} \quad (1.6)$$

dove τ_0 è il tempo di fuga caratteristico per l'elettrone intrappolato. L'esponente contiene l'altezza di barriera all'equilibrio ($E_C - E_T$) corretto per l'abbassamento dovuto al campo elettrico. D'ora in avanti si farà riferimento a un abbassamento di tipo Poole, ovvero lineare con la tensione applicata. La corrente nella direzione del campo (*corrente di forward*) si scrive quindi

$$dI_{\rightarrow} = qAn_T(E_T) \frac{\Delta z}{\tau_0} e^{-[E_C - E_T - qV_a(\Delta z/2u_a)]/kT} dE_T \quad (1.7)$$

dove A è l'area del dispositivo e $n_T(E_T)dE_T$ è la concentrazione di elettroni che si trovano nelle trappole comprese tra E_T e $E_T + dE_T$. Questa espressione è in grado di spiegare l'andamento esponenziale della curva I-V di figura 1.4, però non tiene conto della zona lineare a basse tensioni. Per spiegare questo fenomeno è necessario osservare che esiste una probabilità non nulla che l'elettrone venga emesso termicamente da una buca a quella precedente, ovvero in direzione opposta al campo elettrico. Per bassi campi applicati questa probabilità potrebbe non essere trascurabile, mentre deve esattamente bilanciare la possibilità di emissione in avanti quando la tensione esterna è nulla. Questa corrente inversa si può scrivere in una forma simile alla 1.7,

$$dI_{\leftarrow} = qAn_T(E_T) \frac{\Delta z}{\tau_0} e^{-[E_C - E_T + qV_a(\Delta z/2u_a)]/kT} dE_T \quad (1.8)$$

dove la modifica all'altezza della barriera è di segno opposto rispetto al caso della corrente di forward. La corrente netta si ottiene dalla somma dei due contributi:

$$dI = 2qAN_T(E_T) \frac{\Delta z}{\tau_0} e^{-(E_C - E_F)/kT} \sinh\left(\frac{qV_a}{kT} \frac{\Delta z}{2u_a}\right) dE_T \quad (1.9)$$

In questa equazione la concentrazione di portatori n_T è stata sostituita da

$$n_T(E_T) = N_T(E_T) e^{-(E_T - E_F)/kT} \quad (1.10)$$

che rappresenta la concentrazione di trappole situate a un livello energetico E_T moltiplicata per l'approssimazione di Maxwell-Boltzmann della distribuzione di Fermi. La dipendenza della corrente da E_T si trova solamente all'interno dell'espressione della densità di portatori intrappolati, e non influisce direttamente sul tempo di emissione degli elettroni. Questo è dovuto al perfetto bilanciarsi della probabilità di avere un portatore su un dato livello energetico, espressa dalla 1.10, e della probabilità di emissione, data dall'inverso della 1.6. Quanto appena detto è da ritenersi valido finché il livello di trappole è sufficientemente distante sia dal livello di Fermi, sia dalla banda di conduzione. In questo caso infatti è valida l'approssimazione di Maxwell-Boltzmann, ed è quindi semplice integrare la 1.9 e ottenere la corrente totale:

$$I = 2qAN_{tot} \frac{\Delta z}{\tau_0} e^{-(E_C - E_F)/kT} \sinh\left(\frac{qV_a}{kT} \frac{\Delta z}{2u_a}\right) \quad (1.11)$$

N_{tot} è il risultato dell'integrale della distribuzione di trappole nel gap al di sopra del livello di Fermi. Poiché questa distribuzione non è conosciuta con precisione, il termine N_{tot} verrà trattato come una densità di stati equivalenti, che include tutte le correzioni necessarie per compensare le incertezze sui valori numerici precisi. L'equazione 1.11 rappresenta quindi il modello finale per la conduzione nell'amorfo. Per ottenere l'espressione della corrente nella regione lineare si sostituisce al seno iperbolico la sua approssimazione lineare, che risulta essere

$$I \approx \frac{q^2 AN_{tot} \Delta z^2}{kT \tau_0 u_a} e^{-(E_C - E_F)/kT} V_a \quad (1.12)$$

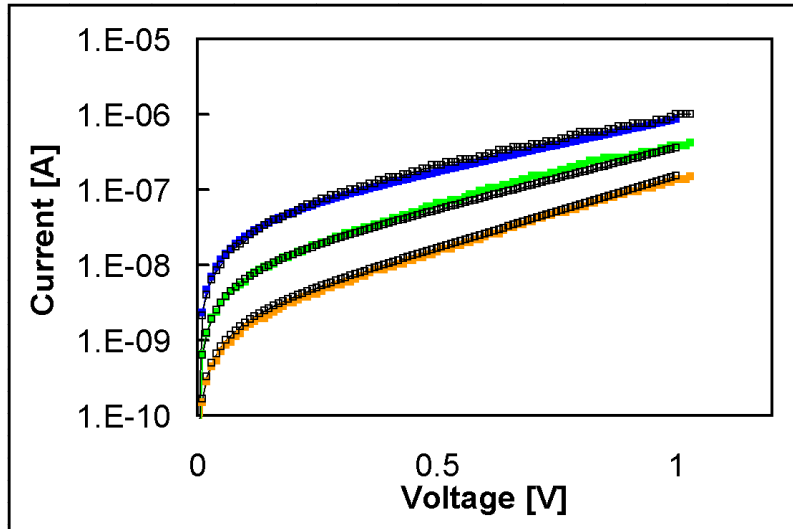


Figura 1.6: Curve I-V sottosoglia per il GST amorfo prese a diverse temperature. Il modello analitico è sovrapposto in nero.

mentre per quanto riguarda la zona esponenziale basta trascurare la componente di corrente inversa, per cui $I \approx I_{\rightarrow}$.

Questo modello è molto utile per estrarre alcuni parametri di interesse a partire dalla caratteristica I-V sottosoglia, come lo spessore della zona amorfa, ad esempio osservando la pendenza di sottosoglia della curva, definita come

$$STS = \frac{d \log I_{\rightarrow}}{dV_a} = \frac{q}{kT} \frac{\Delta z}{2u_a} \quad (1.13)$$

La figura 1.6 mostra le curve sperimentali mostrate in precedenza a cui viene sovrapposto il modello analitico, con i seguenti valori numerici: $\Delta z = 5nm$, $u_a = 22nm$, $E_C - E_F = 0.27eV$, $\tau_0 = 100fs$.

Lo switching di soglia

Il meccanismo di switching di soglia ovonico (*ovonic threshold switching*, OTS) è un fenomeno caratteristico dei vetri calcogenuri ed è fondamentale per le memorie a cambiamento di fase, in quanto consente di avere un'alta densità di corrente nell'area attiva nei

dispositivi in stato di RESET senza sviluppare campi elettrici enormi. Nel corso degli ultimi 40 anni sono state elaborate diverse teorie sull'origine dell'OTS. Inizialmente si pensava che lo switching fosse un meccanismo governato principalmente da fenomeni termici: la tensione applicata fa aumentare la densità di corrente e quindi il calore prodotto per effetto Joule, e di conseguenza la concentrazione di portatori liberi generati termicamente. Questo porta a un aumento della conducibilità termica, e quando la tensione esterna è sufficientemente elevata si innesca una reazione positiva che porta a un enorme aumento dei portatori. A questo punto la tensione sulla cella può calare mantenendo comunque un'elevata densità di corrente, ottenendo curva I-V a forma di S (*S-shaped negative differential resistance*). Degli esperimenti condotti da Adler negli anni '80 hanno però dimostrato che lo switching è riconducibile a fenomeni di natura puramente elettrica, legati all'equilibrio tra generazione e intrappolamento di portatori. Di questo modello non esiste una forma analitica, ma è possibile un'implementazione numerica. Il modello originale di Adler prevede che il calcogenuro nello stato amorfo sia descritto da un semiconduttore omogeneo drogato con un singolo livello di trappole vicino alla banda di conduzione e il livello del drogante vicino alla banda di valenza. La densità del drogante N_a deve essere maggiore della concentrazione delle trappole C_3^{tot} per replicare correttamente il trasporto di tipo p osservato negli esperimenti. Le trappole rappresentano gli stati VAP carichi positivamente, mentre vengono ignorati gli altri stati trappola descritti nel modello a bande, in quanto non strettamente necessari per spiegare il meccanismo OTS. In base a questa approssimazione, il modello è applicabile a tutti i sistemi caratterizzati da difetti che si comportano in questo modo, e non solo ai materiali calcogenuri.

All'equilibrio termico il livello di Fermi si colloca leggermente sotto il centro del gap energetico, per cui la concentrazione di trappole ionizzate C_3^+ è pari a quella totale. Applicando una tensione esterna la concentrazione di portatori aumenta per via della generazione, e le trappole iniziano a catturare elettroni, neutralizzandosi e diventando centri C_3^0 . Viene trascurata la possibilità che una trappola carica positivamente catturi una lacuna e che una trappola neutra catturi un elettrone. Nello stato stazionario, il potenziale Φ e la concentrazione di portatori n e p possono essere calcolati risolvendo accoppiate le equazioni di

continuità della corrente e di Poisson:

$$\frac{dn}{dt} = \frac{1}{q} \nabla \cdot \mathbf{J}_n + G_n - R_n \quad (1.14)$$

$$\frac{dp}{dt} = \frac{1}{q} \nabla \cdot \mathbf{J}_p + G_p - R_p \quad (1.15)$$

$$\nabla \cdot \epsilon \nabla \Phi = -\rho(\Phi) \quad (1.16)$$

Nel sistema di Adler, il meccanismo di ricombinazione prevalente è quello di Shockley-Read-Hall (SRH) attraverso stati trappola. I tassi di ricombinazione di elettroni e lacune si esprimono quindi:

$$R_n = \alpha_n n C_3^+ \quad (1.17)$$

$$R_p = \alpha_p n p_3^0 \quad (1.18)$$

dove α_n e α_p sono, rispettivamente, la probabilità di cattura di elettroni da parte delle trappole cariche e di cattura di lacune da parte delle trappole neutre.

La generazione di portatori è descritta da un meccanismo che è una funzione monotona sia del campo elettrico, sia delle densità di portatori:

$$G = A(n + p)g(\mathbf{E}) \quad (1.19)$$

in cui A è una costante e g è una funzione monotona del campo elettrico E . Il sistema di equazioni differenziali in un materiale omogeneo si traduce nel bilancio di generazione e ricombinazione in condizioni di neutralità di carica, per cui si ottiene

$$G = R_n = R_p \quad (1.20)$$

Ora che le equazioni fondamentali necessarie al modello sono state descritte, si introducono due nuove quantità: la frazione di trappole cariche $f = \frac{C_3^+}{C_3^+ + C_3^0}$ e il rapporto tra portatori minoritari (elettroni) e portatori totali $\eta = \frac{n}{n+p}$. Cominciando dall'equazione 1.20

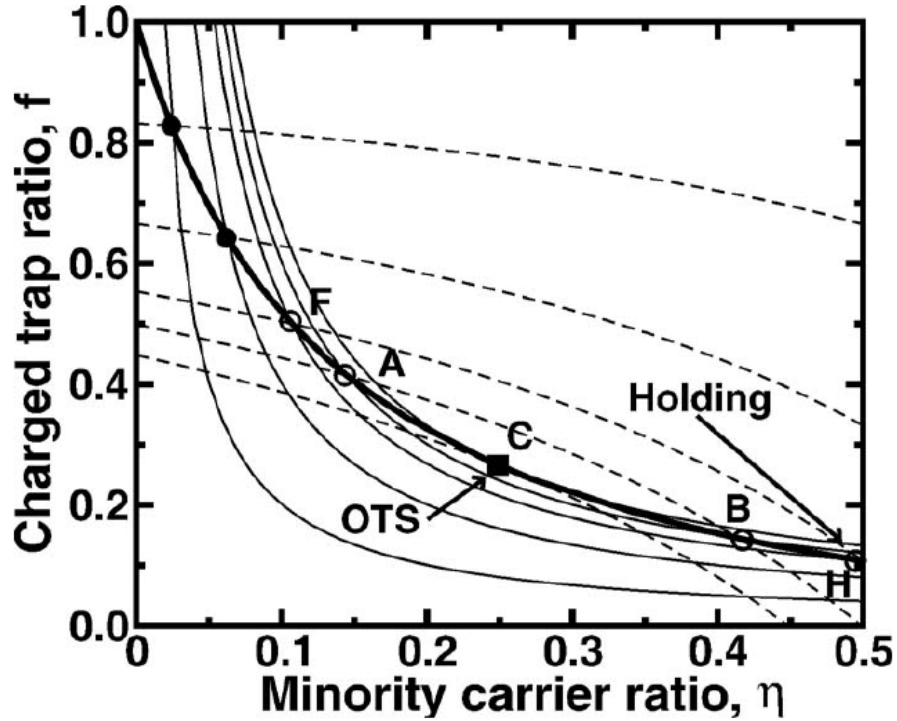


Figura 1.7: Rapporto tra trappole cariche e trappole totali in funzione della frazione dei portatori minoritari. L'intersezione delle linee continue e tratteggiate costituisce la famiglia di punti ammessi come soluzione del sistema.

e uguagliando le equazione 1.19 e 1.17 si ottiene

$$kg(\mathbf{E}) = f\eta \quad (1.21)$$

dove $k = A/(\alpha_n C_3^{tot})$ è una costante. Inoltre, eguagliando le equazione 1.19 e 1.18 si ottiene la seguente espressione:

$$\frac{\alpha_n}{\alpha_p} kg(\mathbf{E}) = (1 - f)(1 - \eta) \quad (1.22)$$

Se si impone la condizione di trasporto di tipo p ($p > 0$ e $N_a > C_3^{tot}$) si ottiene la condizione $\eta < 0.5$, che deve essere messa a sistema con le due equazioni precedenti. La figura 1.7 mostra la famiglia di curve descritte dall'equazione 1.21 con i tratti solidi e dall'equazione 1.22 con le linee tratteggiate per diversi valori del campo elettrico \mathbf{E} . I pun-

ti che risolvono il sistema sono le intersezioni delle due curve, e giacciono su una curva simmetrica rispetto all'asse $f = \eta$, rappresentata con una curva solida spessa. Se si trascura momentaneamente la condizione $\eta < 0.5$, per ogni valore del campo elettrico esistono due configurazioni di sistema che soddisfano le equazioni enunciate, ad esempio, per (A) e (B): (A) un'elevata concentrazione di trappole cariche f e un valore basso di rapporto dei portatori minoritari η oppure (B) un numero elevato di portatori minoritari e una bassa concentrazione di trappole cariche. Nel caso (A) è possibile avere pochi elettroni liberi e molte trappole cariche quando il tasso di generazione è trascurabile e la reazione $C_3^+ + e \rightarrow C_3^0$ è poco probabile. Nel caso (B) si è in presenza di un alto numero di trappole neutre e di numerosi elettroni liberi. Questa condizione è vera per alte concentrazioni di portatori quando avviene un'elevata generazione e il numero di elettroni aumenta. Quindi il sistema di Adler può presentare due diverse configurazioni di carica per lo stesso campo elettrico applicato. Nel caso del punto (C) di figura 1.7 le due diverse configurazioni collassano in una soluzione singola, che corrisponde allo switching di soglia. Solo un valore per f e per η è ammesso, e coincide con

$$f = \eta = \sqrt{kg(\mathbf{E}_{th})} \quad (1.23)$$

in cui \mathbf{E}_{th} è il campo elettrico nelle condizioni di switching. Imponendo la condizione $\eta < 0.5$ una parte del piano non è più ammessa, e le soluzioni del sistema si suddividono in due gruppi: uno per il quale esiste una sola soluzione nel piano ammesso e uno nel quale esistono entrambe le soluzioni. Il punto (H) definisce l'ultimo punto in cui la soluzione è doppia e corrisponde al limite asintotico $I \rightarrow \infty$, che rappresenta la condizione di holding. Pertanto, secondo questo modello lo switching è determinato dall'equilibrio tra la generazione di portatori e la ricombinazione dovuta alle trappole. L'ipotesi fondamentale alla base della descrizione numerica è che la generazione di portatori sia una funzione crescente del campo elettrico e della concentrazione totale di portatori. Un meccanismo noto che abbia queste caratteristiche è la ionizzazione a impatto, a cui viene imputato l'effetto di switching nei materiali calcogenuri.

Cambiamento di fase

Vitrificazione. Quando un calcogenuro viene raffreddato, può solidificarsi in due modi diversi a seconda della velocità di diminuzione della temperatura. Per una velocità di quenching sotto a una certa soglia critica il liquido può avere il tempo necessario a formare i legami reticolari, e si presenterà nella fase cristallina. Se invece il raffreddamento è molto rapido si riporta nello stato solido senza avere formato un reticolo, e rimane in uno stato disordinato. Pertanto per ottenere la transizione di fase dallo stato cristallino allo stato amorfo è necessario portare la temperatura del materiale oltre il punto di fusione (circa 900 K per il $\text{Ge}_2\text{Sb}_2\text{Te}_5$) e raffreddarlo rapidamente per evitarne la cristallizzazione. Valori tipici della velocità di quenching richiesti per questo genere di transizioni variano tra 10^{10} e 10^{11} K/s. Il processo di vitrificazione, nonostante abbia un'importanza vitale per le memorie a cambiamento di fase, non è mai stato studiato in maniera dettagliata, limitandosi sempre a fornire indicazioni di uso pratico, come la corrente necessaria per raggiungere la fusione. Questa mancanza è in parte giustificata dal fatto che in un tipico dispositivo PCM il tempo necessario per la fusione e il quenching sono limitati principalmente dal ritardo tra l'applicazione dell'impulso elettrico e l'aumento di temperatura nell'area attiva della cella. Il tempo di vitrificazione è inferiore a 1 ns, mentre le costanti termiche del dispositivo risultano essere di qualche nanosecondo. Per cui la cinetica di vitrificazione è più rapida dei ritardi termici ed è completamente trascurabile rispetto ai tempi di accesso richiesti dalle applicazioni delle PCM (tra 50 e 150ns).

Cristallizzazione. Il processo di cristallizzazione richiede che il materiale amorfo sia mantenuto a una temperatura inferiori a quella di fusione T_m per un certo periodo di tempo. In queste condizioni avviene prima la nucleazione dei cristalliti e in seguito la loro crescita. Alcuni modelli che descrivono la cristallizzazione sono presenti in letteratura, in particolare negli anni '30 è stata sviluppata la teoria di Avrami, che modella il fenomeno in condizioni di isothermicità (tutto il sistema è alla stessa temperatura) e isocineticità (i processi di nucleazione e crescita hanno la medesima dipendenza dalla temperatura). Sotto queste ipotesi il modello mette in relazione la frazione di materiale cristallizzato alla

temperatura e al tempo trascorso nella forma

$$x(t) = 1 - e^{-A(T)t^n} \quad (1.24)$$

in cui $x(t)$ è la frazione trasformata, n è detto *coefficiente di Avrami* e $A(T)$ è un tasso effettivo di cristallizzazione dipendente dalla temperatura, che può essere espresso come

$$A(T) = \nu e^{\frac{-E_a}{k_B T}} \quad (1.25)$$

dove E_a è l'energia di attivazione del processo e ν un fattore di frequenza. Questo modello è stato originariamente sviluppato per le memorie ottiche: infatti le caratteristiche di rilievo per questi dispositivi (ad esempio la riflettanza del materiale) sono linearmente dipendenti dalla frazione di materiale che ha subito il cambiamento di fase. Questa linearità viene meno nei dispositivi PCM elettrici, per via dell'esistenza di effetti percolativi. Infatti, per poter osservare il passaggio dallo stato alto-resistivo allo stato basso-resistivo, è necessario che si sia formato un percorso conduttivo di materiale cristallizzato che unisca i due elettrodi di contatto. In caso contrario, si continuerà a misurare una resistenza elevata tipica dello stato amorfo. Una volta che questo percorso è stato creato, le successive nucleazioni hanno un effetto notevolmente ridotto sulla conduzione totale del dispositivo. Oltre a questa evidente non linearità, il dispositivo durante la fase di programmazione non opera né in condizioni di isothermicità né di isocinetività, per cui il modello Avrami non può essere usato per studiare questo fenomeno nelle memorie elettriche a cambiamento di fase, ma è necessario svilupparne una descrizione che tenga conto della temperatura locale in ogni punto del dispositivo. Si consideri una generica transizione da una fase l a una fase x a una temperatura T inferiore alla temperatura di fusione T_m e si considera la formazione di una sfera di fase x all'interno di una matrice di fase l . L'energia libera della sfera si può esprimere come

$$\Delta G_{\text{sphere}}^{l \rightarrow x} = 4\pi r^2 \gamma + \frac{4}{3}\pi r^3 \Delta g \quad (1.26)$$

in cui γ è l'energia di superficie, r il raggio della sfera e Δg è il guadagno di energia libera di bulk relativo a un raffreddamento $T - T_m$. Per $T < T_m$ il fattore di bulk diventa negativo ed è possibile la formazione di sfere con certi valori di r . La stabilità della sfera è in

relazione con la possibilità di assorbire nuovi atomi dalla fase l , e quindi con la derivata della 1.26 rispetto a r . Se questa derivata è negativa, allora è possibile l'aumento del raggio della sfera da r a $r + dr$ per via del guadagno di energia libera. Imponendo questa condizione si ottiene un valore critico del raggio e l'energia libera ad esso associata:

$$r_c = \frac{2\gamma}{\Delta g} \quad (1.27)$$

$$\Delta G_c = \frac{16\pi\gamma^3}{3\Delta g^2} \quad (1.28)$$

Se si ipotizza che gli atomi di l siano aggregati casualmente in cluster di dimensioni variabili con una distribuzione di energia data dalla statistica di Boltzmann, la concentrazione di cluster critici si può esprimere come

$$C_n = \rho e^{-\left(\frac{\Delta G_c}{k_B T}\right)} = \rho e^{-\left(\frac{16\pi\gamma^3}{3\Delta g^2 k_B T}\right)} \quad (1.29)$$

dove ρ è la densità dei cluster. Non appena un atomo l si attacca al cluster critico, la cristallizzazione ha luogo e in nucleo risultante è stabile. Il tasso di nucleazione si può quindi esprimere nella forma

$$I = C_n \frac{dn}{dt} \quad (1.30)$$

dove dn/dt è il tasso con cui gli atomi l si aggiungono al cluster critico. Si ha

$$\frac{dn}{dt} = \nu s_c p e^{-\left(\frac{E_{an}}{k_B T}\right)} \quad (1.31)$$

in cui E_{an} è la barriera energetica alla superficie, ν è la frequenza di vibrazione di un atomo l , s_c è il numero di atomi rivolti nella direzione del cluster x e p è la loro probabilità di saltare nel cluster. Per ottenere il tasso di nucleazione finale si inseriscono le equazioni 1.29 e 1.31 nella 1.30. Se si ipotizza inoltre che il guadagno di energia libera di bulk sia proporzionale, tramite una costante, all'entità del raffreddamento, in modo che $\Delta g = \delta(T - T_m)$, si ottiene infine:

$$I = K e^{-\left(\frac{E_{an}}{k_B T}\right)} \cdot e^{-\left(\frac{16\pi\gamma^3}{3[\delta(T-T_m)]^2 k_B T}\right)} \quad (1.32)$$

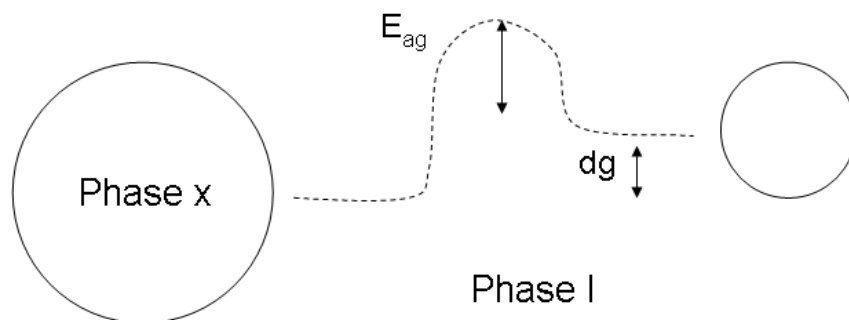


Figura 1.8: Rappresentazione schematica delle barriere energetiche coinvolte nel meccanismo di crescita di un cluster x all'interno della fase l . Il cluster x aumenta di dimensioni non appena un monomero acquisisce energia sufficiente per superare la barriera E_{ag} . L'elevata barriera tra x e l rende il cammino inverso molto meno probabile.

in cui K contiene tutte le precedenti costanti. Il primo termine è attivato in temperatura a un'energia E_{an} e rappresenta la diffusione atomica, mentre il secondo, associato all'esistenza di un cluster critico disponibile, si disattiva per temperature che si avvicinano alla fusione. Questo significa che per basse temperature ci sono numerosi cluster critici disponibili ma non avverrà nessuna diffusione atomica da l a x , mentre a temperature prossime alla fusione molti atomi hanno la possibilità di diffondere ma non ci sarà nessun cluster critico ad accoglierli. Questi semplici calcoli mettono in luce le principali dipendenze dalla temperatura del tasso di nucleazione nel corso del processo di cristallizzazione.

Una volta che un cluster ha raggiunto dimensioni tali da diventare un nucleo cristallino stabile, altri atomi possono diffondere al suo interno, aumentando quindi la dimensione del cristallite. Si consideri una barriera energetica all'interfaccia tra le fasi l e x , come mostrato in figura 1.8, in cui E_{ag} è la barriera di diffusione e dg il guadagno di energia libera. Per questo sistema il tasso di passaggio di atomi da l a x si scrive

$$r_{lx} = Sv_0 e^{-\left(\frac{E_{ag}}{k_B T}\right)} \quad (1.33)$$

in cui v_0 è la frequenza di vibrazione atomica e S il numero di atomi che si trovano all'interfaccia. Esiste anche una probabilità che gli atomi diffondano da x a l , a patto che

riescano a superare una barriera di altezza $E_{ag} + dg$, per cui si può scrivere:

$$r_{xl} = Sv_0 e^{-\left(\frac{E_{ag} + dg}{k_B T}\right)} \quad (1.34)$$

Si può quindi ricavare il tasso netto di diffusione da l a x scrivendo la differenza tra i due flussi in entrata e in uscita, e si ha quindi:

$$r = Sv_0 e^{-\left(\frac{E_{ag}}{k_B T}\right)} \cdot [1 - e^{-\left(\frac{dg}{k_B T}\right)}] \quad (1.35)$$

Se si moltiplica questa espressione per la distanza media di salto λ , che si suppone essere uguale al parametro del reticolo di x , si ottiene la velocità di crescita del cristallo. Poiché il termine dg aumenta all'aumentare dell'entità del raffreddamento, il secondo termine è praticamente trascurabile per la quasi totalità delle temperature in gioco durante il processo di cristallizzazione.

1.3 La cella di memoria Wall

Nella sezione precedente sono state descritte le proprietà salienti del materiale calcogeno, vero e proprio cuore della cella, in cui viene immagazzinata l'informazione. Tuttavia affinché il dispositivo possa essere adoperato come memoria sono necessari altri elementi. Infatti, per via della dipendenza dalla temperatura delle resistività del GST illustrata in precedenza, risulta difficile portare il composto alla temperatura di fusione impiegando impulsi di corrente di ampiezza e durata contenute. È quindi opportuno che una parte del dispositivo sia occupato da un riscaldatore (heater) che abbia lo scopo di generare calore e convogliarlo al materiale attivo, in qualunque condizioni quest'ultimo si trovi. Sono poi necessari contatti di tipo metallico per connettere la cella al circuito di selezione di riga e di colonna e dei materiali isolanti per confinare adeguatamente il calore all'interno della zona attiva. La cella PCM prodotta presso Numonyx è caratterizzata da una struttura di tipo *Wall* [6, 7]. Su un substrato di silicio viene depositato un contatto, chiamato *plug*, di tungsteno, sopra al quale viene depositato uno spesso strato di materiali isolanti (SiN e SiO₂). All'interno di questi viene praticato un foro a pianta rettangolare

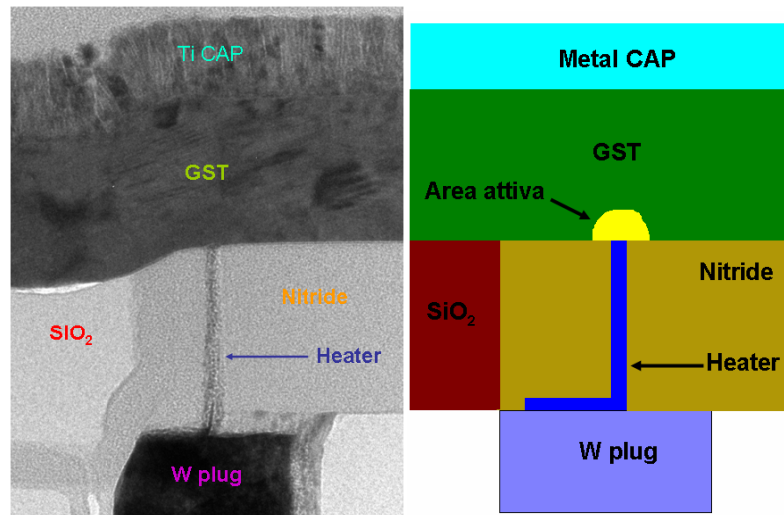


Figura 1.9: Immagine TEM di un cella realizzata con architettura Wall (a) e uno schema semplificato della struttura (b).

all'interno del quale viene depositato l'heater, che possiede caratteristiche simili a quelle di uno semiconduttore degenere di tipo p. Sopra l'heater viene depositato il GST, confinato all'interno degli isolanti, sul quale viene poi depositato un "cap" in metallo che connette alla bit-line. L'area attiva del dispositivo è quindi definita dall'intersezione dell'heater con il calcogenuro. Quest'interfaccia è molto importante in quanto è proprio qui che si avrà il maggiore aumento di temperatura durante l'operazione, e questo processo ha come vantaggio un buon controllo sulle dimensioni dell'area attiva. Per accedere al dato è necessario un elemento di selezione, ovvero un dispositivo che a fronte di un segnale di corrente applicato alla bit-line o alla word-line dell'array di memoria sia in grado di collegare la cella al circuito esterno. In questa struttura viene fatto uso di un transistor bipolare PNP, preferito a un dispositivo di tipo MOSFET al fine di limitare la dimensione totale della cella. Infatti, mentre il MOSFET deve essere realizzato necessariamente con una struttura planare, il BJT opera verticalmente, consentendo di non occupare spazio ai lati della cella. L'emettitore è connesso all'heater tramite il plug di tungsteno, mentre un contatto metallico connette la base alla word-line. Il substrato di silicio forma il collettore, che è quindi posto a massa. Ai fini della nostra simulazione, i contatti saranno trattati come perfetti

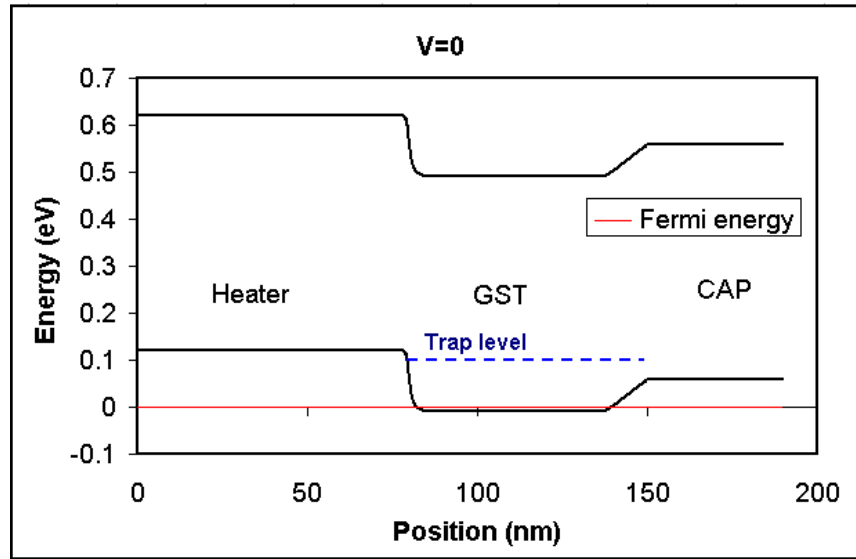


Figura 1.10: Struttura a bande di una cella Wall 90nm in assenza di tensione esterna applicata. L'heater si comporta come un semiconduttore degenere per via dell'elevato drogaggio, mentre il GST come un semiconduttore quasi-degenere a causa del livello trappole in prossimità della banda di valenza.

conduttori termici ed elettrici: la differenza di potenziale e di temperatura ai loro capi sarà perciò trascurata. L'heater ha un comportamento di tipo degenere, e viene quindi modellizzato come un semiconduttore di tipo p pesantemente drogato ($N_a = 2.5 \cdot 10^{20} \text{cm}^{-3}$) con un band-gap di 0.5 eV, e una mobilità dei portatori con una dipendenza power-law dalla temperatura, tale che

$$\mu = \mu_0 \left(\frac{T}{T_0} \right)^\alpha \quad (1.36)$$

I parametri μ_0 e α verranno trattati successivamente. Essi sono molto importanti, in quanto di fatto fissano la resistività dell'heater. Quest'ultimo determina alcune fondamentali caratteristiche della cella: la corrente richiesta per fondere il materiale attivo (infatti il calore generato dipende dalla corrente e dalla resistenza che essa attraversa) e la resistenza totale della cella alle alte tensioni, perché in tali condizioni operative la resistenza del GST diventa trascurabile. La figura 1.10 mostra la struttura a bande della cella appena descritta, in assenza del transistor di selezione.

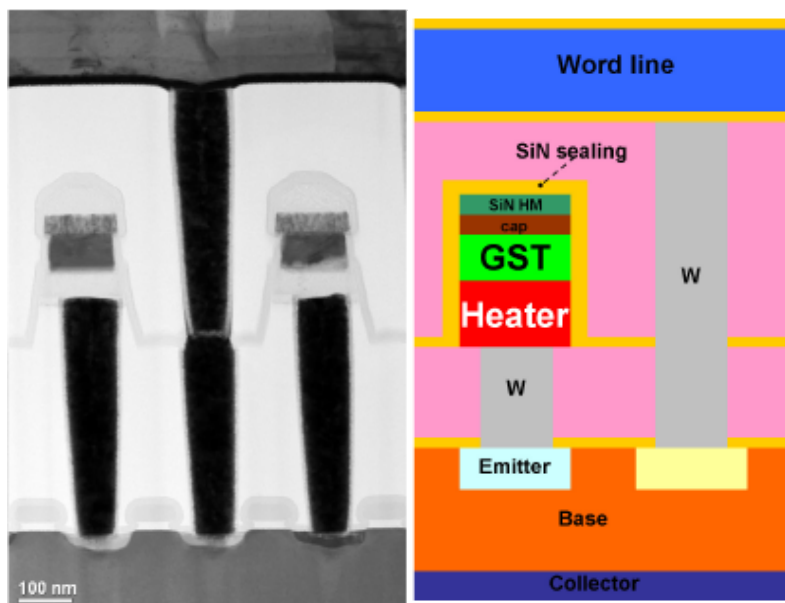


Figura 1.11: Immagine TEM di un cella realizzata in struttura Wall (a) e relativo schema semplificato (b) in cui si può notare la struttura che connette la cella all'array di memoria: il BJT verticale e il contatto di base.

1.3.1 Parametri caratteristici della cella e tecniche di programmazione

Di seguito verrà data la definizione di alcuni parametri che aiutano a caratterizzare le proprietà della cella, e che verranno usati frequentemente in seguito. La figura 1.12 mostra graficamente il loro significato.

Resistenza di lettura. Si tratta della resistenza letta a bassa tensione, e serve a determinare lo stato della cella: essa sarà alta nello stato di RESET e bassa nello stato di SET. A seconda del caso, viene chiamata anche *resistenza di RESET* (R_{RESET}) o *resistenza di SET* (R_{SET}) rispettivamente. La misurazione avviene applicando al dispositivo una tensione relativamente bassa (pari o inferiore a 0.4 V) e leggendo la corrente risultante. Poiché in questo lavoro le simulazioni si concentreranno principalmente sulla fase cristallina, la resistenza di SET sarà quella che verrà trattata più frequentemente. Questo parametro fornisce informazioni sulla finestra di lettura della PCM, definita come la differenza tra la le

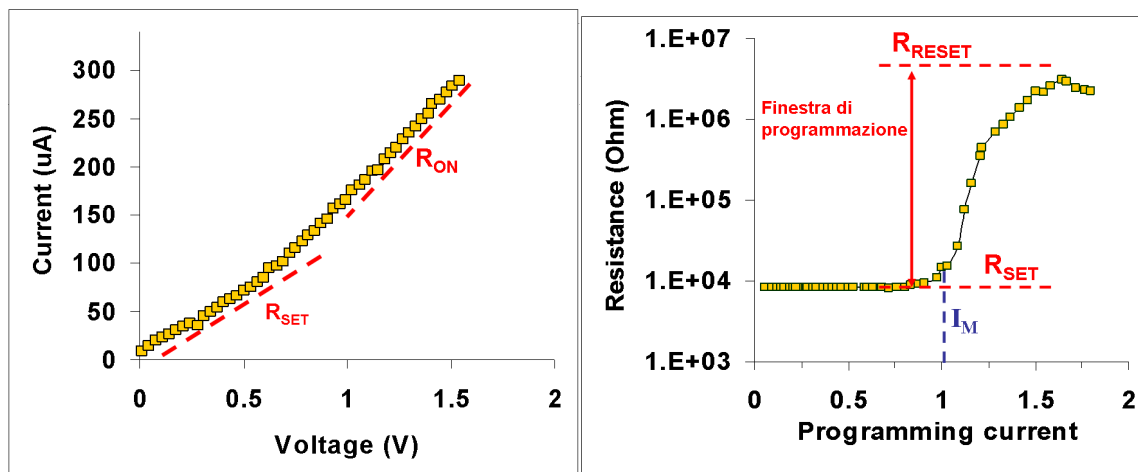


Figura 1.12: Curva I-V (sinistra) e curva di programmazione (destra) di una tipica cella a 45nm. Sono mostrati i parametri caratteristici del dispositivo: resistenza a bassi e ad alti campi, resistenza di RESET, corrente di fusione e finestra di lettura. Nella curva R-I le correnti sono normalizzate alla I_M .

resistenza di SET e RESET, e si vuole che sia il più ampia possibile. Infatti, più i due livelli sono separati, meno la memoria è sensibile a rumori di origine elettronica e termica.

Resistenza di ON. Nella fase cristallina, in presenza di alte tensioni applicate alla cella, la resistenza totale tende a essere dominata dall'heater. La resistenza di ON (R_{ON}) è la resistenza differenziale misurata ad alti campi elettrici (tipicamente per correnti superiori a 300-500 μA , a seconda della dimensione della cella), vale a dire la pendenza della parte alta della curva I-V nella fase di SET. Questo parametro fornisce informazioni sull'heater e sulle resistenze interfacciali, e risulta molto utile per la loro caratterizzazione. Quando il dispositivo opera a tensioni vicine a quella che permette la fusione del GST, la resistenza di quest'ultimo è già calata notevolmente per via della generazione di portatori al suo interno, per cui la R_{ON} contribuisce in modo determinante al determinare la temperatura all'interfaccia GST/Heater.

Corrente di fusione. La corrente di fusione (o di melting, I_M) è la più piccola ampiezza dell'impulso di programmazione in grado di fondere una parte del GST. In verità, per avere

un margine di riparo da errori di carattere numerico, essa viene definita come l'ampiezza dell'impulso che fa aumentare la resistenza di lettura del 20%. Tutte le correnti al di sotto della I_M non sono in grado di creare uno strato amorfo, per cui lasceranno del tutto invariate le caratteristiche della cella. Al di sopra di questa corrente, al crescere dell'impulso aumenterà anche la dimensione della zona vetrosa, e si avrà quindi un progressivo aumentare della resistenza del dispositivo. La corrente di melting è un parametro molto importante perchè determina la potenza necessaria per il cambiamento di fase, e quindi l'efficienza energetica della memoria.

Tensione di soglia. È la tensione che bisogna applicare al dispositivo quando si trova nello stato di RESET per provocarne lo switching di soglia. Di fatto la tensione di soglia, indicata con V_{th} , indica la potenza necessaria per il passaggio di stato da RESET a SET. Infatti, dopo che è avvenuto lo switching la resistenza del GST crolla ed è pertanto più facile iniettare corrente e aumentare la temperatura del calcogenuro sopra quella di cristallizzazione. La V_{th} aumenta all'aumentare della dimensione della calotta amorfa, la quale a sua volta è funzione dell'impulso di programmazione, per cui si ha una correlazione diretta tra tensione di soglia e resistenza di RESET.

Impulsi di programmazione. Per indurre il cambiamento di fase nella cella si devono applicare degli impulsi elettrici che provocano passaggio di corrente, e quindi dissipazione di potenza per effetto Joule e aumento di temperatura. I segnali che verranno descritti di seguito sono tutti impulsi di tensione, che si traducono, mediante la resistenza della cella, in segnali di corrente. Per programmare la cella nello stato di RESET è necessario fornire un impulso di ampiezza relativamente elevata, seguito da un fronte di discesa molto rapido, tale da permettere al materiale di raffreddarsi senza che i legami cristallini abbiano il tempo di formarsi: in questo modo il GST si stabilizza nella sua fase vetrosa. Questo *impulso di RESET* deve avere un'ampiezza di circa 2V, una durata di qualche decina di nanosecondi e un tempo di discesa che non deve superare i 10ns. Per quanto riguarda la programmazione nello stato di SET sono possibili due approcci. È possibile applicare un impulso simile a quello di RESET, detto *impulso di SET minimo*, che si differenzia da quest'ultimo solamente per la durata del fronte di discesa, decisamente più lunga (nel-

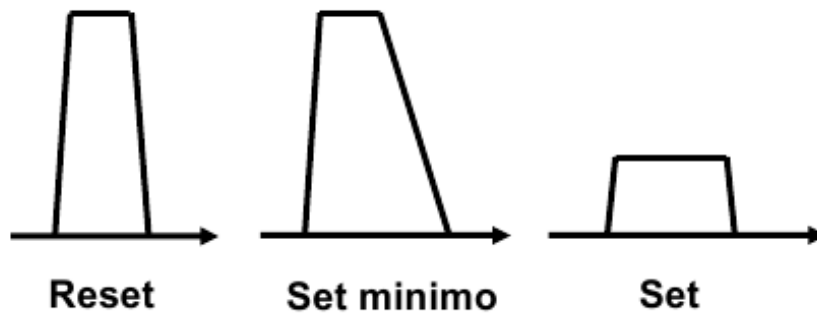


Figura 1.13: Forma degli impulsi di programmazione: RESET, SET minimo e SET.

l'ordine dei microsecondi). In questo modo il calcogenuro viene portato a fusione e poi raffreddato lentamente, in modo da formare i legami cristallini, e si stabilizza nella sua fase fcc. Il nome di questo impulso deriva dal fatto che consente di riportare la cella nello stato di SET migliore, ovvero quello a resistenza più bassa, fondendo e ricristallizzando completamente la calotta amorfa. È anche possibile utilizzare gli *impulsi di SET*, caratterizzati da un'ampiezza più contenuta (1.2 V), che non consente alla cella di raggiungere la temperatura di fusione, ma che le fornisce energia sufficiente a superare la temperatura di cristallizzazione. A questo punto l'energia termica all'interno del GST consente ai legami di riorganizzarsi secondo un ordine reticolare, senza dover passare dalla fase liquida. La PCM così programmata si troverà in uno stato basso resistivo, che non corrisponde al minimo ottenibile con l'impulso di SET minimo, ma può comunque bastare per avere una finestra di lettura ragionevole.

Per caratterizzare la cella è molto utile tracciare le cosiddette **curve di programmazione**. È possibile applicare una sequenza opportuna di impulsi di programmazione seguiti da impulsi di lettura, e tracciare su un grafico la resistenza letta in funzione dell'ampiezza dell'impulso applicato in precedenza. La curva risultante, detta anche **R-I**, fornisce alcune informazioni preziose sulla cella, in particolare riguardo la corrente di fusione e l'ampiezza della finestra di lettura. La curva che verrà utilizzata più frequentemente in questo lavoro è la R-I parziale, costruita leggendo la resistenza di una cella a cui viene applicata una serie di impulsi di RESET di ampiezza crescente, partendo inizialmente dallo stato di SET minimo. Finché l'ampiezza dell'impulso non supera I_M non avviene nessun cambi-

amento rilevante nella PCM. Man mano che la corrente iniettata aumenta, il dispositivo si scalda e si crea una calotta di materiale amorfo nei pressi dell'interfaccia GST/Heater, la cui dimensione aumenta con l'ampiezza dell'impulso applicato. Per correnti elevate si avrà una saturazione della resistenza letta, per via di una limitazione dell'apparato sperimentale utilizzato: infatti non è in grado di leggere resistenze troppo elevate (superiori a qualche $M\Omega$) di questo problema si discuterà più in dettaglio nel prossimo capitolo. La scaletta di impulsi può anche essere applicata al contrario, cioè partendo da uno stato di RESET e diminuendone l'ampiezza: ogni impulso che porti la temperatura della cella tra la temperatura di cristallizzazione e quella di fusione provocherà il passaggio alla fase fcc di parte della calotta amorfa.

È anche possibile tracciare una curva di programmazione completa intervallando impulsi di SET di ampiezza crescente a impulsi di pieno RESET. In questo modo, lo stato di partenza della cella sarà sempre la fase amorfa. L'ampiezza dell'impulso di SET cresce, e quando la temperatura raggiunta supera quella di cristallizzazione, la resistenza letta comincia a calare, fino a raggiungere quella tipica di uno stato di SET. Se però ora la corrente aumenta a tal punto da portare il calcogeno a fusione si avrà l'equivalente di una sequenza di RESET, e la resistenza totale torna ad aumentare fino ai valori di pieno RESET. La curva ottenuta presenta una forma a vasca, tipica delle memorie a cambiamento di fase.

1.3.2 Modello analitico della cella Wall

Nonostante i fenomeni che governano la cella PCM siano complessi, per via dell'ampia escursione di temperatura abbinata, è utile elaborare un modello analitico semplificato che sia in grado di dare una stima del prim'ordine delle sue proprietà salienti. Questo modello viene utilizzato operativamente per stimare i parametri da inserire nel simulatore numerico e per avere un'idea delle prospettive di scaling della memoria. Per avere un preciso risultato bisognerebbe risolvere le equazioni dei semiconduttori e del trasporto di calore, tuttavia a causa della struttura tridimensionale e della presenza di numerose interfacce tra materiali diversi, il calcolo analitico risulta impossibile. Per questa ragione la soluzione di queste equazioni viene effettuata per via numerica, come descritto nel prossimo capitolo. In ogni caso, è possibile, in base a ragionevoli approssimazioni, giungere a delle descrizioni sem-

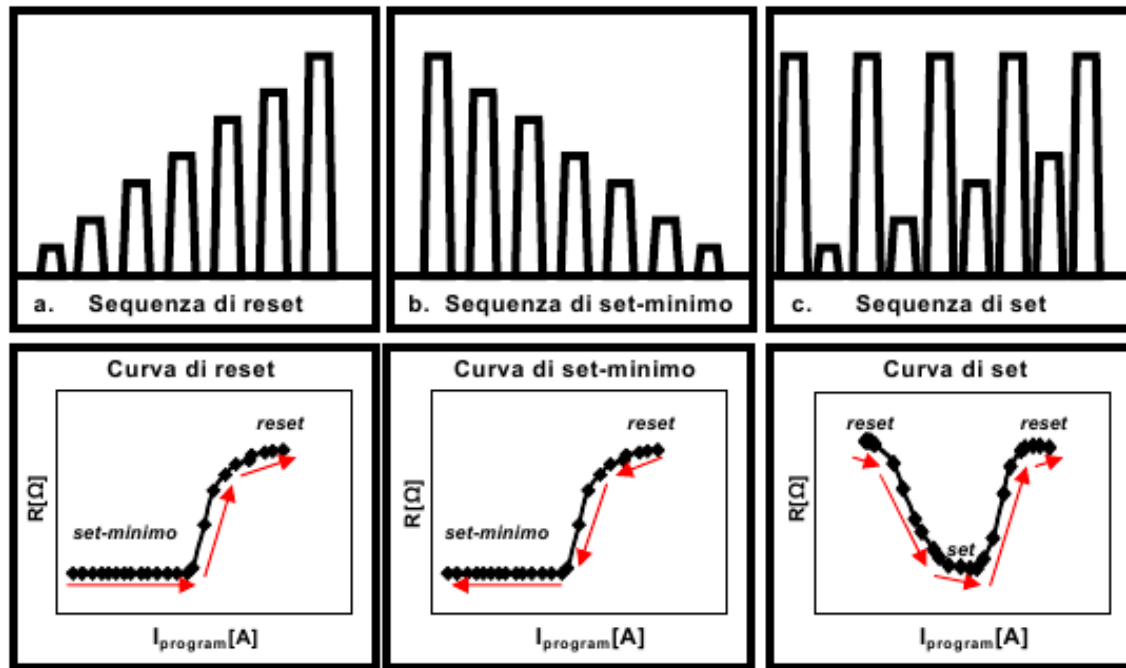


Figura 1.14: Curve di programmazione ottenute con sequenze di impulsi differenti. Dopo ogni impulso viene applicato un impulso di lettura (circa 0.4 V).

plificate dei fenomeni in gioco. In particolare, il comportamento termico ed elettrico della fase cristallina può essere analizzato, mentre per quel che riguarda il profilo di temperatura che il dispositivo raggiunge a seguito di un impulso (utile per determinare lo spessore della calotta amorfa, e quindi la resistenza di RESET) non si è in grado di giungere a una forma analitica semplice ma al contempo accettabile. Il comportamento della cella Wall nella fase amorfa è dominato dall'elevata resistenza del calcogenuro vetroso, rendendo trascurabile il contributo dell'heater, almeno fino a che non avviene lo switching di soglia. Si noti che in entrambi si ipotizza di avere a disposizione una cella analitica, ovvero priva del transistor di selezione.

La cella Wall è schematizzabile in prima approssimazione come una serie di tre resistenze elettriche, che rappresentano l'heater, l'interfaccia GST/Heater e il calcogenuro. La dipendenza dalla temperatura di resistenza di heater e di interfaccia verrà trascurata in questo

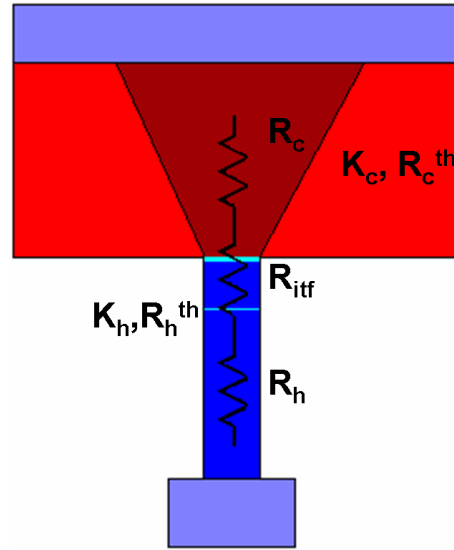


Figura 1.15: Modello analitico resistivo per la cella PCM. I tre contributi elettrici sono dovuti all'heater, all'interfaccia GST/Heater e al GST, mentre i contributi termici sono dovuti solamente a heater e GST.

modello, mentre quella del GST sarà descritta a tratti: avrà un valore finito a bassi campi elettrici e sarà nulla per alti campi. La resistenza dell'heater è data da:

$$R_h = \rho_h \frac{L_h}{A_h} \quad (1.37)$$

dove ρ indica la resistività, L_h l'altezza dell'heater e A_h la sua area. La resistenza di interfaccia è invece data da

$$R_{if} = \rho_{if}^{\square} \frac{1}{A_c} \quad (1.38)$$

in cui ρ_{if}^{\square} è la resistività interfacciale per unità di spessore dello strato e A_c l'area di contatto tra il GST e l'heater, che delimita la zona attiva. Le ultime due equazioni sono state formulate nell'ipotesi che la densità di corrente sia omogenea in tutto il materiale. In realtà l'area dell'heater è più grande dell'area di contatto di una decina di nanometri, per motivi legati al processo tecnologico, tuttavia l'effetto di spread della corrente può essere trascurato in quanto l'altezza dell'heater è maggiore della differenza di larghezza tra i due elementi. Questa ipotesi non vale quando si tratta il materiale calcogenuro, in

quanto la corrente uscente dall'interfaccia si ritrova in uno strato di larghezza infinita in confronto con la sua altezza. Si può ipotizzare che la corrente scorra aprendosi con forma trapezoidale ad un angolo α rispetto al piano contenente l'interfaccia. In questo caso la resistenza infinitesima di uno strato di GST alto dx con pianta rettangolare vale:

$$dR = \rho_c \cdot \frac{dx}{W_c(t + 2nx)} \quad (1.39)$$

dove ρ_c è la resistività del GST, W_c la sua larghezza, t lo spessore e n un parametro che dipende dall'angolo di fuoriuscita della corrente. Se α è l'angolo formato dalla retta verticale perpendicolare all'interfaccia e la retta che rappresenta il flusso di corrente, n vale 1 se $\alpha=45^\circ$, 0 se $\alpha=0$ e tende a $+\infty$ se $\alpha \rightarrow 90^\circ$. In sostanza, si ha $n = tg(\alpha)$. Per avere la resistenza totale si integra lungo tutta l'altezza L_c del materiale attivo:

$$R_c = \int_0^{L_c} \rho_c \cdot \frac{dx}{W_c(t + 2nx)} = \frac{\rho_c}{2nW_c} \cdot \ln\left(1 + \frac{2nL_c}{t}\right) \quad (1.40)$$

È da osservare come in questo caso il valore della resistenza dipenda solo logicamente dall'altezza e dallo spessore dello strato, mentre mantiene la dipendenza iperbolica dalla larghezza. Con l'approssimazione descritta in precedenza, ovvero ipotizzando che ad alti campi la resistività del GST sia nulla, è possibile scrivere in maniera semplice le resistenze R_{ON} e R_{SET} che caratterizzano la cella:

$$R_{ON} = R_h + R_{itf} \quad (1.41)$$

$$R_{SET} = R_c + R_h + R_{itf} \quad (1.42)$$

Per descrivere l'andamento della temperatura all'interfaccia GST/Heater, dove si avrà il massimo del riscaldamento Joule e dove comincia a formarsi la calotta amorfa, è utile ricorrere all'equivalente elettrico dei fenomeni termici, in cui la temperatura viene raffigurata come il prodotto della potenza dissipata per la resistenza termica. All'interfaccia, quest'ultima risulta essere il parallelo delle resistenze termiche di heater e di GST. Ipotizzando che l'interfaccia abbia lo stesso valore di conducibilità termica dell'heater, queste

sono scrivibile in una forma molto simile alla resistenza elettrica:

$$R_h^{\text{th}} = \frac{1}{K_h} \frac{L_h}{A_h} \quad (1.43)$$

$$R_c^{\text{th}} = \frac{1}{K_c} \frac{1}{2nW_c} \cdot \ln\left(1 + \frac{2nL_c}{t}\right) \quad (1.44)$$

La temperatura raggiunta dall'interfaccia si può scrivere

$$T = T_0 + PR_{\text{th}} = T_0 + I^2 RR_{\text{th}} \quad (1.45)$$

In questa equazione R non è noto a priori. Infatti, come osservato in precedenza, la resistenza totale che la cella presenta è dipendente dalla temperatura, per cui la soluzione di questo problema non è immediata, e richiede calcoli numerici. Tuttavia, è possibile utilizzarla in alcuni casi che consentono una notevole semplificazione. Ad esempio, è possibile ricavarne un modello per la corrente di fusione, ipotizzando di essere nella zona a campi elettrici elevati in cui la resistenza del GST è trascurabile. In questo caso R coincide con R_{ON} e si giunge a una forma chiusa per la corrente, di fusione, imponendo $T=T_m$:

$$I_m = \sqrt{\frac{T_m - T_0}{\alpha R_{\text{ON}}} \left(K_h \frac{A_h}{L_h} + K_c \frac{2nW_c}{\ln\left(1 + \frac{2nL_c}{t}\right)} \right)} \quad (1.46)$$

Il coefficiente α tiene conto della non idealità del riscaldatore: infatti a causa dei materiali circostanti, non perfettamente isolanti, c'è un flusso di calore uscente ai lati dell'heater, di cui è difficile tenere conto analiticamente. Inoltre, si vedrà in seguito che l'ipotesi di trascurare il self-heating all'interno del GST non è del tutto verificata, soprattutto per le celle scalate rispetto al nodo a 90nm, e in seguito verrà proposta una modifica al modello. Nonostante le equazioni appena mostrato siano frutto di semplificazioni, è comunque possibile utilizzarle per sviluppare un modello analitico per lo scaling delle celle, utile per stimare le prestazioni dei futuri nodi tecnologici di queste memorie, nonché per verificare gli andamenti delle simulazioni numeriche.

1.4 Scaling della cella PCM Wall

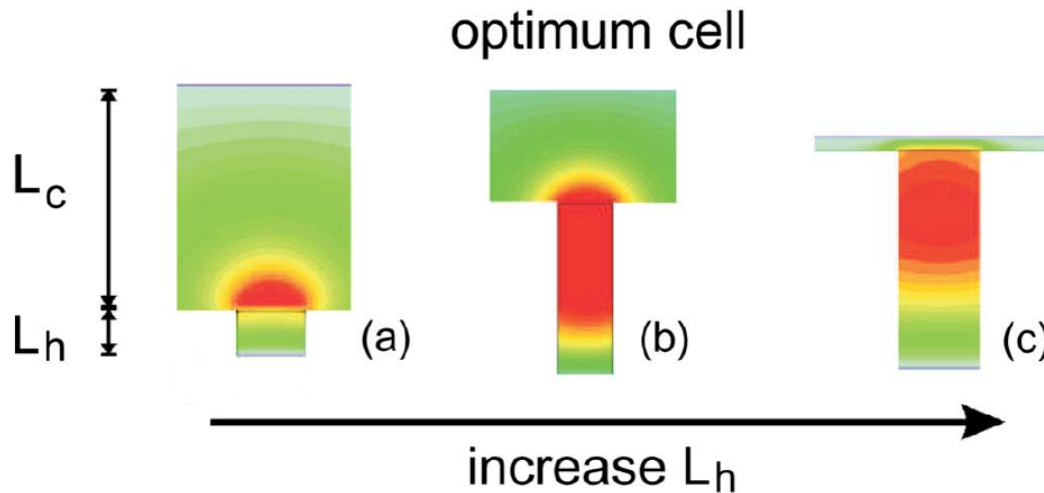


Figura 1.16: Profilo di temperatura e metodologia di ottimizzazione della corrente di programmazione. L_c e L_h vengono modificate in modo da tenere costante la resistenza di SET. Si osserva che l'hot-spot tende a spostarsi all'interno dell'heater all'aumentare di L_h .

Le memorie a cambiamento di fase hanno ricevuto grande interesse negli ultimi anni grazie alle possibilità di scaling che offrono. La scalabilità è infatti uno dei principali fattori che determinano la realizzabilità di una tecnologia, in quanto stabilisce quanto la memoria potrà rimanere competitiva in futuro. Le PCM richiedono una corrente di programmazione relativamente alta, dovuta alla necessità di portare il calcogeno a fusione durante l'operazione di RESET. Questo rappresenta un limite per le applicazioni alta densità e bassa potenza, in quanto pone un vincolo sulle dimensioni del transistor di selezione, che deve necessariamente portare tutta la corrente richiesta dalla cella e sulla possibilità di effettuare più operazioni di scrittura in parallelo, per via dell'alta potenza richiesta. Pertanto lo scaling ha come obiettivo principale quello della riduzione della corrente di programmazione. Quest'ultima è strettamente legata alla corrente di melting, quindi tutti i ragionamenti riguardanti lo scaling saranno effettuati sulla I_m , sapendo che approssimativamente $I_{\text{reset}} \approx 1.5I_m$.

La riduzione di I_m richiede un aumento dell'efficienza del riscaldamento per effetto Joule. Questo a sua volta necessita dell'incremento delle resistenze elettriche e termiche, nonché

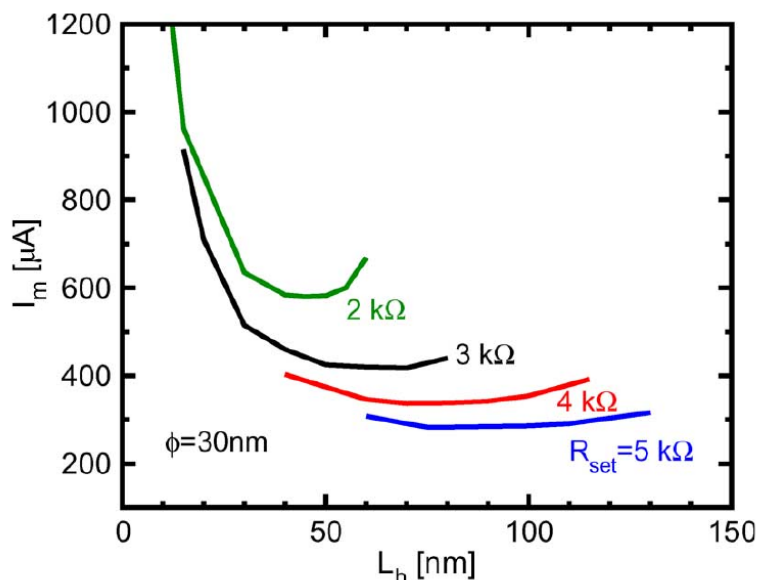


Figura 1.17: Corrente di fusione in funzione di L_h , per diversi valori di R_{SET} fissata. Per ogni valore di resistenza è presente un minimo che rappresenta la geometria ottima della cella.

di un migliore confinamento del calore all'interno dell'area attiva. Questi risultati sono ottenibili con una diminuzione dell'area attiva, tuttavia questo porta anche a un aumento della resistenza di lettura R_{SET} , effetto indesiderato in quanto porta a un aumento di tempi di readout, in quanto la corrente in uscita, risultante da una tensione di lettura applicata di qualche centinaio di mV, si riduce, e così anche la finestra di discernimento dei due livelli logici. Una prima approccio allo scaling consiste nel mantenere fissa R_{SET} agendo opportunamente su L_c e L_h a parità di area di contatto A_c . In particolare, se si aumenta L_h si diminuirà di conseguenza L_c e viceversa. La figura 1.17 mostra l'andamento della corrente di programmazione al variare dell'altezza dell'heater. Si nota che per qualunque valore fissato di R_{SET} si ha un minimo di I_m , che corrisponde a una geometria ottima. Infatti il miglioramento della corrente di RESET è dovuto a due contributi: l'entità del riscaldamento per effetto Joule, che aumenta con l'altezza dell'heater per via dell'incremento di R_{ON} e di R_{ON}^{th} , e la distanza del punto che raggiunge la massima temperatura dall'interfaccia GST/heater. Infatti questo punto, detto *hot-spot*, si allontana all'aumentare di L_h , come viene mostrato in figura 1.16, e tende a spostarsi sempre più all'interno dell'heater. Il punto

di minimo nella curva I_m - L_h corrisponde dunque all'ottimizzazione di questi due contributi. Questo dimostra che, fissate le prestazioni di lettura richieste, è possibile ottimizzare la geometria della cella in modo da avere la più bassa corrente di programmazione possibile. Per avere una stima quantitativa della scalabilità dei dispositivi PCM si prende come riferimento la cella ottimizzata del nodo tecnologico 90nm e si applicano delle regole di scaling arbitrarie alle dimensioni geometriche. In generale sono possibile due approcci: lo scaling isotropo, nel quale tutte le dimensioni della cella (W , L_c , L_h , t) vengono moltiplicate per il medesimo fattore di scala F , e lo scaling anisotropo, in cui alcuni parametri rimangono costanti o vengono scalati di fattori diversi da F . Nella famiglia degli scaling anisotropi si possono effettuare numerose scelte riguardo a quali parametri debbano rimanere costanti o essere modificati diversamente dagli altri. Per prima cosa si analizzerà lo scaling isotropo. Dal modello analitico illustrato in precedenza si ricavano i facilmente gli andamenti di R_{SET} , R_{ON} , I_m in funzione di F :

$$R_{ON} \propto \frac{aF + b}{F^2} \quad (1.47)$$

$$R_{SET} \propto \frac{(a + c)F + b}{F^2} \quad (1.48)$$

$$I_m \propto \sqrt{\frac{F^3}{aF + b}} \quad (1.49)$$

dove a, b e c sono costanti che rappresentano rispettivamente il contributo dell'heater, dell'interfaccia GST/Heater e del GST. La figura 1.18 mostra graficamente questi andamenti. Si osserva che fintanto che l'interfaccia gioca un ruolo trascurabile nel computo della resistenza totale ($b=0$), si ha $R_{ON}, R_{SET} \propto F^{-1}$ e $I_m \propto F$. Lo scaling ha quindi l'effetto di ridurre la corrente di programmazione a scapito di un aumento della resistenza di lettura. Per dimensioni estremamente ridotte si arriverà al punto in cui la resistenza dell'interfaccia, che scala proporzionalmente a F^{-2} diventa dominante, portando a un aumento maggiore di R_{SET} . La figura 1.19 mostra l'andamento di I_m in funzione di R_{SET} , in cui si osserva che la pendenza per il caso di scaling isotropo è di -1 su tutto il range di fattore di scala preso in considerazione.

Lo scaling anisotropo può essere effettuato in svariati modi, a seconda di quali parametri rimangano invariati. Per motivi tecnologici inoltre, le dimensioni che scalano non lo fan-

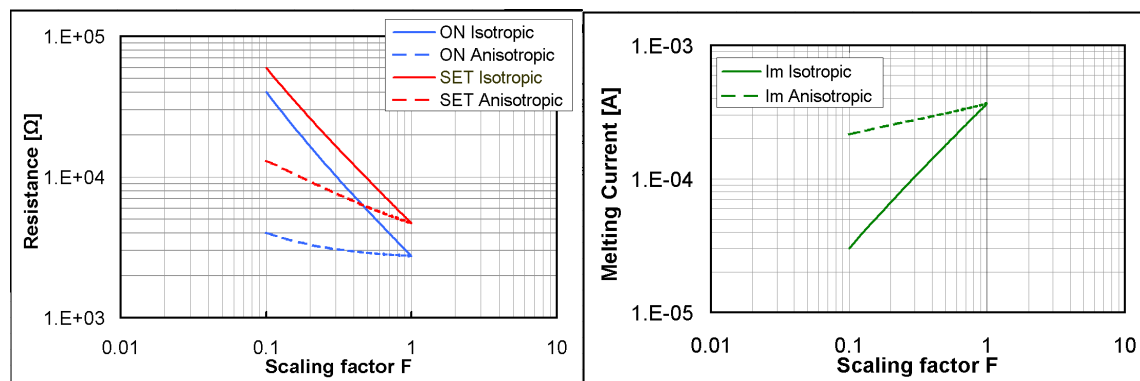


Figura 1.18: Prestazioni di scaling delle resistenze di ON e di SET (sinistra) e della corrente di fusione (destra) nei casi di scaling isotropo e anisotropo.

no con un fattore F uguale per tutte, rendendo più difficoltoso il calcolo analitico. Come esempio si porterà il caso in cui scalano altezza e larghezza della cella, mentre lo spessore t rimane invariato. Per questo approccio si ottiene quindi:

$$R_{ON} \propto \frac{aF + b}{F} \quad (1.50)$$

$$R_{SET} \propto \frac{aF + b + c \ln(1 + dF)}{F} \quad (1.51)$$

$$I_m \propto \sqrt{\frac{F}{b + aF} \left(\frac{1}{a} + cF \ln(1 + dF) \right)} \quad (1.52)$$

Come si osserva in figura 1.18, l'aumento di resistenza è più contenuto, in quanto dipende unicamente dalla resistenza d'interfaccia: la resistenza dell'heater rimane infatti costante perchè il rapporto lunghezza/larghezza rimane invariato. Le resistenze di ON e SET hanno quindi un andamento asintotico pari a F^{-1} . La dipendenza logaritmica all'interno dell'espressione di R_{SET} è trascurabile fino a valori di F molto bassi (tali che $2nL_c < t$), oltre i quali tende ad abbattere la resistenza del GST. Nel grafico I_m - R_{SET} la pendenza bilogaritmica è di -0.51, indicando che la resistenza di lettura aumenta più rapidamente di quanto diminuisca la corrente di fusione, peggiorando quindi le prestazioni del dispositivo. La spiegazione di questo andamento va cercata nel diverso scaling delle resistenze termiche

nel caso isotropo e anisotropo; infatti nel caso isotropo sia R_h^i che R_c^i scalano di un fattore F^{-1} . Essendo le due resistenze in parallelo, questo significa che anche diminuendo le dimensioni geometriche della cella, il flusso di calore dissipato si suddivide allo stesso modo tra heater e calcogenuro, per cui il profilo di temperatura è sostanzialmente dettato dall'entità del riscaldamento Joule all'interfaccia, a sua volta dominato dalla dissipazione nell'heater. Con lo scaling anisotropo, invece, la resistenza termica del riscaldatore tende a rimanere costante, mentre quella del GST scala come F^{-1} . Il calore tenderà quindi a scorrere attraverso la resistenza termica più bassa, portando a una diminuzione dell'efficienza termica della cella. Infatti si arriverà al punto in cui tutto il calore generato all'interfaccia viene dissipato attraverso l'heater, e non si ha più miglioramento della resistenza termica del GST con il semplice scaling delle dimensioni geometriche. Alla luce di quanto appena detto, è chiaro che lo scaling isotropo fornisce prestazioni migliori alla cella scalata, dato che a parità di resistenza di lettura è caratterizzato da una corrente di programmazione più bassa. Tuttavia questo approccio non è possibile per motivi tecnologici. Nei capitoli successivi sarà analizzato in dettaglio lo scaling reale delle celle sperimentali prese in esame, che non ricade in nessuno dei due casi appena analizzati, trattandosi di uno scaling anisotropo con fattori di scala differenti per le diverse dimensioni geometriche. Inoltre, questi fattori sono anche differenti a seconda del nodo tecnologico di riferimento, per cui le simulazioni numeriche sono utili per prevedere l'andamento delle celle scalate.

Per migliorare le prestazioni del dispositivo è anche possibile lavorare sull'ottimizzazione dei materiali. Nell'ambito di interesse, essi vengono descritti principalmente dalla resistività elettrica e dalla conducibilità termica. La resistività dell'heater ad esempio ha un effetto diretto sulla resistenza ad alti campi, oltre che sulla R_{SET} , e aumentandola si ottiene una riduzione della corrente di programmazione. La resistività della lega calcogenura, secondo il modello proposto, non ha un'influenza rilevante sulla corrente di melting, per cui è opportuno cercare di ridurla per limitare la resistenza di readout. Tuttavia, secondo la legge di Wiedemann-Franz questo provoca anche un incremento della conducibilità termica, modificando negativamente la resistenza termica del GST. Per questa ragione, se si segue questo approccio è anche necessario ridurre la conducibilità termica per non degradare eccessivamente il confinamento termico della cella.

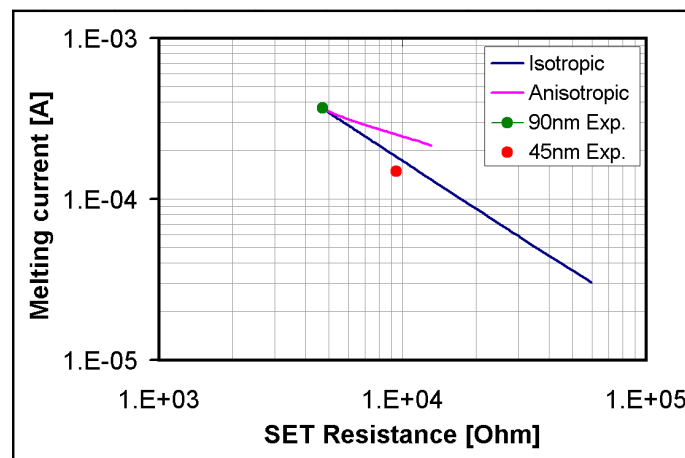


Figura 1.19: Andamento di I_m in funzione di R_{SET} per i due casi di scaling isotropo e anisotropo presi in considerazione. I punti sperimentali sono stati ottenuti con un approccio di scaling ancora differente, che tuttavia riporta risultati più simili al caso isotropo.

Capitolo 2

Analisi dei dati sperimentali e calibrazione del modello

Nel precedente capitolo sono state illustrate le proprietà fondamentali e i principi di funzionamento della cella di memoria a cambiamento di fase. È evidente che un'accurata caratterizzazione del dispositivo deve tenere conto di numerosi fenomeni, di natura sia elettrica che termica, tra loro interagenti. Ad esempio, durante il funzionamento la temperatura nell'area attiva può passare dai 300 gradi Kelvin della temperatura ambiente al migliaio di gradi richiesto per fondere il materiale. Una simile escursione termica ha inevitabilmente ripercussioni sul trasporto elettrico, in particolare sulla concentrazione e la mobilità dei portatori, e sulla conduzione termica, ad esempio per via della dipendenza della conducibilità termica dal numero di portatori liberi. Una stima quantitativa delle prestazioni del dispositivo (corrente di fusione e finestra di resistenza) è impossibile senza l'ausilio di una simulazione numerica che prenda in considerazione tutti i fenomeni citati. Grazie a questa simulazione è possibile non solo perfezionare i modelli esistenti, ma anche studiare le proprietà di celle con geometrie e fattori di scaling non ancora disponibili per un'analisi sperimentale. In questo capitolo verrà prima descritto il modello numerico utilizzato, poi verranno analizzati dei dati sperimentali riguardanti dispositivi prodotti presso Numonyx, grazie ai quali sarà possibile calibrare i parametri dei materiali da inserire nel simulatore.

2.1 Il modello elettrotermico

2.1.1 Fase cristallina

La simulazione avviene utilizzando il software SDevice di Synopsys. Questo programma risolve mediante il metodo degli elementi finiti una serie di equazioni differenziali su una griglia tridimensionale. In ogni punto i termini noti delle equazioni variano per via della dipendenza spaziale e dalla temperatura delle caratteristiche dei materiali. Le equazioni da risolvere sono le classiche equazioni del modello drift-diffusion per i semiconduttori:

$$\nabla \cdot (\varepsilon \nabla \phi) = -q(p - n + N_d - N_a) - \rho_{trap} \quad (2.1)$$

$$\nabla \cdot J_n = qR_{net} + \frac{dn}{dt} \quad (2.2)$$

$$\nabla \cdot J_p = qR_{net} + \frac{dp}{dt} \quad (2.3)$$

$$J_n = -nq\mu_n \nabla \phi_n \quad (2.4)$$

$$J_p = -nq\mu_p \nabla \phi_p \quad (2.5)$$

e l'equazione di trasporto del calore:

$$\frac{d}{dt} C_L T - \nabla \cdot (k \nabla T) = -\nabla \cdot (\phi_n J_n + \phi_p J_p) \quad (2.6)$$

La 2.1 è l'equazione di Poisson, in cui ε è la costante dielettrica del materiale, ϕ il potenziale elettrostatico, p ed n la densità di lacune ed elettroni rispettivamente, N_d ed N_a la densità di droganti ionizzati, ρ_{trap} la densità di trappole. Le 2.2 e 2.3 sono le equazioni di continuità di elettroni e lacune, in cui J è la densità di corrente e R_{net} il tasso di ricombinazione netto. Le 2.4 e 2.5 esprimono la corrente in funzione del campo elettrico in base al modello drift-diffusion, e μ_n e μ_p sono le mobilità di elettroni e lacune, ϕ_n e ϕ_p i rispettivi livelli di quasi-Fermi. Nella 2.6 C_L è la capacità termica e k la conducibilità termica del materiale. Molti di questi termini dipendono dalla temperatura, per cui queste equazioni devono essere risolte a sistema tra loro. Come illustrato nel capitolo precedente, ai fini della conduzione nella fase cristallina il contributo degli elettroni è trascurabile rispetto

a quello delle lacune, per cui in seguito si illustreranno per semplicità solo le equazioni relative a queste ultime. Nell'equazione di Poisson, la densità di portatori viene esplicitata facendo uso della statistica di Fermi-Dirac applicata al modello a bande della struttura. Non viene usata l'approssimazione di Maxwell-Boltzmann in quanto sia il GST nella fase cristallina che l'heater si comportano come semiconduttori degeneri o quasi-degeneri, e il livello di Fermi giace molto vicino alla banda di valenza. Il simulatore quindi risolve numericamente l'integrale di Fermi-Dirac

$$p = N_v \int_0^{\infty} \frac{\sqrt{E}}{1 + \exp\left(\frac{E - E_{f,p}}{k_b T}\right)} dE \quad (2.7)$$

dove N_v è un parametro fornito dall'utente, ed è stato calibrato in base ai dati reperiti in letteratura. Con un procedimento simile, il programma calcola anche la densità di stati trappola occupati ρ_{trap} . Il livello di Fermi si posiziona nei pressi della banda di valenza, e la distanza dipende dalla densità di trappole associate alle vacanze strutturali e dal loro livello energetico. È quindi possibile calibrare l'energia di attivazione della conducibilità della fase cristallina agendo sulle caratteristiche degli stati trappola. Nelle equazioni del modello drift-diffusion compare la mobilità dei portatori. Si è deciso di adottare un modello a mobilità costante per il GST cristallino, in quanto i dati sperimentali raccolti e reperiti in letteratura fanno riferimento alla resistività totale del GST, senza separare i contributi di mobilità e densità di portatori. Perciò per semplificare la calibrazione si impone un valore di mobilità costante entro una finestra di accettabilità e si riconduce alla concentrazione di portatori la totalità delle variazioni di conducibilità del materiale. Nella 2.6 la conducibilità termica k ha un impatto significativo sul profilo spaziale di temperatura, e quindi su tutte le caratteristiche della cella, e sarà uno dei più importanti parametri da calibrare. Nel modello numerico si tiene anche conto della dipendenza di k dalla concentrazione di portatori disponibili. Infatti il calore, oltre a essere trasmesso mediante le vibrazioni degli atomi reticolari (trasmissione fononica del calore) può anche essere portato dai portatori liberi (trasmissione elettronica del calore). Questo fenomeno viene descritto mediante la legge empirica detta di Wiedemann-Franz, formulata inizialmente per i metalli ma osservata

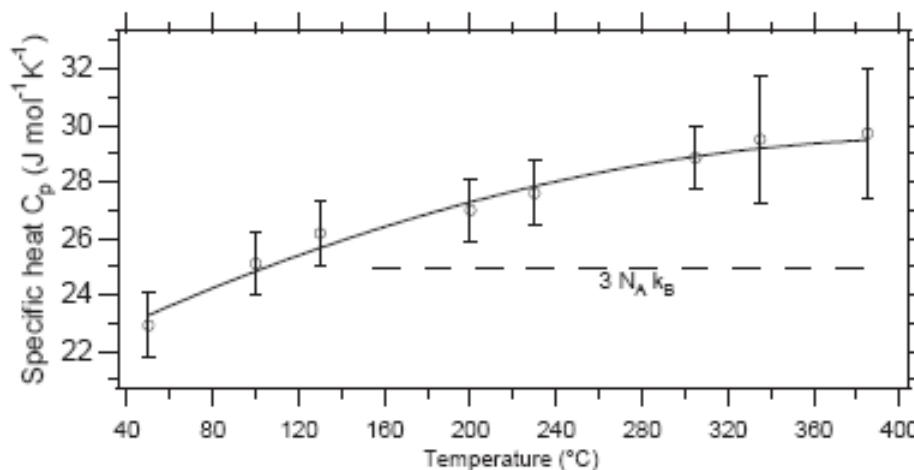


Figura 2.1: Capacità termica molare del $Ge_2S B_2Te_5$ all'interno del range termico di funzionamento tipico. Questo valore può essere usato per stimare i tempi caratteristici di risposta termici del calcogenuro

anche nei calcogenuri. Questa legge descrive il fenomeno mediante la formula

$$K = K_0 + \sigma \frac{\pi^2}{3} \left(\frac{k_b}{q} \right)^2 T \quad (2.8)$$

dove K_0 è un termine indipendente dalla temperatura, σ è la conducibilità elettrica e K_b la costante di Boltzmann. Si avrà quindi che la conducibilità termica sarà più alta nei pressi dell'interfaccia che nel resto del calcogenuro. La capacità termica C_L ha importanza solamente nei transistori, e non entra in gioco in situazioni stazionarie, in cui $dT/dt = 0$. La maggior parte delle simulazioni che verranno effettuate in questo lavoro saranno in regime stazionario, per cui il valore numerico di C_L non avrà alcun effetto. Questa scelta è giustificata dal fatto che i tempi di risposta termici caratteristici del dispositivo sono molto rapidi rispetto al tempo in cui gli impulsi di programmazione si mantengono al loro valore massimo, rendendo quindi possibile un'analisi in regime quasi-stazionario, per la quale ad ogni punto della rampa di tensione la cella riesce a portarsi nel suo stato di equilibrio. A conferma di questa ipotesi, la figura 2.1[9] mostra alcune misure sperimentali della capacità termica del calcogenuro, reperite in letteratura. Una tipica cella del nodo

tecnologico a 45nm ha un'area attiva di 250 nm^2 e un'altezza di GST di circa 50nm. Il tempo caratteristico di risposta dei fenomeni termici è dato da

$$\tau_{th} = R_{th}C_L \quad (2.9)$$

dove R_{th} è la resistenza termica del calcogenuro, e con valori tipici di conducibilità termica e il volume sopra riportato, è dell'ordine di $5 \cdot 10^7 \text{ KW}^{-1}$. Inserendo i valori numerici si ottiene un tempo caratteristico di alcuni nanosecondi, mentre gli impulsi di programmazione e di lettura hanno una zona di plateau che può durare dalle decine di nanosecondi a qualche microsecondo, per cui la simulazione DC fornisce una ragionevole approssimazione di questi impulsi. Le simulazioni che invece terranno conto dei transistori (in particolare quelle riguardanti i fenomeni di ricristallizzazione parziale) utilizzeranno valori di C_L ricavati dalla letteratura.

Le condizioni a contorno delle equazioni differenziali sono quelle dei contatti elettrici e termici. In questa simulazione si suppone che i contatti siano ohmici e ideali, per cui non si hanno termini aggiuntivi dovuti alla ricombinazione superficiale o altri effetti del secondo ordine, e le condizioni a contorno si riducono alla neutralità di carica e alle condizioni di equilibrio, ovvero

$$n_0 - p_0 = N_D - N_A n_0 p_0 = n_i^2 \quad (2.10)$$

Per quanto riguarda i contatti termici, si usano delle condizioni di Dirichlet semplici ($T=T_0$). In entrambi i casi, sulle superfici dove non sono definiti contatti si impongono condizioni a contorno di Neumann, ovvero riflesse, del tipo $\nabla f = 0$. Sulle superfici in cui si hanno cambiamenti di materiali, o comunque di costanti elettriche e termiche, vengono imposte le condizioni di continuità delle variabili (potenziale e temperatura) e dei flussi (campo elettrico e flusso di calore). La figura 2.2 mostra una sezione della struttura 3D usata per la simulazione della cella. Il taglio è stato effettuato lungo il piano verticale ($x=0$) a metà del dispositivo, in quanto esso è simmetrico rispetto a questo piano. I colori trasparenti rappresentano i materiali isolanti (ossidi e nitruri). La struttura è stata disegnata tenendo conto del processo tecnologico per quanto riguarda la posizione e le dimensioni degli isolanti. In rosso viene rappresentato il calcogenuro, mentre il blu contraddistingue il riscaldatore (heater). Si può osservare la forma ad L dell'heater caratteristica del processo

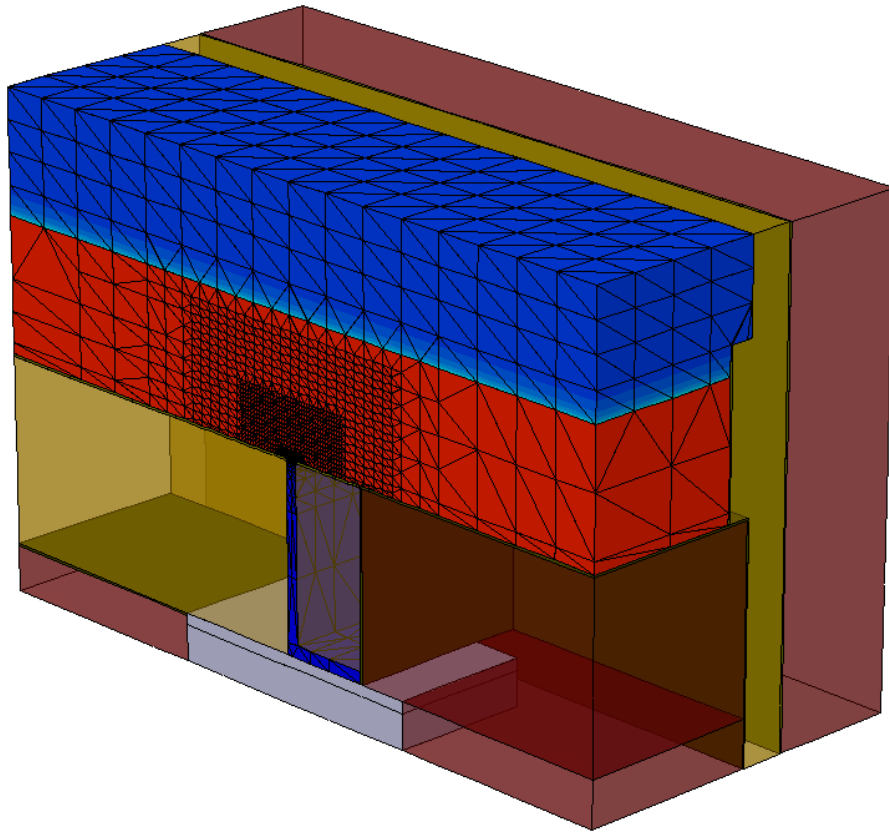


Figura 2.2: Struttura della cella PCM usata per la simulazione. I materiali isolanti sono trasparenti.

Wall di Numonyx. Sopra al GST e sotto il riscaldatore sono posti due strati di materiale conduttivo che fungono da contatto: all'interno di questi non si ha praticamente né caduta di tensione né gradiente di temperatura. Nella stessa immagine è anche mostrata la griglia a elementi finiti usata per la simulazione, particolarmente fitta nei pressi dell'interfaccia GST/heater in quanto è qui che avvengono i fenomeni più critici per il funzionamento della PCM.

Una particolarità di questo modello sono le interfacce termiche tra GST e isolanti. Il simulatore SDevice non ne tiene conto autonomamente, tuttavia in letteratura è stato dimostrato che la resistenza aggiuntiva da esse introdotta non è trascurabile rispetto a quella di bulk, e quindi risulta molto importante tenerne conto. Per questo motivo sono stati inseriti degli

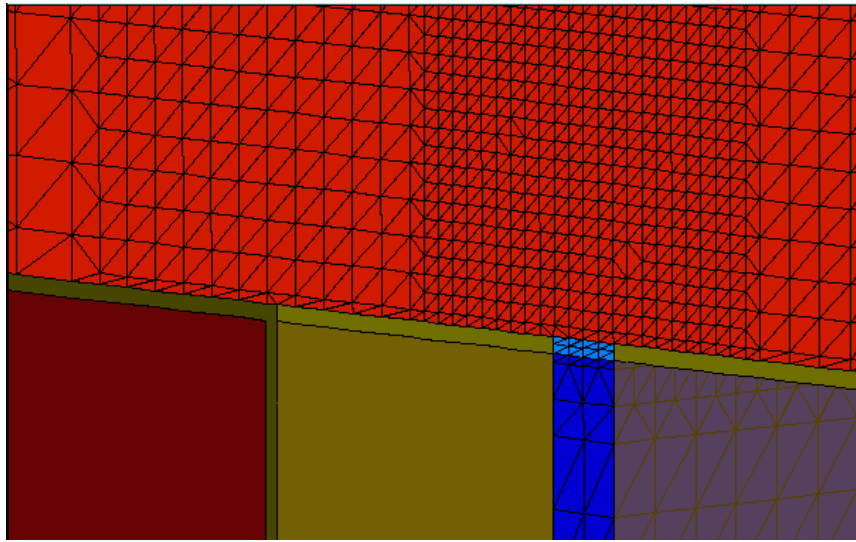


Figura 2.3: Particolare della struttura usata per la simulazione, in cui vengono evidenziate: in verde le interfacce termiche di spessore 1nm tra GST e i materiali isolanti; in blu chiaro la presenza di un'interfaccia elettrica di spessore 1nm tra GST e riscaldatore, che serve a rappresentare un'ulteriore resistenza in serie all'heater. Entrambi i tipi di interfaccia sono importanti per la corretta descrizione del comportamento della cella in temperatura.

strati di spessore molto ridotto (1nm), ai quali è stata assegnata una conducibilità termica ridotta rispetto agli isolanti circostanti. Tenere conto delle interfacce è fondamentale per comprendere correttamente il funzionamento del dispositivo. In prima approssimazione, esse hanno l'effetto di avvicinare le condizioni al contorno di tipo Neumann, aumentando notevolmente il confinamento del calore all'interno della cella. Nella figura 2.3 in verde sono evidenziate queste interfacce nella struttura. Nella figura 2.4 invece è illustrato il profilo di temperatura di una cella attraversata da una corrente fissata, con presenza o meno dell'interfaccia. Sono colorate solamente le temperature superiori a T_m . La cella a sinistra non utilizza le interfacce termiche, quindi il calore può essere ben dissipato attraverso gli ossidi e i nitruri circostanti, e si avrà quindi una temperatura massima più bassa. Nella cella a destra sono presenti gli stati interfacciali, e l'area attiva si scalda molto di più. Si nota inoltre come il profilo di temperatura ha una brusca discontinuità in prossimità delle interfacce. Il valore numerico della conducibilità termica di questi strati è indicato con K_{if} e verrà calibrato nelle sezioni successive sulla base di misure sperimentali effettuate

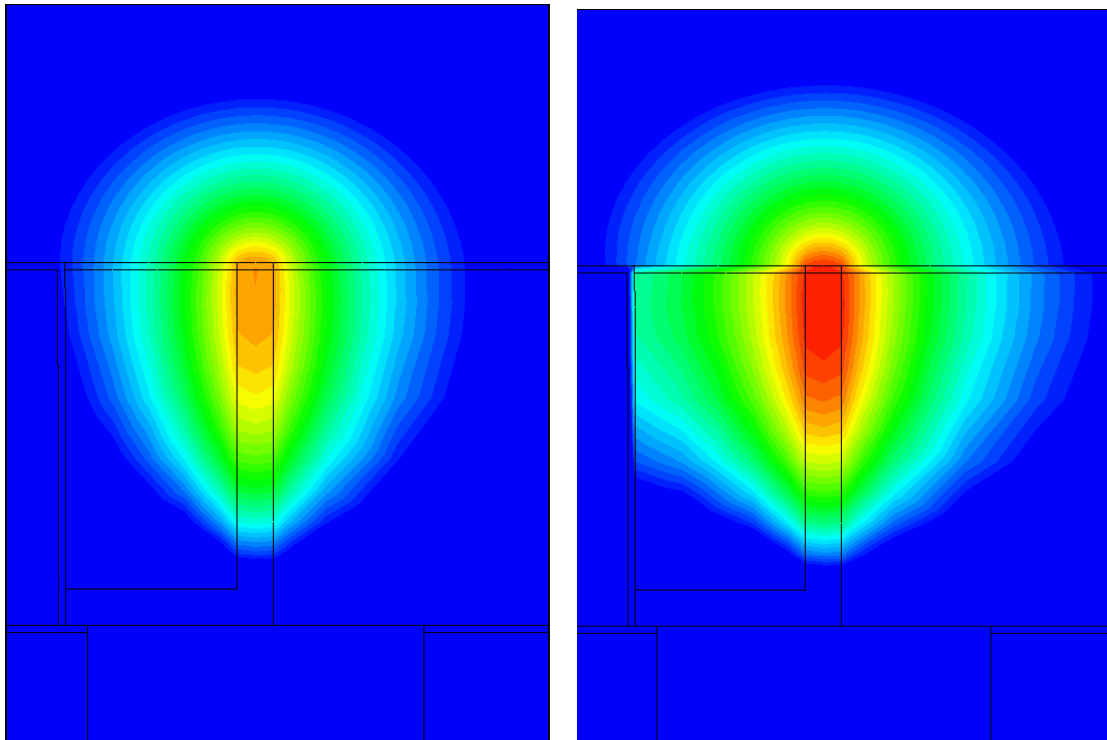


Figura 2.4: Profilo di temperatura all'interno della stessa struttura, attraversata dalla medesima corrente, in due casi: assenza (sinistra) e presenza (destra) dell'interfaccia. Questa ha l'effetto di confinare il calore all'interno del GST.

presso la Numonyx. Come verrà mostrato in seguito, le interfacce termiche sono indispensabili per riprodurre correttamente le curve di programmazione R-I della cella PCM. Oltre alle interfacce termiche, nel modello è anche presente un'interfaccia elettrica tra GST e heater, mostrata in figura 2.3, e rappresenta una resistenza aggiuntiva rispetto a quella dell'heater. Anch'essa verrà inserita nel simulatore mediante un sottile strato (1 nm) di materiale identico a quello dell'heater ma dotato di una mobilità ridotta. La presenza di questa resistenza di contatto è evidenziata dai dati sperimentali che verranno mostrati in seguito, ed è necessaria per simulare correttamente la cella, soprattutto per quanto riguarda la dipendenza dalla temperatura della corrente portata dal dispositivo. Alla struttura appena descritta viene applicata quindi la rampa di tensione quasi-stazionaria. Lo scopo è quello di ottenere la caratteristica I-V della cella nella fase cristallina e il profilo di tempe-

ratura al variare della tensione e della corrente di programmazione. Il profilo di temperatura viene poi utilizzato dal calcolatore per elaborare l'algoritmo di cambiamento di fase. Questo permette di simulare la creazione dello strato di amorfo all'interno del dispositivo, preparandolo per le simulazioni successive, che riguarderanno la fase amorfa. Questa prima simulazione è molto importante, in quanto molte delle cifre di merito della cella si riferiscono alla fase cristallina: a basse tensioni la corrente letta è quella che determina il livello di resistenza associato allo stato di "SET" della memoria, alle alte tensioni la corrente è quella necessaria alla fusione del materiale, ed ha un notevole impatto sul consumo energetico e le prestazioni della memoria. È pertanto su questa fase che si concentrerà la maggior parte del lavoro di calibrazione.

2.1.2 Cambiamento di fase e fase amorfa

Dopo aver effettuato una rampa di tensione ai capi della cella, abbiamo ottenuto una curva I-V del dispositivo. Ogni punto di questa curva corrisponde ad un impulso di programmazione di durata molto lunga rispetto ai tempi di risposta termici e di cambiamento di fase, per via delle ipotesi di quasi-stazionarietà enunciate in precedenza. Ad ognuno di essi è inoltre associato il profilo di temperatura all'interno del dispositivo. All'interno del simulatore è stata inserita una variabile ausiliaria che indica in che fase si trova un dato punto della griglia a elementi finiti. A ogni punto che nel corso della simulazione ha raggiunto la temperatura di fusione T_m , che vale 888 K per il composto $Ge_2Sb_2Te_5$, viene assegnato un valore corrispondente alla fase amorfa. In questo modo ricaviamo una struttura contenente una calotta di materiale vetroso, le cui dimensioni dipendono dall'ampiezza dell'impulso di programmazione. In figura 2.5 viene mostrato il valore massimo della temperatura all'interfaccia tra GST e heater, ricavate dalle simulazione, e viene confrontato con il modello analitico, nella forma:

$$T = T_0 + I^2 R_{th} R_{ON} \alpha \quad (2.11)$$

Si osserva una discrepanza tra il modello e la simulazione, dovuto alla dipendenza del calore generato dalla temperatura: infatti all'aumentare della corrente I la temperatura

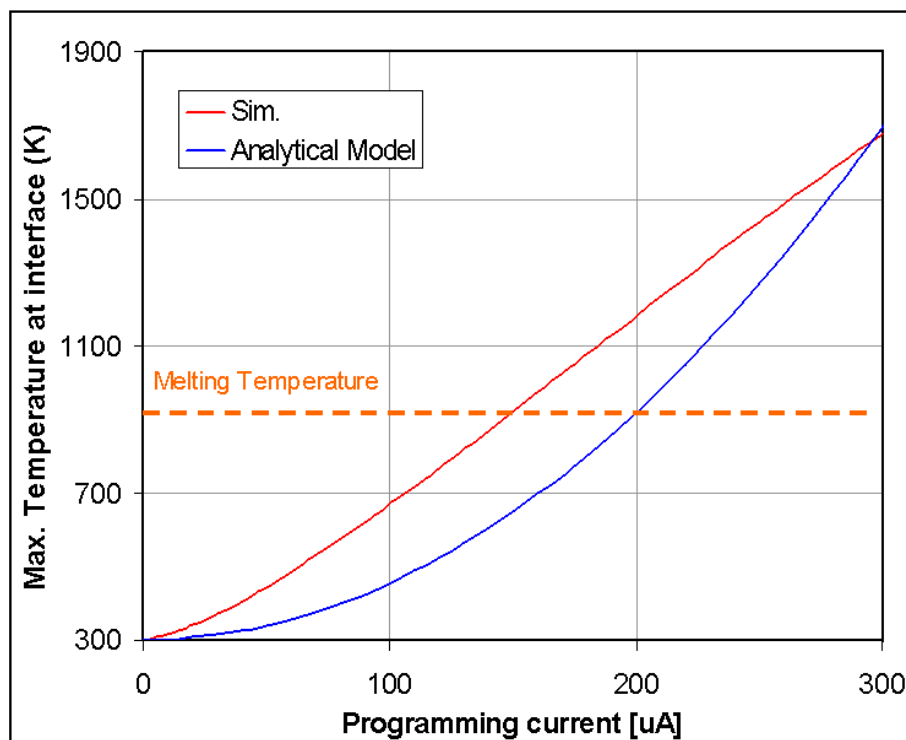


Figura 2.5: Simulazione e calcolo analitico della temperatura massima raggiunta dall'interfaccia GST/heater in funzione dell'ampiezza dell'impulso di RESET applicato. La simulazione si discosta notevolmente dal modello analitico (quadratico) per via della dipendenza dalla temperatura di tutti i parametri dei materiali in gioco. Ogni punto che nel corso della simulazione supera il valore $T_m = 888$ viene considerato fuso e raffreddato, e alla simulazione successiva viene inizializzato come amorfo.

crece e provoca a sua volta un'abbassamento della resistenza della cella. L'equilibrio di questi due fenomeni porta a una dipendenza di T dalla corrente inferiore a quella quadratica prevista dal modello analitico.

Un impulso che riesce a portare la temperatura dell'interfaccia sopra la T_m fa fondere una parte del calcogeno. Il modello poi assume che il quenching sia talmente rapido da vitrificare la totalità del materiale che ha raggiunto la fase liquida, per cui tutti i punti della griglia che si trovano a una temperatura superiore agli 888 K vengono segnalati come amorfi dal simulatore. A questo punto viene ripetuta la rampa di tensione, per ottenere la caratteristica I-V della cella programmata. Il simulatore qui risolve numericamente le

stesse equazioni descritte per la fase cristallina, con la differenza che nei punti dove è presente l'amorfo viene utilizzata la struttura di Adler, cioè con la presenza degli stati trappola relativi ai siti C_3^+ e C_1^- e alla coda di Urbach. Inoltre questi punti della griglia hanno tutte le caratteristiche del materiale amorfo: una diversa conducibilità termica, mobilità dei portatori e gap energetico.

Per simulare il fenomeno dello switching di soglia è poi necessario inserire nel modello un meccanismo di ricombinazione e uno di generazione di portatori come descritto in precedenza. Per quanto riguarda la ricombinazione è stato usato il modello Shockley-Read-Hall (SRH). Esso prevede un tasso di ricombinazione legato agli stati trappola che presente le caratteristiche richieste dal modello numerico descritto nel capitolo precedente. Poiché la presenza di un'elevata densità di stati trappola in un semiconduttore amorfo è non solo attesa, ma ampiamente riportata in letteratura, la scelta del modello SRH è ragionevole anche dal punto di vista fisico, oltre che da quello numerico. Il tasso di ricombinazione netto è pertanto espresso dalla formula

$$R_{net}^{SHR} = \frac{np - n_{i,eff}^2}{\tau_h(n + n_1) + \tau_e(p + p_1)} \quad (2.12)$$

in cui τ_h e τ_e sono i tempi di vita dei portatori e $n_{i,eff}$ la densità intrinseca di portatori. I parametri n_1 e p_1 si possono scrivere come

$$n_1 = n_{i,eff} e^{\frac{E_{traps}}{k_b T}} \quad (2.13)$$

$$p_1 = n_{i,eff} e^{\frac{-E_{traps}}{k_b T}} \quad (2.14)$$

dove E_{traps} è il livello energetico delle trappole. Per quanto riguarda il modello di generazione dei portatori, il modello dello switching di soglia richiede che il tasso di generazione sia una funzione monotona della densità di portatori e del campo elettrico. La ionizzazione a impatto possiede questa caratteristica, e viene pertanto implementata nel simulatore mediante il modello empirico di Okuto-Crowell. In base a questo modello, il tasso di generazione di portatori viene descritto dalla forma

$$G = A_n n v_n + A_p p v_p \quad (2.15)$$

dove v_n e v_p sono le velocità di deriva dei portatori e $A_{n,p}$ sono i coefficienti di ionizzazione, che si esprimono con la formula

$$A_{n,p} = a[1 + c(T - T_0)] \cdot |E|^\gamma e^{(-\frac{b(1+d(T-T_0))}{|E|})^\delta} \quad (2.16)$$

in cui $T_0=300$ K, $\gamma=1$, $\delta=2$ e a,b,c sono dei parametri da calibrare. Per semplificare la calibrazione, come già accennato in precedenza, si è ipotizzato che la mobilità dei portatori nel calcogenuro amorfo sia costante al variare della temperatura e del campo elettrico. Questa ipotesi trae fondamento dall'utilizzo che si vuole fare del modello. Infatti, questo lavoro si concentrerà principalmente sulla simulazione delle curve di programmazione della cella nel passaggio dalla fase di SET alla fase di RESET. In queste condizioni, l'operazione principale da svolgere sul materiale vetroso è la lettura della corrente, per determinare la finestra di resistenza tra le due fasi. Questa lettura avviene applicando alla cella tensioni inferiori a 400mV, ben al di sotto delle tipiche tensioni che portano allo switching. Applicando tensioni così ridotte alla cella nello stato di RESET il calore generato è molto basso per via dell'elevata resistenza dell'amorfo, per cui gli effetti di temperatura in questa fase vengono trascurati. La figura 2.6 mostra un esempio di simulazione numerica della curva I-V dell'amorfo, in cui si può osservare lo snap-back della caratteristica. Utilizzando la medesima struttura è stata anche lanciata una simulazione con la moltiplicazione a valanga disattivata, e in questo caso non si osserva alcuno switching della curva I-V, a conferma della natura essenziale della generazione per ionizzazione a impatto nella corretta simulazione numerica del GST amorfo. In entrambi i casi il modello riesce comunque a prevedere le due zone, lineare e ed esponenziale, della caratteristica. Il dato più importante ai fini di questo lavoro è il valore assunto dalla corrente nella zona esponenziale intorno ai 0.4V. È infatti a questo valore di tensione che verrà letta la resistenza della cella con la quale si procederà poi a costruire la curva di programmazione.

2.2 Calibrazione del modello mediante i dati sperimentali

La corretta calibrazione del modello a elementi finiti richiede l'inserimento nel simulatore di numerosi parametri, come visto nella sezione precedente. Uno degli scopi principali

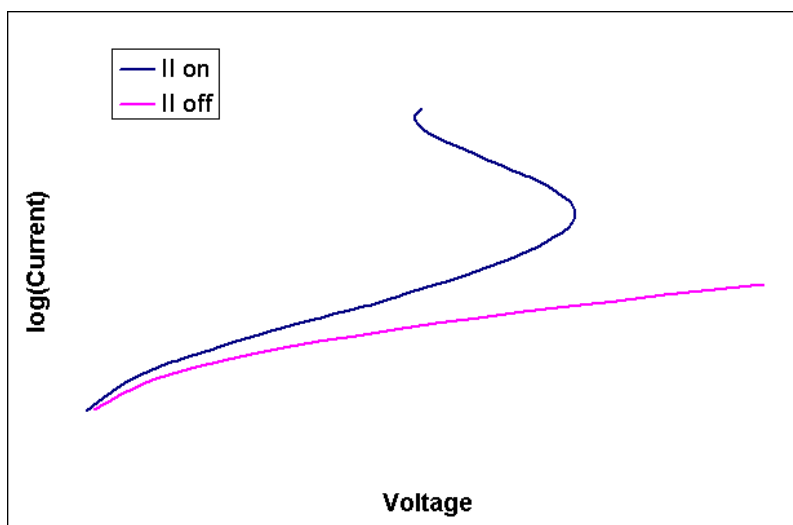


Figura 2.6: Caratteristiche I-V di una generica cella PCM nello stato di RESET. a) Modello completo con ionizzazione a impatto, è presente lo switching di soglia. b) Modello senza ionizzazione a impatto, lo switching non avviene.

di questo lavoro è appunto quello di ricavare un insieme di parametri che sia in grado di riprodurre correttamente il funzionamento di celle PCM appartenenti a diversi nodi tecnologici a partire dai principi primi fisici che la governano. I parametri saranno ricavati sia da dati sperimentali ottenuti presso i laboratori Numonyx, sia dalla letteratura disponibile. Questo lavoro si concentrerà sulle proprietà dei materiali in gioco, in particolare:

- **Resistività dell'heater** (ρ_h), che determina il funzionamento della cella ad alti campi elettrici applicati e l'entità del riscaldamento Joule. Il riscaldatore viene descritto come un semiconduttore quasi-degenere il cui drogaggio, di tipo p, è noto a priori. Per determinarne la resistività sarà dunque necessario inserire nel simulatore un valore di mobilità delle lacune (μ_h) corretto. È anche necessario calibrarne la dipendenza dalla temperatura per riprodurre correttamente le curve sperimentali.
- **Resistività del GST** (ρ_c) e la sua dipendenza dalla temperatura. Il GST dà il suo contributo alle caratteristiche della cella in ogni fase del suo funzionamento, eccetto per tensioni applicate molto elevate, ed è quindi importante modellizzarlo correttamente. Per determinarne la resistenza si agisce sulla sua energia di attivazione E_a ,

regolata dalla distanza del livello trappole dalla banda di valenza, e sulla mobilità μ_c .

- **Resistività specifica dell'interfaccia GST/heater** (ρ_{if}^{\square}) e sua dipendenza dalla temperatura. L'esistenza di questa resistenza è stata dimostrata in letteratura, e influisce sulla quantità di calore generato all'interfaccia. I suoi effetti inoltre tendono a diventare più evidenti allo scalare delle dimensioni della cella. Nel modello viene rappresentata da uno strato sottile (1nm) di materiale identico a quello dell'heater, di cui si calibrerà la mobilità efficace (μ_{if}) in modo da ottenere i valori di resistenza elettrica d'interfaccia (EBR, *electrical boundary resistance*) che si ritrovano nei dati sperimentali.
- **Conducibilità termica del GST** (K_c), importante per la generazione di calore all'interno del GST, ha un effetto evidente sulle curve di programmazione RI. La sua dipendenza dalla temperatura è data dalla legge di Wiedemann-Franz riportata nel capitolo precedente.
- **Resistenze d'interfaccia termica** (TBR , *thermal boundary resistance*). Misure sperimentali dimostrano che le interfacce termiche tra GST e materiali isolanti e GST e heater contribuiscono in maniera non trascurabile alla resistenza termica totale, ed è pertanto necessario tenerne conto, inserendole nella struttura mediante sottili strati (1nm) di materiale isolante con una conducibilità termica effettiva (K_{if}) da calibrare in base ai dati.

Poiché la tesi si concentrerà principalmente sulla fase cristallina della cella, la calibrazione di parametri che hanno effetto sulla conduzione nell'amorfo e sullo switching di soglia, come le caratteristiche delle trappole nell'amorfo e i parametri di generazione e ricombinazione di portatori, saranno trattati a margine servendosi principalmente di dati reperiti in letteratura.

Per calibrare i parametri ci si servirà, inizialmente, del modello analitico applicato ai dati sperimentali. Come si vedrà, il modello non è in grado di descrivere correttamente tutti i fenomeni in gioco nel corso dell'operazione della cella PCM, tuttavia offre un utile punto di partenza per la calibrazione. Una volta ricavati i parametri si procederà a eseguire

una serie di simulazioni numeriche per migliorare le scelte effettuate in precedenza. La maggior parte dei dati viene ricavata da misure eseguite su *cella analitica*, ovvero un dispositivo PCM privo del transistor di selezione e della circuiteria che accompagna un array di memoria. Pur non disponendo delle funzionalità di un blocco di memoria, l'assenza del transistor permette un'analisi più accurata delle caratteristiche intrinseche della cella. A questo punto si procederà con la stima delle proprietà dei materiali sopra elencate partendo dai dati sperimentali.

2.2.1 La cella standard

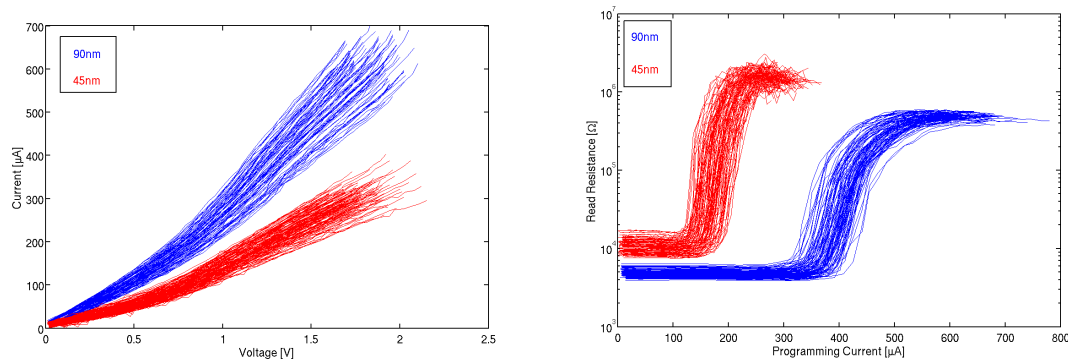


Figura 2.7: Caratteristiche IV e RI di circa 200 celle standard, per i nodi 90nm e 45nm.

Tutte le celle PCM che verranno analizzate appartengono ai nodi tecnologici 90nm e 45nm. La figura 2.7 mostra le loro caratteristiche I-V ed R-I, mentre la tabella 2.1 riporta le dimensioni nominali che caratterizzano i dispositivi dei due nodi tecnologici, insieme alle proprietà di rilievo ricavate dalle curve. Questi valori sono un utile riferimento per avere ben presente i valori di resistenza e corrente di fusione che caratterizzano una normale cella PCM. Si può anche notare l'andamento di queste caratteristiche al variare del nodo tecnologico, aspetto che verrà trattato in dettaglio più avanti. Queste curve vengono ricavate applicando impulsi di tensione, come descritto nel capitolo precedente. Questo setup sperimentale impulsato tuttavia non è in grado di leggere correnti al di sotto di una certa soglia. È per questa ragione che le curve RI di ogni cella saturano allo stesso valore

	90nm	45nm
W_h [nm]	160	58
W_c [nm]	140	50
L_h [nm]	85	50
L_c [nm]	70	50
t [nm]	5	5
R_{SET} [Ω]	4971	10943
R_{ON} [Ω]	2567	4659
I_m [μA]	365	148

Tabella 2.1: Dimensioni geometriche e parametri caratteristici delle celle standard 90nm e 45nm.

di resistenza: infatti la corrente portata dalla cella nella fase amorfa è molto bassa (nA), soprattutto a bassi campi applicati, quando il dispositivo è ancora in regime lineare. Per poter misurare correttamente la resistenza nella fase di RESET è necessario ricorrere a un setup che effettui misure in corrente continua, restando però sotto la tensione di switching del dispositivo, per non riportarlo inavvertitamente nella fase di SET. Questo genere di misura è stata effettuata solamente su un paio di celle standard, la cui curva RI è riportata in figura 2.8.

2.2.2 Resistività dell'heater

La resistività dell'heater può essere ricavata grazie a misure di R_{ON} e R_{SET} effettuate su celle che differiscono tra loro unicamente per la lunghezza dell'heater L_h . Infatti, secondo il modello analitico, si ha

$$R_{SET} = R_{GST} + R_{if}^{lf} + \rho_h^{lf} \cdot \frac{L_h}{A_h} \quad (2.17)$$

$$R_{ON} = R_{if}^{hf} + \rho_h^{hf} \cdot \frac{L_h}{A_h} \quad (2.18)$$

in cui gli apici hf e lf (high-field e low-field) indicano che i valori di resistenza sono presi rispettivamente ad alti e bassi campi. Infatti la temperatura nel dispositivo aumenta con la tensione applicata, e con essa si riduce la resistività dei materiali. Si ricava

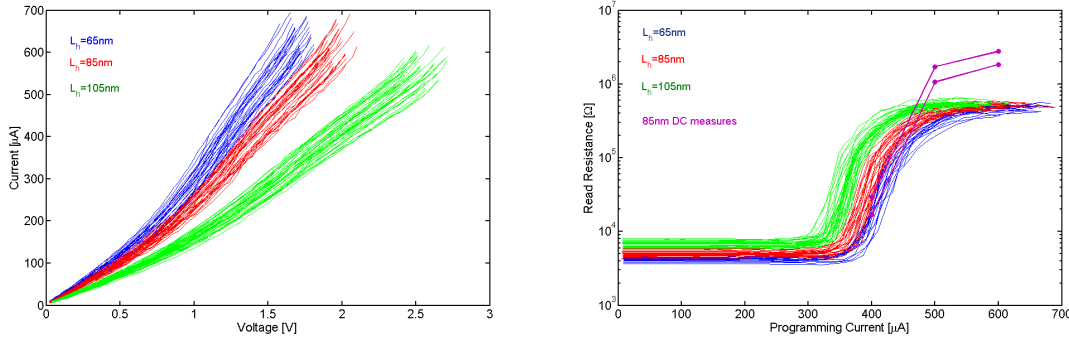


Figura 2.8: Caratteristiche IV ed RI dello split L_h . Sono mostrate le curve di circa 50 celle per ogni lunghezza di heater. Tra le curve RI della cella standard (85nm) compaiono le misure di resistenza di sottosoglia effettuate in DC per evitare il problema della saturazione della resistenza dovuta all'apparato di misura.

immediatamente che:

$$\frac{dR_{SET}}{dL_h} = \rho_h^{lf} \cdot \frac{1}{A_h} \quad (2.19)$$

$$\frac{dR_{ON}}{dL_h} = \rho_h^{hf} \cdot \frac{1}{A_h} \quad (2.20)$$

Quindi, nota la pendenza della curva R vs. L_h e la geometria (che si ipotizza essere quella nominale) è possibile ricavare la resistività dell'heater ad alte e basse temperature. Celle di questo tipo sono disponibili per il nodo 90nm. Per queste celle vengono riportate in figura 2.8 le caratteristiche IV e RI. Dalle caratteristiche IV si estrapolano poi le informazioni necessarie per l'estrazione di ρ_h . Le lunghezze di heater disponibili sono, nominalmente, 65nm, 85nm (standard) e 105nm. Si nota che il punto a 105nm non è allineato con gli altri due: infatti il modello prevede che essi abbiano un andamento lineare. Per spiegare questa differenza sono state formulate alcune ipotesi:

- Sono presenti forti non-linearità non previste dal modello analitico. Queste sono certamente presenti: infatti nel modello non si tiene conto del fatto che all'aumentare di L_h (e quindi della resistenza dell'heater) si ha anche una maggiore resistenza termica, e le temperature più elevate raggiunte all'interno del dispositivo possono causare una deviazione dalla linearità. Tuttavia esse dovrebbero avere l'effetto di ridurre

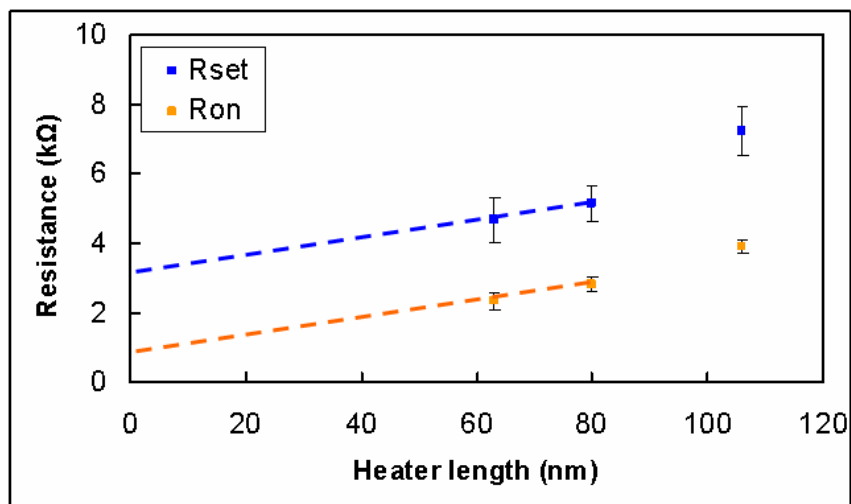


Figura 2.9: Valori di R_{SET} e R_{ON} al variare della lunghezza dell'heater (L_h). Ogni punto rappresenta la media delle misure su circa 50 celle. La pendenza delle curve viene utilizzata per estrapolare la resistività dell'heater, mentre l'intercetta della curva R_{ON} fornisce una stima delle resistenza interfacciale GST/heater per alte tensioni applicate.

la resistività dell'heater, mentre in questi dati si osserva un aumento di resistenza, pertanto questa ipotesi è da scartare.

- La lunghezza di heater è diversa da quella nominale. Se essa fosse pari a 115nm invece che a 105nm il modello sarebbe corretto. Tuttavia questa variazione (di circa 10%) è piuttosto elevata, e inoltre le simulazioni che verranno descritte nelle sezioni successive sono in grado di riprodurre correttamente altre caratteristiche della cella, in particolare la corrente di melting, utilizzando un valore di L_h pari a quello nominale. È quindi improbabile che la discrepanza sia dovuta a una così rilevante variazione di lunghezza rispetto al valore nominale.
- È presente un'interfaccia elettrica aggiuntiva sul contatto di bottom. Questo potrebbe verificarsi per via di problemi tecnologici che insorgono quando si fabbrica un dispositivo di dimensioni verticali elevate. Questa resistenza aggiuntiva fornirebbe un contributo diretto a R_{ON} e R_{SET} , mentre non contribuirebbe al riscaldamento del dispositivo in quanto si trova molto vicino al plug di tungsteno, che costituisce il bound-

	dR/dL_h [Ω/nm]	ρ_h [$\text{m}\Omega\text{cm}$]
SET (low-field)	26.6	2.12
ON (high-field)	29.47	2.36
R_{itf} (R_{ON} intercept)	230 Ω	

Tabella 2.2: Misure sperimentali relative allo split L_h .

ary termico inferiore. La caratteristica RI ne rimarrebbe dunque non influenzata, per cui quest'ultima ipotesi è quella che appare più probabile.

La tabella 2.2 mostra i valori sperimentali di dR/dL_h e i valori di ρ_h da essi ricavati. Si può immediatamente osservare che i valori di resistività ad alti e bassi campi sono molto simili tra loro. A causa della dispersione statistica osservata sulla resistenza di SET e di ON risulta difficile stabilire l'andamento qualitativo della resistività del riscaldatore rispetto alla temperatura. Questo esperimento ha comunque portato utili informazioni sul valore di ρ_h , che verrà utilizzato come punto di riferimento per gli esperimenti e le simulazioni successive. Un'altra utile informazione ricavabile dalla curva $R_{\text{ON}}-L_h$ è il valore della resistenza elettrica di interfaccia GST/heater. Infatti si ha che

$$R_{\text{ON}}(0) = R_{\text{itf}}^{\text{hf}} \quad (2.21)$$

Il valore risultate dall'esperimento è anch'esso riportato nella tabella 2.2, e dimostra l'esistenza di questa resistenza interfacciale. Anche la sua dipendenza dalla temperatura non è intuibile da queste misure, in quanto l'intercetta del grafico $R_{\text{SET}}-L_h$ contiene anche un termine riguardante la resistenza del calcogenuro, che è tuttora incognita. Per indagare meglio la resistenza dell'heater e la sua dipendenza dalla temperatura sono stati raccolti dei dati relativi a celle PCM prive del GST.

Celle no-GST

Per il nodo 45nm sono disponibili delle celle analitiche denominate *no-GST*, in cui la cella è priva del materiale calcogenuro. Questi dispositivi non hanno nessuna delle proprietà delle memorie PCM, e costituiscono essenzialmente un resistore costituito dall'heater. Celle di questo tipo sono molto utili per indagare le caratteristiche del riscaldatore senza

	R [Ω]	ρ_h [m Ω cm]
SET (low-field)	4262	2.47
ON (high-field)	2968	2.14

Tabella 2.3: Valori di resistenza caratteristica del dispositivo no-GST estrapolati dalle caratteristiche IV.

che il GST influenzi le misure ottenute. Analizzandone le caratteristiche I-V si possono ricavare le resistenze di ON e SET, definite allo stesso modo delle celle standard. In questo caso è evidente la diminuzione di resistività dell'heater all'aumentare della temperatura. Come mostrato in precedenza descrivendo il modello analitico, si ipotizza che la mobilità dell'heater abbia una dipendenza dalla temperatura di tipo power-law, scrivibile come

$$\rho_h(T) = \rho_0 \left(\frac{T}{T_0} \right)^\alpha \quad (2.22)$$

di cui si sta cercando il parametro α . Esso viene ricavato eseguendo una serie di simulazioni a parametri variabili, nel tentativo di trovare un fit adeguato per le curve IV delle celle no-GST. Per queste simulazioni si è scelto un valore di ρ_0 pari a 2.45m Ω cm, in accordo con le resistenze misurate in questo ultimo esperimento e non molto discorde dai risultati dello split di L_h . La struttura simulata è priva di interfacce elettriche e termiche in quanto il GST è assente e vengono trascurate le resistenze aggiuntive tra l'heater e gli elettrodi di top e bottom. La figura 2.10 mostra i risultati di queste simulazioni al variare di α . Il valore ricavato è pari ad $\alpha=0.03$. Questo parametro verrà impiegato in tutte le successive simulazioni. Dalla figura 2.10 si osserva che il comportamento della cella no-GST è molto lineare, anche se attraversato da correnti elevate rispetto a quelle usate in una cella PCM 45nm. Gli effetti di temperatura sono infatti meno evidenti, in quanto in questo caso gli elettrodi termicamente conduttivi (equivalenti a delle condizioni a contorno di Dirichlet) sono molto vicini tra loro (50nm) per via dell'assenza dello strato di calcogenuro di 50nm, quindi la cella si scalda meno. La figura 2.11 mostra il profilo di temperatura simulato in una sezione verticale presa nel centro del dispositivo confrontato con quello di una cella completa, allo scopo di evidenziare l'entità della differenza di riscaldamento nei due casi. Il punto $x=0$ è posto all'interfaccia tra GST e heater, e nel caso della cella no-GST rap-

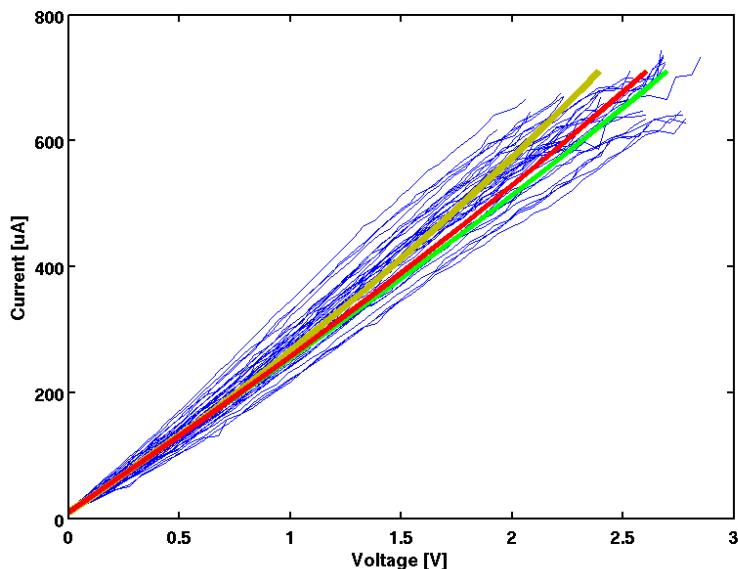


Figura 2.10: Caratteristiche IV di circa 50 celle no-GST (blu). Le simulazioni della struttura no-GST al variare del parametro α sono sovrapposte alle curve. Sono rappresentate le caratteristiche caratterizzate da $\alpha=0.03$ (verde), 0.05 (rosso), 0.1 (giallo).

presenta il contatto tra l'heater e il contatto metallico superiore. La forma parabolica del profilo di temperatura è attesa dalla soluzione dell'equazione del calore in presenza di un solo materiale con generazione di calore spazialmente uniforme.

2.2.3 Proprietà elettriche e termiche del GST

La calibrazione delle proprietà del calcogenuro è essenziale per la corretta simulazione del dispositivo, tuttavia durante lo svolgimento di questo lavoro sono sorti alcuni problemi a riguardo. Per prima cosa in ogni dato sperimentale tra quelli raccolti la variazione statistica delle proprietà influenzate dal GST è decisamente più elevata rispetto alle altre. Questo è dovuto sia a variazioni nel corso del processo di fabbricazione (deposizione del GST), sia alle diverse tecniche di *seasoning* (stagionatura) utilizzate per l'attivazione del calcogenuro. Infatti, diverse combinazioni di temperatura e tempo di bake hanno effetti diversi sullo stato del materiale attivo durante il funzionamento. La dispersione statistica

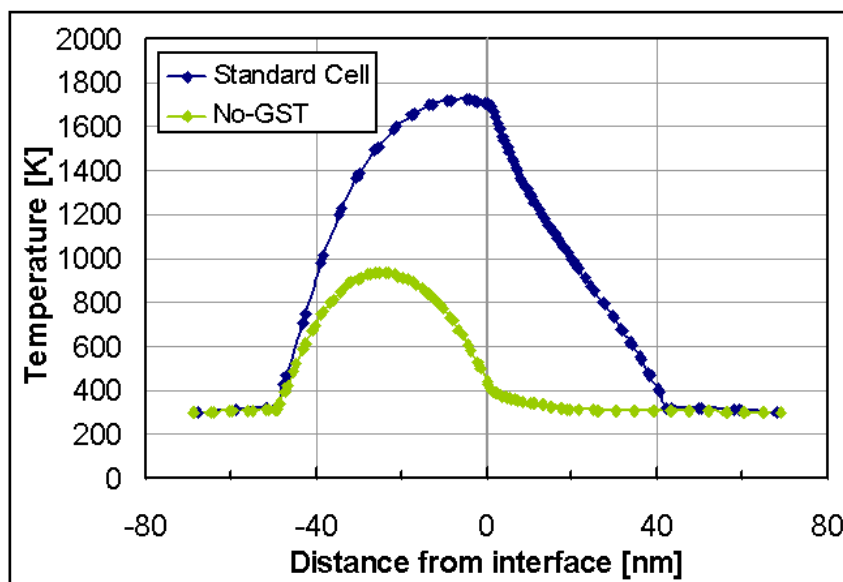


Figura 2.11: Profilo di temperatura simulato della cella standard e no-GST lungo l'asse verticale del dispositivo.

dei parametri del GST sarà affrontata in dettaglio nei capitoli successivi.

Un altro ostacolo alla corretta calibrazione del dispositivo è dato dal contributo che il GST fornisce a quasi ogni parametro caratteristico della cella, in particolare la resistenza di SET e la corrente di melting, rendendo complicato separare gli effetti dovuti al materiale attivo da quelli degli altri materiali presenti. Inoltre, per via della particolare geometria della cella Wall (uno strato largo di GST depositato sopra un heater molto stretto) e degli effetti termici, le relazioni che governano il trasporto di corrente e di calore nel calcogenuro sono per lo più non lineari. Perciò anche in questo caso si procederà alla calibrazione analizzando i dati sperimentali mediante il modello analitico, per poi raffinare la scelta effettuata mediante simulazioni, confrontando il tutto con valori reperiti in letteratura al fine di verificarne la validità.

2.2.4 Resistività del GST

Come illustrato in precedenza, la struttura a bande del GST è caratterizzata da un livello di trappole in prossimità della banda di valenza (a una distanza E_{traps} da quest'ultima)

che rappresenta le vacanze strutturali tipiche della forma cristallina del calcogenuro. Trattandosi di un centro di cattura per gli elettroni questo livello di trappola ha l'effetto di posizionare il livello di Fermi tra E_{traps} e la banda di valenza, conferendo al materiale una conduzione di tipo p. La resistività dipende dunque sia dalla posizione del livello trappole, che determina in ultima istanza la densità di portatori liberi, sia dalla mobilità di questi ultimi. Il valore di E_{traps} determina direttamente anche l'energia di attivazione E_a del GST, parametro importante per valutare correttamente la dipendenza dalla temperatura del calcogenuro.

Per valutare l'energia di attivazione del GST sono state raccolte delle misure in temperatura su celle analitiche del nodo 45nm. Con questo tipo di esperimento si impongono alcuni livelli di temperatura al dispositivo e si misura la resistenza di bassi campi (SET) al variare di T. I risultati vengono riportati sul cosiddetto *Arrhenius plot*, un grafico che ha in ascissa l'inverso dell'energia ($1/kT$) e in ordinata il logaritmo della resistenza misurata. Infatti, se la dipendenza della temperatura del dispositivo è di tipo esponenziale si osserva subito che

$$R = \rho_0 e^{\frac{E_a}{kT}} \quad (2.23)$$

$$\ln(R) = K + \frac{E_a}{kT} \quad (2.24)$$

da cui si deduce che la pendenza del grafico di Arrhenius rappresenta proprio l'energia di attivazione. La figura 2.12 a sinistra mostra un esempio di Arrhenius plot per una singola cella, mentre a destra viene mostrato l'andamento dell'energia di attivazione della cella all'aumentare della resistenza programmata (ovvero per impulsi di RESET più ampi). Il valore di E_a nella fase di SET (la resistenza più bassa) risulta essere tra i 14 e i 18 meV. Va osservato che questa è l'energia di attivazione della resistenza equivalente relativa all'intera cella Wall, e comprende i contributi dell'heater e dell'interfaccia heater/GST. Il coefficiente della power law del riscaldatore è già stato indipendentemente calibrato, mentre per il momento si ipotizza che la resistenza interfacciale sia un termine costante in temperatura, per semplificare la calibrazione iniziale, che verrà raffinata in seguito. Sono state lanciate alcune simulazioni per stabilire una relazione tra l'energia di attivazione e la posizione del livello delle trappole. Come in ogni semiconduttore descritto da una struttura a bande, E_a è pari alla distanza tra la banda di valenza e il livello di Fermi. Tuttavia, poiché

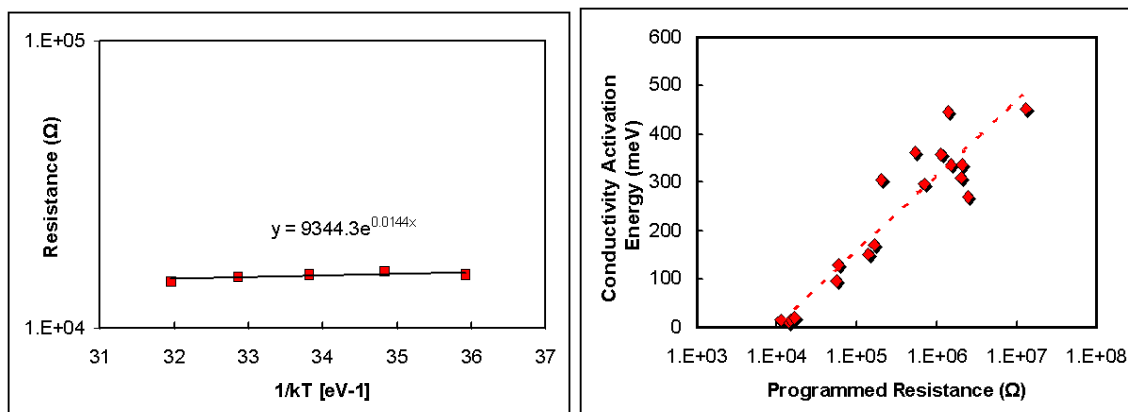


Figura 2.12: Sinistra: Arrhenius plot della resistenza di SET di una singola cella 45nm. La pendenza della curva rappresenta l'energia di attivazione E_a del dispositivo. Destra: energia di attivazione della cella al variare della resistenza programmata (ovvero per impulsi di programmazione di ampiezza crescente).

quest'ultimo si colloca in prossimità del bordo della banda, non è possibile utilizzare l'approssimazione di Boltzmann nel calcolo della densità di portatori, e bisogna ricorrere alla statistica di Fermi vera e propria. Questa complica notevolmente i calcoli analitici, per cui si ricorre alle simulazioni numerica per stabilire con precisione la posizione del livello di Fermi. La figura 2.13 riporta i risultati delle simulazioni. Un valore di E_{traps} pari a 50meV dà come risultato un'energia di attivazione dell'intera cella Wall di 18.4meV, corrispondente a una E_a del solo GST di 40meV, ed è stato pertanto selezionato come valore da utilizzare nelle simulazioni successive.

Celle con altezza di GST variabile

Come per l'heater, anche per valutare la resistività del GST si può ricorrere a celle con altezza di GST L_c variabile. Tuttavia in questo caso il calcolo è più complesso, in quanto, a causa della geometria del dispositivo, la relazione tra resistenza e lunghezza non è lineare ma logaritmica. Anche in questo caso però si può procedere con una stima che verrà

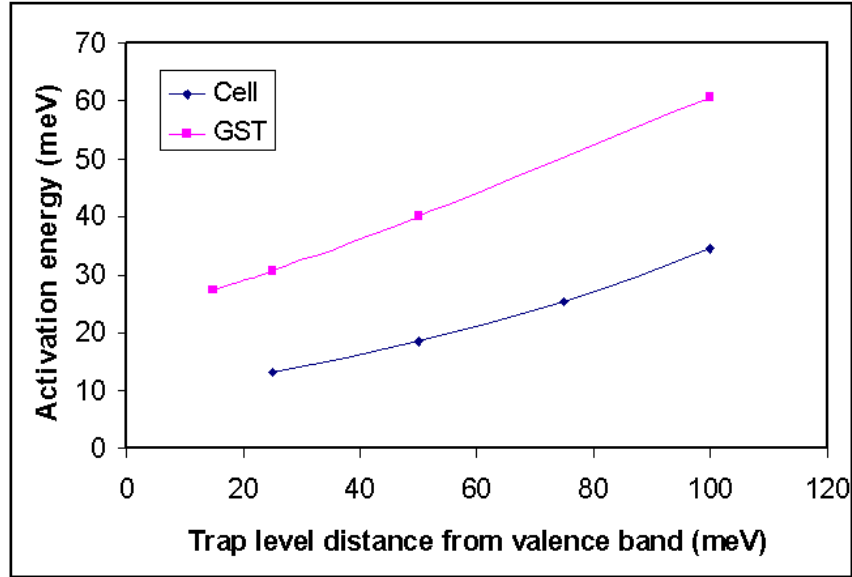


Figura 2.13: Energia di attivazione della cella 45nm e del GST ricavate dalle simulazioni numeriche con una tensione applicata di 0.4 V.

successivamente migliorata. Stando al modello analitico, si ha che

$$R_{SET} = R_c + R_h + R_{itf} \quad (2.25)$$

$$R_c = \rho_c \frac{1}{2nW_c} \left(1 + \frac{2nL_c}{t} \right) \quad (2.26)$$

$$\frac{dR_{SET}}{dL_c} = \rho_c \frac{1}{W_c(t + 2nL_c)} \quad (2.27)$$

Per cui anche in questo caso è possibile, nota la geometria della cella, risalire alla resistività, a patto di effettuare un paio di ulteriori approssimazioni: innanzitutto si linearizza il logaritmo nell'intorno dei valori di L_c disponibili nelle misure, poi si ipotizza che il fattore n sia pari a 1. Quest'ultimo tiene conto dell'angolo che il flusso di corrente forma con la retta perpendicolare all'interfaccia, e nel caso di $n=1$ si ha uno spread della corrente di 45 gradi.

Per questo esperimento sono disponibili vari lotti di dispositivi PCM, sia 90nm che 45nm. Nella figura 2.14 vengono mostrati i valori di R_{SET} al variare di L_c . Oltre alle difficoltà

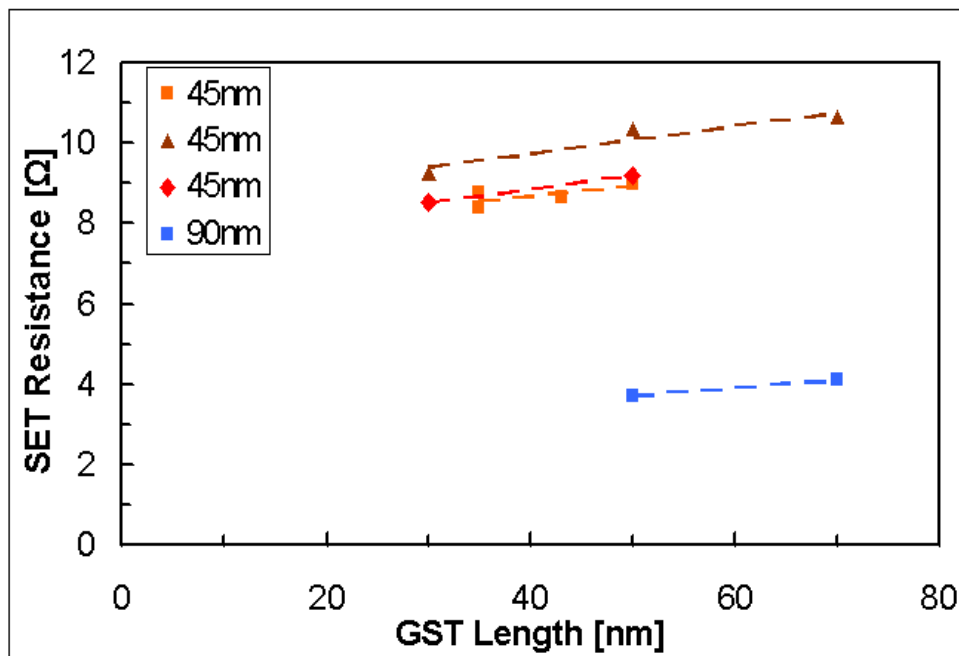


Figura 2.14: Resistenza di SET in funzione dell'altezza dello strato di GST L_c per celle dei nodi 90nm e 45nm. Ogni punto rappresenta la media di circa 50 celle. Questi dati sono usati per stimare la resistività del GST.

riportare in precedenza dovute alle non-linearità intrinsecamente presenti nel sistema, è noto da letteratura che la resistività del GST è caratterizzata da un forte spread statistico, e il suo valore è stato riportato come compreso tra i $15\text{m}\Omega\text{cm}$ e i $30\text{m}\Omega\text{cm}$. Anche le misure raccolte in questo lavoro prevedono risultati comparabili, come riportato nella tabella 2.4. Per le simulazioni successive viene adottato il valore di $27\text{m}\Omega\text{cm}$, che consente di ottenere un fit corretto per il maggior numero di dati sperimentali possibile. Per ottenere questo valore, considerando la calibrazione dell'energia di attivazione già effettuata, è necessario porre la mobilità delle lacune pari a $\mu_c = 7$. Nel corso di questo lavoro si trascura la dipendenza dalla temperatura della mobilità delle lacune nel calcogenuro.

n.	Nodo [nm]	ρ_c [m Ω cm]
1	90	30.12
2	45	21.43
3	45	12.23
4	45	15.78

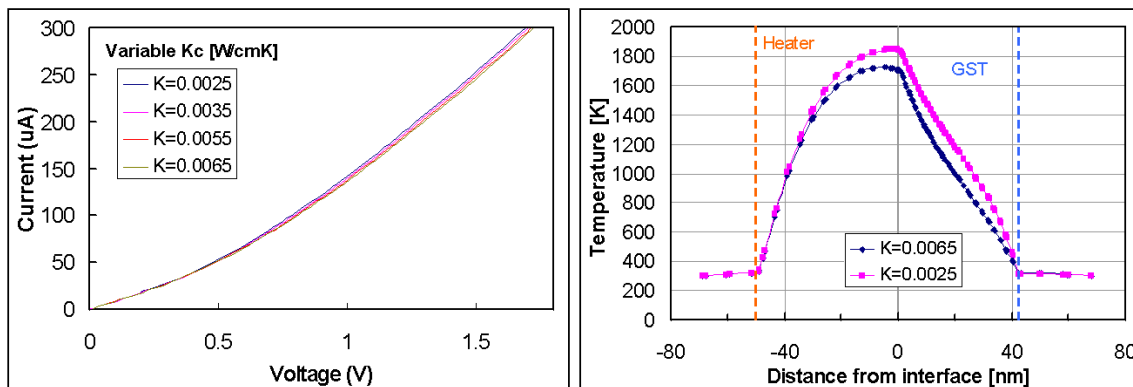
Tabella 2.4: Resistività del GST ricavata dalle misure su celle con L_c variabile.

Figura 2.15: Sinistra: caratteristiche IV simulate al variare di K_c . Non si notano sostanziali differenze. Destra: profilo di temperatura lungo la direzione verticale al centro del dispositivo. La temperatura nel bulk dell'heater non viene molto influenzata, mentre si ha un effetto non trascurabile sulla temperatura all'interfaccia GST/heater e nel bulk del calcogenuro.

2.2.5 Conducibilità termica

La conducibilità termica del GST influisce sulla generazione e sul trasporto del calore all'interno del materiale attivo. Ne vengono quindi influenzate tutte le caratteristiche dovute a fenomeni dipendenti dalla temperatura. Nel modello analitico semplificato questo contributo viene descritto dalla dipendenza della resistenza termica del GST dalla conducibilità, tale per cui

$$R_c^{\text{th}} = \frac{1}{K_c} \cdot \frac{1}{2nW_c} \left(1 + \frac{2nL_c}{t} \right) \quad (2.28)$$

La figura 2.15 mostra le caratteristiche IV simulate al variare di K_c . La differenza tra le varie curve è molto ridotta. Infatti, per basse tensioni applicate la temperatura non ha ancora raggiunto livelli tali da fare apprezzare la variabilità indotta da K_c , mentre per alti

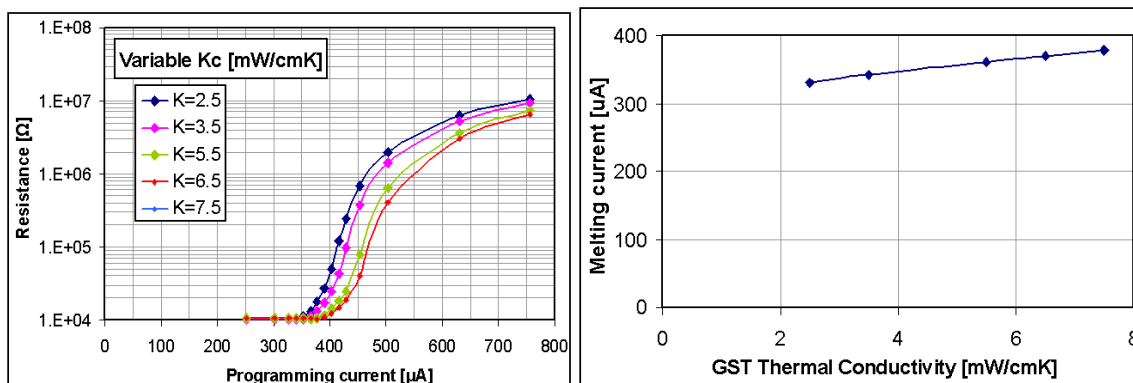


Figura 2.16: Sinistra: curve di programmazione RI in funzione di K_c . Le curve subiscono una traslazione al variare della conducibilità. Destra: Valore della corrente di melting I_m in funzione della conducibilità termica di bulk del GST.

campi elettrici la resistenza serie del dispositivo è dominata dall'heater, le cui proprietà sono solo marginalmente influenzate dalla temperatura. A dimostrazione di quest'ultima affermazione, nella figura 2.15 a sinistra viene riportato il profilo di temperatura lungo una sezione verticale del dispositivo, collocata nel suo centro. Al variare di K_c cambia la forma del profilo all'interno del materiale attivo e nei pressi dell'interfaccia, mentre nella zona di bulk del riscaldatore non si notano differenze sostanziali. Questo permette di concludere che il valore della conducibilità termica ha poco effetto sulle prestazioni di readout del dispositivo.

La corrente di fusione I_m ne è invece influenzata. Il modello analitico prevede che

$$I_m = \sqrt{\frac{T_m - T_0}{\alpha R_{ON} R_{tot}^{th}}} \quad (2.29)$$

in cui R_{tot}^{th} rappresenta il parallelo delle resistenze termiche di GST e heater. La corrente di melting dipende quindi dalla radice di K_c . Questo rende difficile l'estrazione diretta del parametro come nel caso della resistività dell'heater. Tuttavia, sapendo che le curve IV non ne vengono influenzate, è possibile eseguire una calibrazione cercando un fitting per le caratteristiche RI delle celle. Si è deciso di procedere simulando prima le celle del nodo 90nm, e verificare i valori così ottenuti sia simulando anche le celle a 45nm, sia

confrontandoli con quelli reperiti in letteratura. La figura 2.16 mostra le simulazioni sulle celle a 90nm, in cui si vede che al variare di K_c si ha una variazione della corrente di melting, che coincide con una traslazione della curva RI. Con questo metodo si giunge a un valore di K_c di 5.7mW/cmK. Questo è in accordo con quelli trovati in letteratura per la fase cubica a facce centrate: 5.7mW/cmK in [10], tra 5 e 5.3mW/cmK in [12] e [13], mentre sono stati anche riportati valori più bassi, come 2.8mW/cmK in [11]. La dipendenza dalla temperatura è data dalla legge di Wiedemann-Franz (equazione 2.8), che tiene in considerazione il contributo al trasporto del calore dei portatori liberi generati termicamente nei pressi dell'interfaccia GST/heater. Nel corso dell'operazione del dispositivo quindi la conducibilità termica aumenterà, anche fino a un valore 15mW/cmK durante l'operazione di RESET, come ricavato dalle simulazioni.

2.2.6 Calibrazione delle interfacce termiche ed elettriche

Nel simulare correttamente il dispositivo PCM la presenza delle interfacce termiche ed elettriche è essenziale. Le interfacce termiche servono per descrivere le non regolarità nel trasporto di calore che si hanno quando due materiali diversi entrano in contatto. In un modello analitico compatto questi effetti sono sintetizzabili con una resistenza termica serie equivalente, e nella struttura che viene simulata in questo lavoro sono realizzate aggiungendo dei sottili strati (1 nm) di materiale elettricamente isolante caratterizzato da una conducibilità termica diversa (generalmente inferiore) da quella degli elementi circostanti. In questo ambito è soprattutto interessante la conducibilità superficiale, espressa in $\frac{W}{cm^2K}$. Il simulatore SDevice non è in grado di simulare questo tipo di interfacce, e da qui deriva la necessità di inserire uno piccolo strato di dimensione finita. Nel corso della calibrazione saranno dunque trovati dei valori di conducibilità superficiale che verranno poi inseriti nel programma moltiplicandoli per lo spessore di 1 nm, ottenendo così la conducibilità di bulk richiesta dal simulatore.

Le interfacce hanno un effetto sostanziale sul profilo di temperatura all'interno della zona attiva. Infatti esse tendono a confinare il calore generato nel GST al suo interno. Se non ci fossero zone interfacciali, il profilo termico sarebbe determinato solamente dalle

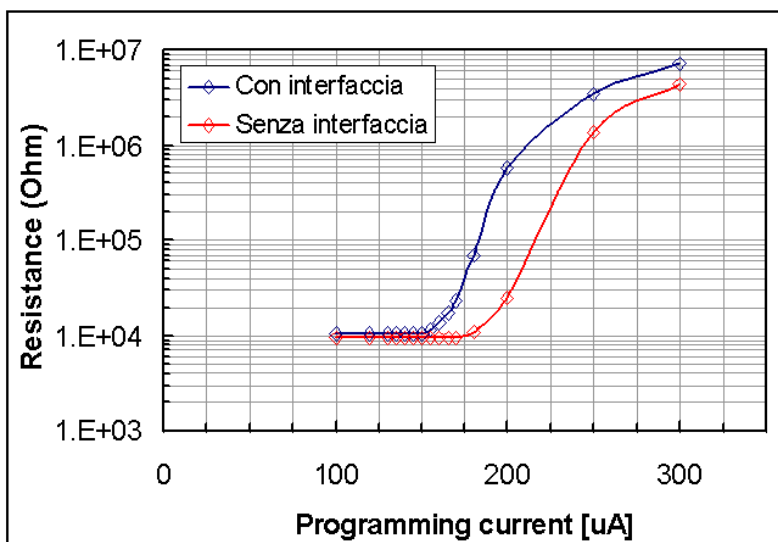


Figura 2.17: Curve di programmazione RI simulate in presenza e in assenza delle resistenze termiche alle interfacce GST/ossido. Si nota che le curve raggiungono circa lo stesso valore per alte correnti di programmazione, ma hanno pendenza diversa. Questa è determinata dall'entità del confinamento termico, a sua volta dipendente dalla presenza e dal valore delle resistenze termiche interfacciali.

proprietà di bulk del calcogenuro e dei materiali isolanti che lo circondano. La figura 2.17 riporta la curva di programmazione simulata di una cella 45nm nel caso appena descritto, confrontata con la stessa struttura provvista di interfacce. Il calore generato può lasciare il GST tramite i contatti di top e bottom e tramite le perdite dei materiali isolanti. Maggiore è il calore perso, più alta sarà la corrente necessaria a portare a fusione il materiale: le interfacce hanno quindi effetto sulla I_m . Dalla figura 2.17 si osserva anche una differenza nella pendenza della curva R-I nei due casi. Questa pendenza fornisce un'utile indicazione sull'entità del confinamento termico. La transizione non ideale dallo stato basso resistivo a quello alto resistivo è dovuta al fatto che alla corrente di melting non si viene a creare uno strato uniforme di GST amorfo che ricopra interamente l'area dell'interfaccia GST/heater. In questo caso infatti si avrebbe un passaggio netto della resistenza dal valore di SET a quello di RESET, superiore di un paio di ordini di grandezza. In realtà, anche se la corrente di programmazione consente di creare una zona amorfa nei pressi del centro del riscaldatore, le zone laterali sono caratterizzate da un profilo di temperatura esponen-

ziale decrescente, come dettato dall'equazione di trasporto del calore. Pertanto, esiste una finestra di ampiezza di correnti di programmazione entro la quale si ha una non completa amorfizzazione della zona attiva a ridosso del contatto GST/heater: questo fenomeno può essere schematizzato inserendo una resistenza in parallelo a quella del calcogenuro amorfo, che rappresenta la parte di materiale cristallino rimasta dopo la programmazione. All'aumentare delle correnti questa zona laterale cristallina si riduce sempre di più (aumentando la resistenza parallelo), fino ad avere un'area tendente a zero. In quest'ultima condizione l'area attiva è interamente sovrastata dall'amorfo, e l'aumento di resistenza non è più esponenziale, ma è dettato dall'incremento verticale della dimensione della calotta. Se si indica con x_m la dimensione laterale della zona amorfa ($x_m < W_c$), si ottiene:

$$R_a = \rho_a \frac{1}{2nx_m} \ln \left(1 + \frac{2nL_a}{t} \right) \quad (2.30)$$

$$R_x = \rho_x \frac{L_c}{t(W_c - x_m)} \quad (2.31)$$

in cui i pedici a e x indicano, rispettivamente, il GST amorfo e cristallino. La resistenza totale è data dal parallelo di R_a e R_x . Il calcolo di x_m richiede la soluzione dell'equazione del calore in tre dimensioni, particolarmente onerosa se ottenuta per via analitica. È però possibile, mediante simulazioni a elementi finiti, avere un'idea qualitativa del fenomeno. La figura 2.19 mostra quanto appena detto. Nel riquadro superiore viene mostrato il profilo laterale di temperatura, in cui sono colorate solamente le temperature superiori a T_m , corrispondenti a tre ampiezze crescenti dell'impulso di programmazione, il cui valore è indicato nella curva RI del riquadro inferiore. Nel punto 1, al di sotto di I_m , il materiale amorfizzato ha un volume molto basso, infatti la resistenza totale della cella si discosta di poco dal valore di SET. Nel punto 2, corrispondente circa a I_m , la quantità di GST amorfo è tale da provocare un innalzamento della resistenza letta, tuttavia è ancora presente un cammino a bassa resistività in parallelo che ne limita l'incremento. Nel punto 3, ben oltre I_m , l'amorfo ricopre interamente l'area dell'interfaccia GST/heater, per cui il cammino di conduzione incontra solamente la resistività caratteristica del GST vetroso. Da questo punto in avanti l'aumento di resistenza con l'ampiezza dell'impulso di programmazione è provocato solo dalla maggiore dimensione verticale della calotta amorfa, per cui ha ter-

	TBC [W/cm ² K]
GST/SiO ₂	4·10 ³ [12]
	11-20·10 ³ [13]
GST/TiN	3.85·10 ³ [15]

Tabella 2.5: Misure del valore della conducibilità termica interfacciale reperite in letteratura.

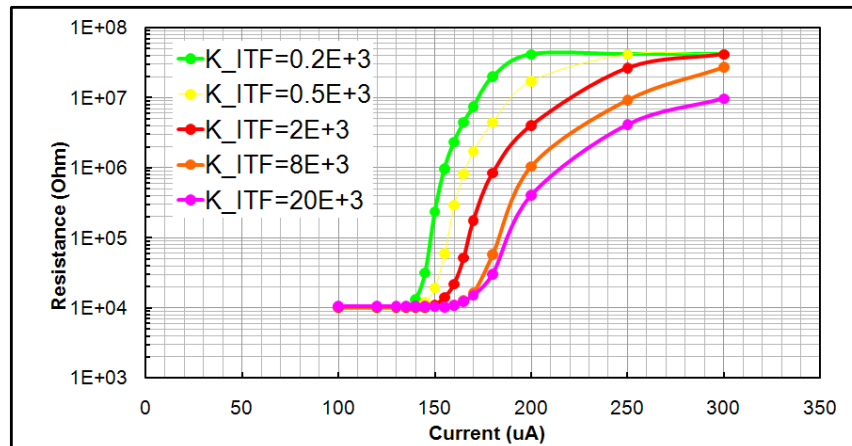
mine il tratto esponenziale della curva RI.

Anche senza essere approdati a una forma analitica chiusa che descriva accuratamente il processo di confinamento del calore, si può comunque concludere che esso ha un effetto determinante sulla pendenza (slope) della curva di programmazione RI. La calibrazione delle interfacce prenderà quindi il via raccogliendo dati dalla letteratura e da misure effettuate presso i laboratori Numonyx riguardanti la conducibilità delle interfacce e proseguirà con il fitting della pendenza della curva di programmazione.

La tabella 2.5 mostra le misure raccolte dalla letteratura. Esse riguardano le interfacce GST/SiO₂ e GST/TiN. Le misure effettuate localmente per l'interfaccia GST/SiO₂ tramite le tecniche 3 ω e PTR (*Photo Thermal Radiometry*) danno un valore di 1.43·10³ W/cm²K, mentre quelle relative alle interfacce con i contatti di top e bottom sono mostrate in figura 2.20. La figura 2.18 mostra delle curve RI simulate in cui viene fatto variare il valore della resistenza d'interfaccia tra GST e cap metallico superiore. Per calibrare correttamente le interfacce occorre dare una definizione più precisa della slope della curva di programmazione. Detto R_1 il valore di resistenza letta in corrispondenza della corrente di fusione I_m e definito $R_2 = 20 \cdot R_1$, si definisce la slope

$$S = \frac{\log_{10}(I_2) - \log_{10}(I_m)}{\log_{10}(R_2) - \log_{10}(R_1)} \quad (2.32)$$

Si ha che la slope per le celle del nodo 45nm vale 25.48 μ A/dec. In base alle simulazioni appena mostrate e ai dati sperimentali si è deciso di adottare una resistenza interfacciale di 2·10³ W/cm²K per quel che riguarda l'interfaccia GST/Ossido e GST/Nitrato (valore preso direttamente dai dati sperimentali) e 8·10³ W/cm²K per l'interfaccia GST/cap superiore. L'interfaccia elettrica (EIR) tra GST e heater viene descritta in un modo analogo: un



K_{itf} [$W/cm^2K \cdot 10^3$]	Slope [$\mu A/dec$]
0.5	8.43
1	10.58
2	14.48
3.5	15.91
8	22.22

Figura 2.18: Curve di programmazione RI simulate per diversi valori di conducibilità termica interfacciale. Si nota come l'interfaccia influisca sia sul valore di I_m , sia sulla pendenza della curva.

sottile strato di materiale (1 nm) identico a quello del riscaldatore eccetto per un diverso valore di resistenza elettrica. L'esistenza di questa interfaccia, oltre a essere documentata in letteratura, è dimostrata anche dalle misure di resistenza della cella Wall in tecnologia 90nm al variare di L_h , come mostrato nelle sezioni precedenti. Il suo valore ad alti campi è già stato stimato in quell'occasione a circa 300Ω . Per bassi campi elettrici la sua stima è più ardua in quanto in quella zona operativa la resistenza totale della cella è influenzata anche dal GST oltre che dall'heater e dalle resistenze interfacciali. Si riesce comunque ad avere una stima qualitativa del suo valore osservando l'intercetta del grafico $R_{SET}-L_h$ in figura 2.9 e sottraendo la resistenza del GST R_c ricavata utilizzando il valore di resistività trovato in precedenza. Si ottiene un valore di R_{itf} per le celle da 90nm a bassi campi pari a 436Ω , che corrisponde a una resistività (su uno strato di 1 nm) di $30.52m\Omega cm$. La sua dipendenza dalla temperatura (anch'essa di tipo power-law) viene calibrata tenendo

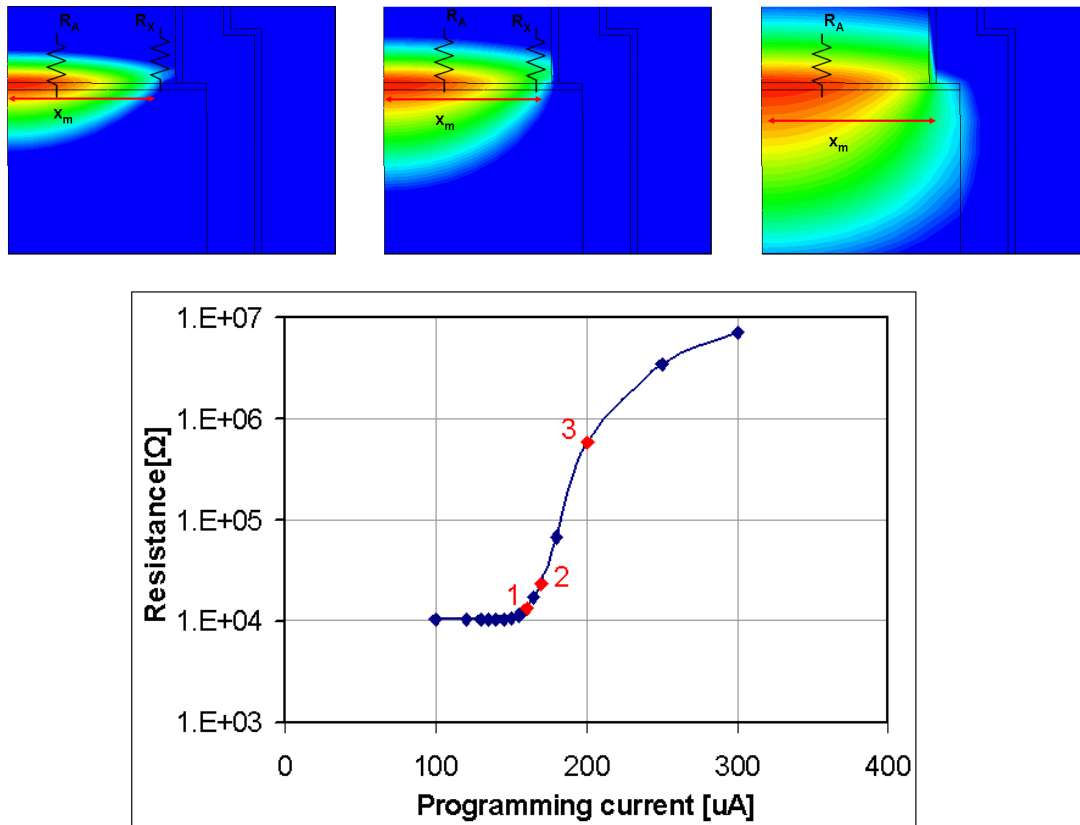


Figura 2.19: Illustrazione dell'origine della pendenza della curva RI. Nel *punto 1* (sinistra) la corrente applicata è circa pari a I_m . Una parte della regione di GST al di sopra dell'interfaccia non ha raggiunto la temperatura di fusione, per cui esiste un percorso basso resistivo in parallelo alla calotta amorfa. Nel *punto 2* (al centro) la corrente è superiore a I_m . La quasi totalità della zona GST a ridosso dell'heater a raggiunto la temperatura T_m e si è amorfizzata: la resistenza totale aumenta esponenzialmente man mano che il percorso parallelo basso resistivo si restringe. Nel *punto 3* la corrente applicata è ancora più elevata. La resistenza della cella è dovuta interamente alla calotta amorfa, caratterizzata da una resistività molto elevata. Da qui in avanti la resistenza aumenta, non linearmente, per via dell'incremento dell'altezza della calotta.

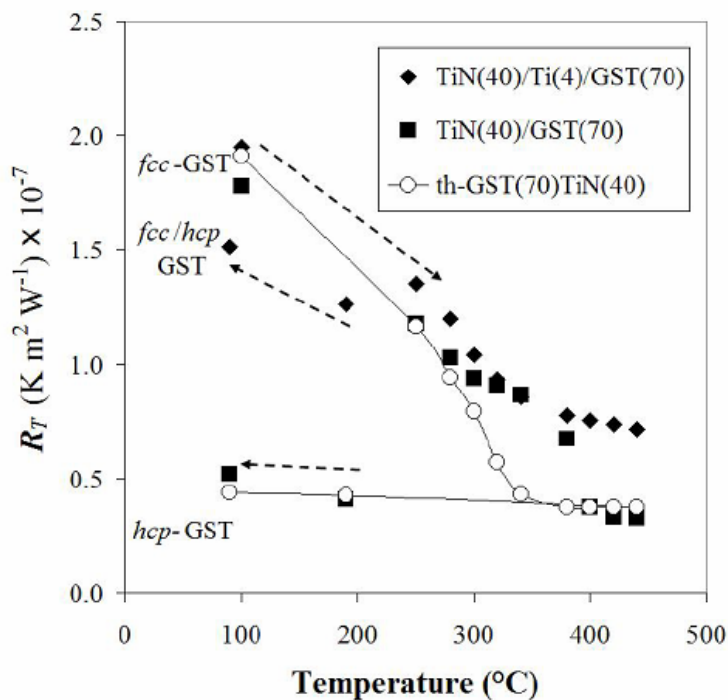


Figura 2.20: Misure di resistenza termica interfaciale prese presso Numonyx utilizzando la tecnica MPTR.

conto che la resistenza deve calare di circa un fattore 2 nel corso di un'escursione termica compresa tra i 1300 e 1500 °K, tale infatti è la temperatura raggiunta dall'interfaccia per correnti di programmazione molto elevate. In base a questa considerazione a un fitting delle caratteristiche I-V delle celle standard 90nm si ottiene un valore di α_{if} pari a 0.15.

2.3 Validazione del modello numerico

Nella sezione precedente si è calibrato il modello numerico sulla base dei dati sperimentali. Lo scopo di questo lavoro è quello di ottenere un set di parametri unico che sia in grado di riprodurre le caratteristiche della cella PCM a 90nm e 45nm, in modo tale da avere a disposizione uno strumento che consenta la simulazione di nodi tecnologici successivi e geometrie alternative con un ottimo grado di affidabilità. La tabella 2.6 riassume tutti i

Resistività dell'heater ai bassi campi (ρ_h)	2.45	m Ω cm
Coefficiente di dipendenza dalla temperatura dell'heater (α_h)	0.03	
Resistività superficiale interfaccia GST/heater (ρ_{itf}^{\square})	$3 \cdot 10^{-8}$	m Ω cm ²
Coefficiente di dipendenza dalla temperatura dell'interfaccia (α_{itf})	0.15	
Mobilità delle lacune nel GST (μ_c)	7	cm ² /Vs
Distanza del livello trappola dalla banda di valenza nel GST (E_{traps})	0.05	eV
Conducibilità termica superficiale all'interfaccia GST/ossido	$2 \cdot 10^3$	W/cm ² K
Conducibilità termica superficiale all'interfaccia GST/cap	$8 \cdot 10^3$	W/cm ² K

Tabella 2.6: Riassunto dei parametri calibrati.

parametri che sono stati calibrati. Per validare il modello si è lanciata la simulazione e si sono confrontati i risultati (in particolare le curve IV ed RI) con le caratteristiche di svariate celle, differenti da quelle usate per la calibrazione. Le immagini successive mostrano i risultati ottenuti. La figura 2.21 mostra il modello applicato alle celle standard del nodo 45nm; le caratteristiche I-V ed R-I vengono entrambe ben riprodotte dalla simulazione a elementi finiti. Questo consentirà di effettuare altre analisi sulle celle standard di entrambi i nodi tecnologici, mostrate nel capitolo successivo. Le figure 2.22 e 2.23 mostrano tutte le curve delle celle 90nm con diversi valori di L_h . Eccetto la cella con altezza di heater pari a 105nm, della cui discrepanza si è parlato nel capitolo relativo, tutte le altre curve vengono simulate correttamente, segno che la dipendenza da L_h viene descritta bene dal modello numerico. La figura 2.24 mostra i parametri caratteristici della cella simulata sovrapposti ai dati sperimentali per celle con altezza di GST variabile. Queste celle sono differenti da quelle utilizzate nel corso della calibrazione della resistività del GST, e l'andamento viene ben riprodotto. La figura 2.25 mostra i medesimi parametri relativi a cella con diversa larghezza dell'area attiva W . Anche in questo caso la dipendenza dalla geometria è riprodotta piuttosto bene senza bisogno di ulteriori calibrazioni.

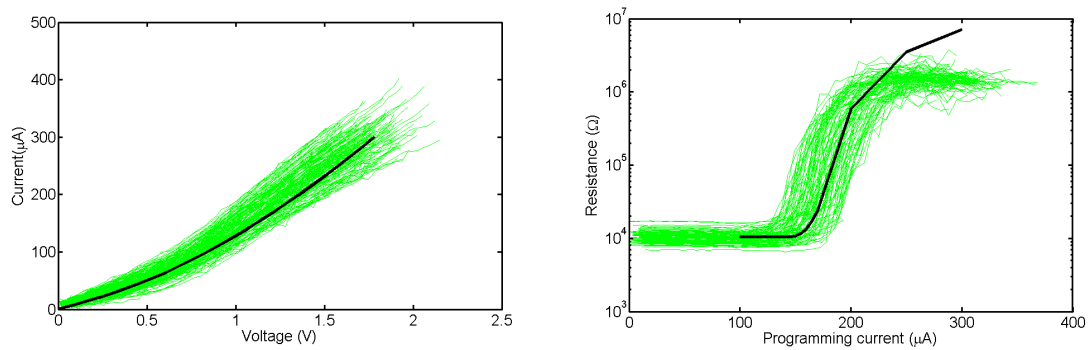


Figura 2.21: Validazione del modello numerico calibrato. In verde le curve IV e RI sperimentali delle celle standard 45nm e in nero le caratteristiche simulate. Ottimo fit dei dati con il set di parametri unico.

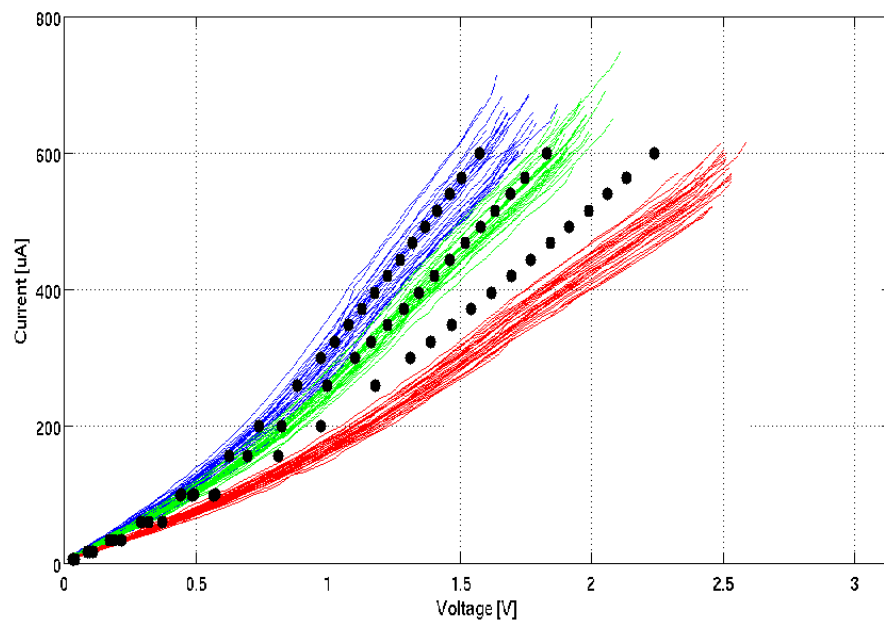


Figura 2.22: Validazione del modello numerico calibrato. Fitting delle caratteristiche IV per diversi valori di lunghezza di heater L_h : 65nm (blu), 85nm (verde), 105nm (rosso). Simboli: simulazioni. La curva 105nm non viene ben riprodotta per via di un'interfaccia aggiuntiva presente sul contatto di bottom, come illustrato nel paragrafo relativo.

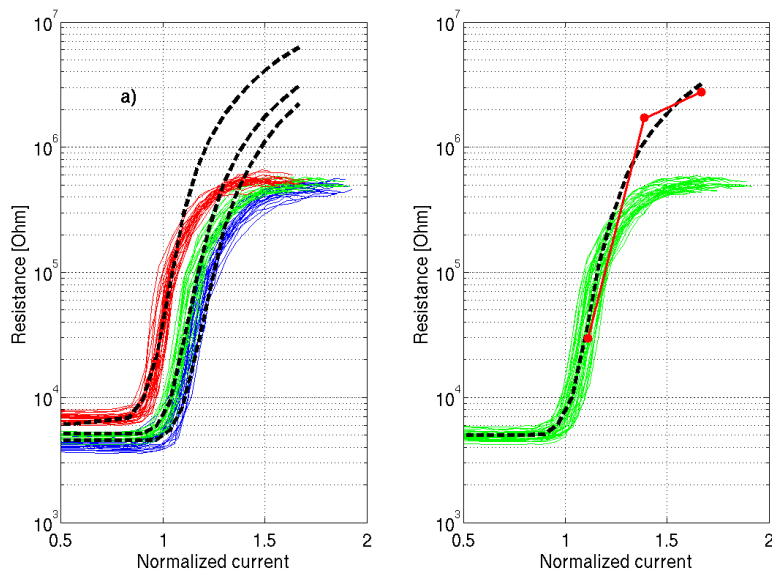


Figura 2.23: Validazione del modello numerico calibrato. Sinistra: fitting delle curve di programmazione per diversi valori di lunghezza di heater L_h : 65nm (blu), 85nm (verde), 105nm (rosso). Simboli: simulazioni. Destra: fitting della curva RI del nodo 85nm in cui sono mostrati anche i dati sperimentali della misura parametrica in DC, che rimuove il problema della saturazione della resistenza misurata. Anche in questo caso si ha un buon fitting.

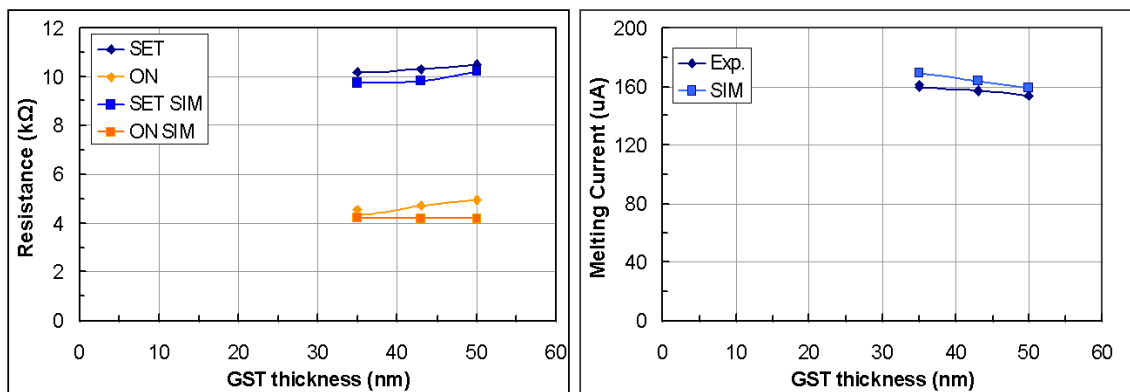


Figura 2.24: Validazione del modello numerico calibrato. Simulazione dei parametri R_{SET} , R_{ON} e I_m per delle celle con diversa altezza di GST L_c . Queste celle appartengono a un lotto differente da quello utilizzato in precedenza per la calibrazione del GST.

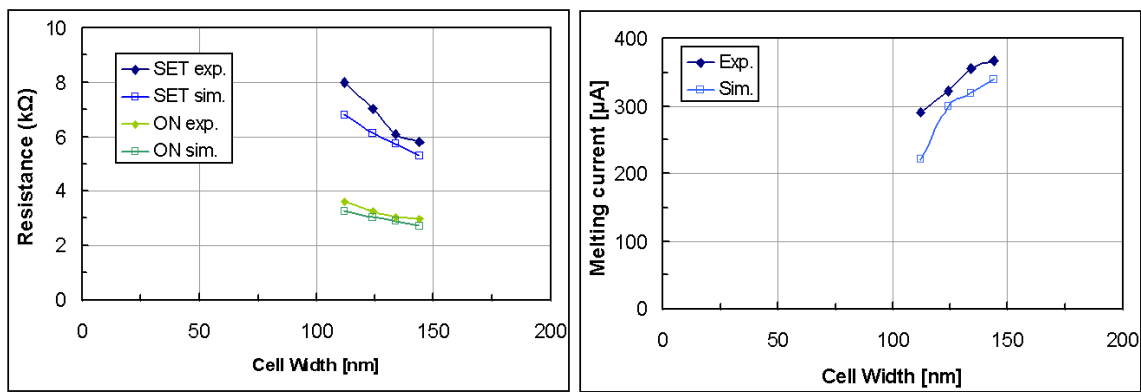


Figura 2.25: Validazione del modello numerico calibrato. Simulazione dei parametri R_{SET} , R_{ON} e I_m per celle con una diversa larghezza dell'area attiva (W).

Capitolo 3

Applicazioni del modello

3.1 Analisi di sensibilità

Come si è visto, le cifre di merito principali di una cella di memoria PCM (resistenza di lettura, corrente di fusione, ecc.) presentano tutte un certo numero di dipendenze dalle caratteristiche geometriche e fisiche della cella. Ognuna di queste è affetta da variabilità statistica: le dimensioni della cella potrebbero non essere quelle desiderate, il GST depositato può non presentare le sue proprietà nominali, ecc. È di grande importanza riuscire ad avere una stima dell'incidenza di ciascuna di queste variabili sul dispositivo, al fine di individuare le maggiori criticità. A questo scopo il modello numerico appena descritto può risultare uno strumento estremamente utile.

La figura 3.1 mostra lo scatter plot delle resistenze di ON e SET in funzione della corrente di melting per circa 100 celle standard a 90nm e 45nm. Tutti e tre i parametri sono affetti da dispersione statistica, la cui deviazione standard è indicata nella tabella 3.1. Si nota che R_{SET} risulta meglio correlata con I_m rispetto a R_{ON} . Questa relazione tra le grandezze non è tuttavia attesa dal modello analitico, il quale prevederebbe una maggiore correlazione tra R_{ON} e I_m . Questo perché per correnti di programmazione prossime al melting la resistenza del GST è già notevolmente calata, e l'heater domina la serie dei due elementi. I dati tuttavia contraddicono questa ipotesi: è possibile che il contributo del GST non sia in realtà trascurabile, così come possono essere rilevanti effetti del secondo ordine dovuti alle

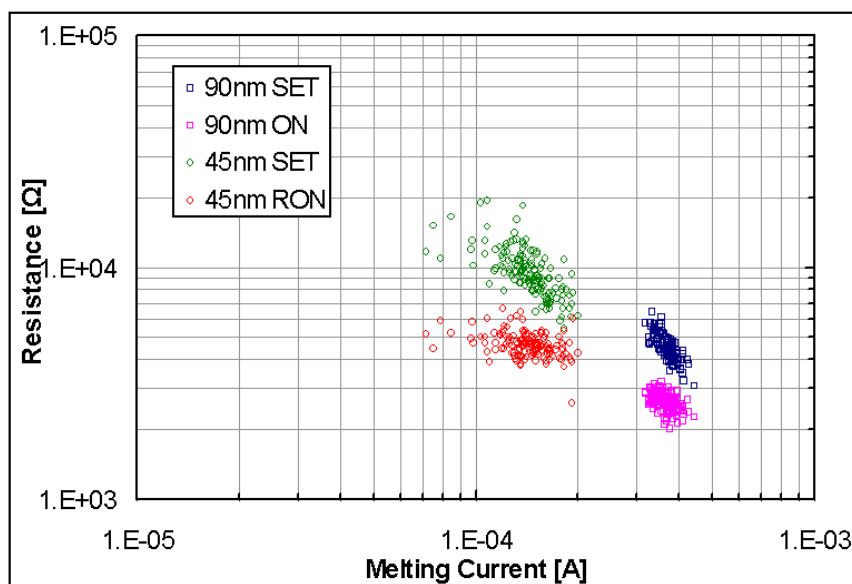


Figura 3.1: Scatter plot delle resistenze di SET e ON in funzione della corrente di melting I_m . Si evidenzia lo spread statistico che caratterizza ciascuna delle tre figure di merito e in particolare la maggiore correlazione della resistenza a bassi campi con I_m rispetto a R_{ON} , chiaramente visibile per le celle 45nm. Questa dipendenza non è prevista dal modello analitico.

alte temperature, non previsti dal modello analitico; la figura 3.2 mostra le correlazioni ottenute analiticamente, inserendo per ogni grandezza fisica e geometrica in gioco uno spread statistico del 5% se appartenente all'heater e del 10% se appartenente al GST. I risultati non coincidono con i dati sperimentali: la correlazione $R_{ON}-I_m$ è molto maggiore che nelle misure. Se effettivamente queste discrepanze fossero dovute a effetti termoelettrici del secondo ordine, il simulatore sarebbe in grado di tenerne conto. È quindi utile eseguire un'analisi di sensibilità mediante il modello numerico. Non è tuttavia possibile eseguire una simulazione statistica completa utilizzando solamente SDevice a causa dei tempi di elaborazione proibitivamente lunghi. Si è quindi pensato di adottare un approccio misto al problema: mediante il simulatore si ricavano dei coefficienti di variabilità, che mettono in relazione lo scostamento percentuale di una delle tre caratteristiche della cella rispetto al valore nominale con la variazione di alcuni parametri fisici e geometrici. Le variabilità che si andranno ad analizzare saranno quelle dovute ai seguenti parametri: distanza del livello trappola dalla banda di valenza del GST (E_{traps}), mobilità delle lacune

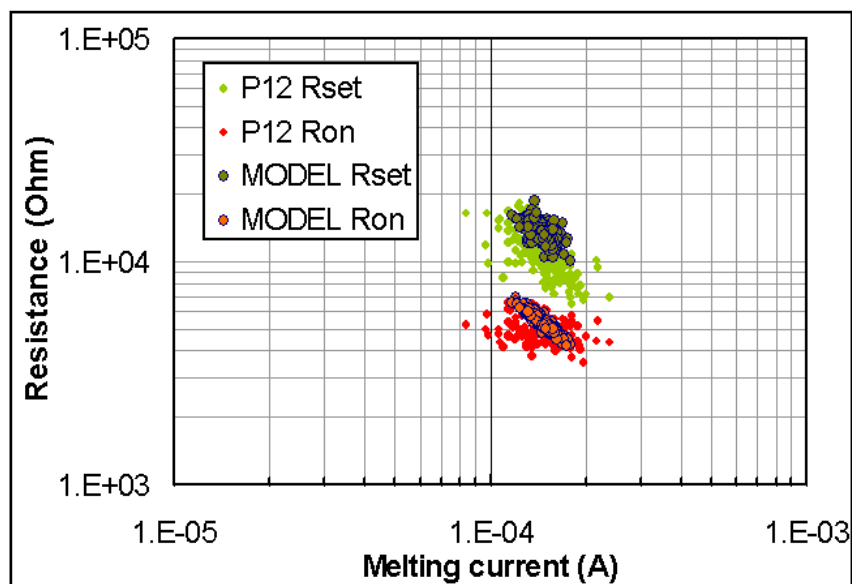


Figura 3.2: Scatter plot R - I_m ricavato dal modello analitico inserendo uno spread gaussiano del 5% (heater) o del 10% (GST) su ogni grandezza fisica e geometrica incontrata nelle formule, sovrapposto ai dati sperimentali delle celle 45nm. Si nota come il modello non riproduca qualitativamente le correlazioni degli esperimenti.

(μ_c) e conducibilità termica (K_c) nel GST, larghezza (W) e spessore (t) dell'area di contatto, lunghezza dell'heater (L_h), mobilità della lacune nell'heater (μ_h) e nell'interfaccia elettrica GST/Heater (μ_{if}). Grazie a questi coefficienti è possibile effettuare delle analisi statistiche con strumenti numerici più veloci semplicemente effettuando un'approssimazione polinomiale del secondo ordine delle caratteristiche della cella:

$$R_{SET} = R_{SET}^0 + \sum_i \Delta x_i \alpha_i^{SET} + \sum_{ij} \Delta x_i \Delta x_j H_{ij}^{SET} \quad (3.1)$$

$$R_{ON} = R_{ON}^0 + \sum_i \Delta x_i \alpha_i^{ON} + \sum_{ij} \Delta x_i \Delta x_j H_{ij}^{ON} \quad (3.2)$$

$$I_m = I_m^0 + \sum_i \Delta x_i \alpha_i^m + \sum_{ij} \Delta x_i \Delta x_j H_{ij}^m \quad (3.3)$$

	Deviazione standard	
	90nm	45 nm
R_{SET} [Ω]	550 (10%)	2052 (19%)
R_{ON} [Ω]	234 (7.5%)	681 (14%)
I_m [μA]	23 (7%)	42 (28%)

Tabella 3.1: Dispersione statistica dei parametri delle celle standard 45nm e 90nm. Lo spread più elevato nel nodo 45nm è atteso per via delle ridotte dimensioni.

In queste equazioni Δx è la variazione del parametro fisico o geometrico della cella, ed è un ingresso del sistema. α è il coefficiente di variabilità del primo ordine mentre H è la matrice Hessiana delle variabilità, che determina i coefficienti delle derivate seconde miste. I coefficienti α vengono ricavati in questo modo: si effettua una serie di simulazioni della cella standard 45nm, nelle quali vengono modificati uno per volta i parametri in ingresso, nell'intorno del valore nominale e in un range del +/- 20%. Valutando lo scostamento percentuale, approssimato al secondo ordine, delle caratteristiche rispetto a quello degli ingressi si ottengono i coefficienti del primo ordine e quelli del secondo ordine non misti. Per quanto riguarda gli elementi fuori dalla diagonale della matrice Hessiana, il procedimento è il medesimo ma vengono fatti variare due ingressi simultaneamente. La figura 3.3 mostra graficamente i risultati delle simulazioni del primo ordine, mentre la tabella 3.2 mostra i coefficienti α così ottenuti. Da questi si osserva che: le variazioni che hanno un effetto più marcato sono quelle relative all'area attiva e all'heater (W , t , L_h e μ_h), che influenzano pesantemente tutte e tre le caratteristiche della cella; le variazioni relative al GST influenzano R_{SET} e I_m , ma solo in maniera molto lieve R_{ON} . Questo conferma quanto qualitativamente già ipotizzato in precedenza riguardo al crollo della resistenza del GST per campi elettrici elevati.

Queste costanti, essendo state ottenute con la simulazione numerica calibrata nel capitolo precedente, portano con sé informazioni sulla totalità delle interazioni elettrotermiche inserite nel programma, per cui non risentono delle approssimazioni che sono necessarie nel ricavare un modello analitico. L'approccio misto simulativo-Monte Carlo ha quindi il duplice vantaggio di essere poco oneroso dal punto di vista computazionale e ben descrittivo della fisica del dispositivo. Ora l'obiettivo è quello di individuare le variazioni in ingresso

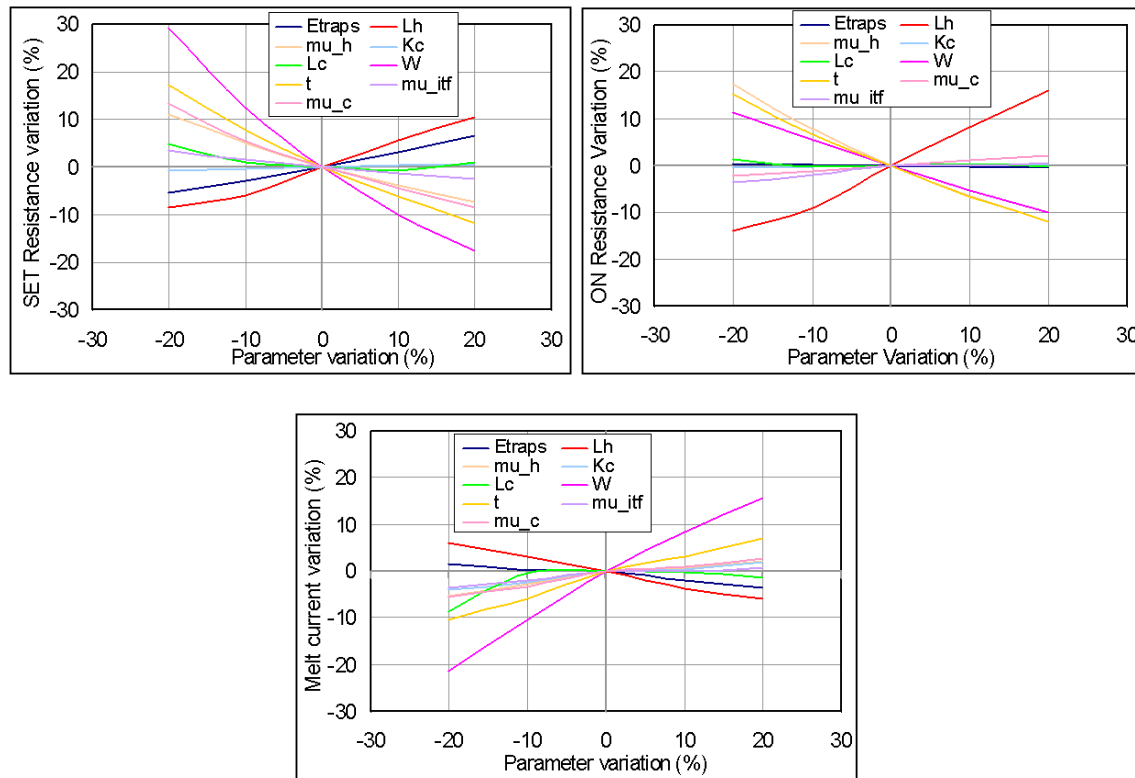


Figura 3.3: Sensibilità dei parametri R_{SET} , R_{ON} e I_m alle variazioni di alcune proprietà fisiche e geometriche della cella, espresse in termini percentuali.

al sistema che riproducono i dati sperimentali, in modo da determinare i contributi principali che portano a correlazioni del tipo illustrato dalla figura 3.1.

Come prima prova si impongono degli spread gaussiani relativamente bassi (inferiori al 5% del valore nominale) su tutte le variabili in ingresso (*caso 1*). Dalla figura 3.4.a si osserva che i risultati non sono compatibili con i dati sperimentali: la dispersione statistica risulta essere troppo bassa e la correlazione $R-I_m$ non viene ben riprodotta. È quindi necessario introdurre delle variazioni differenti. Si è dunque provato a inserire delle variazioni più elevate nelle proprietà fisiche del GST (*caso 2*). Come si nota nella figura 3.4.b si riesce a riprodurre correttamente la correlazione. Altri valori di deviazione standard della resistività e dell'energia di attivazione sono documentati in letteratura [16] e prevedibili

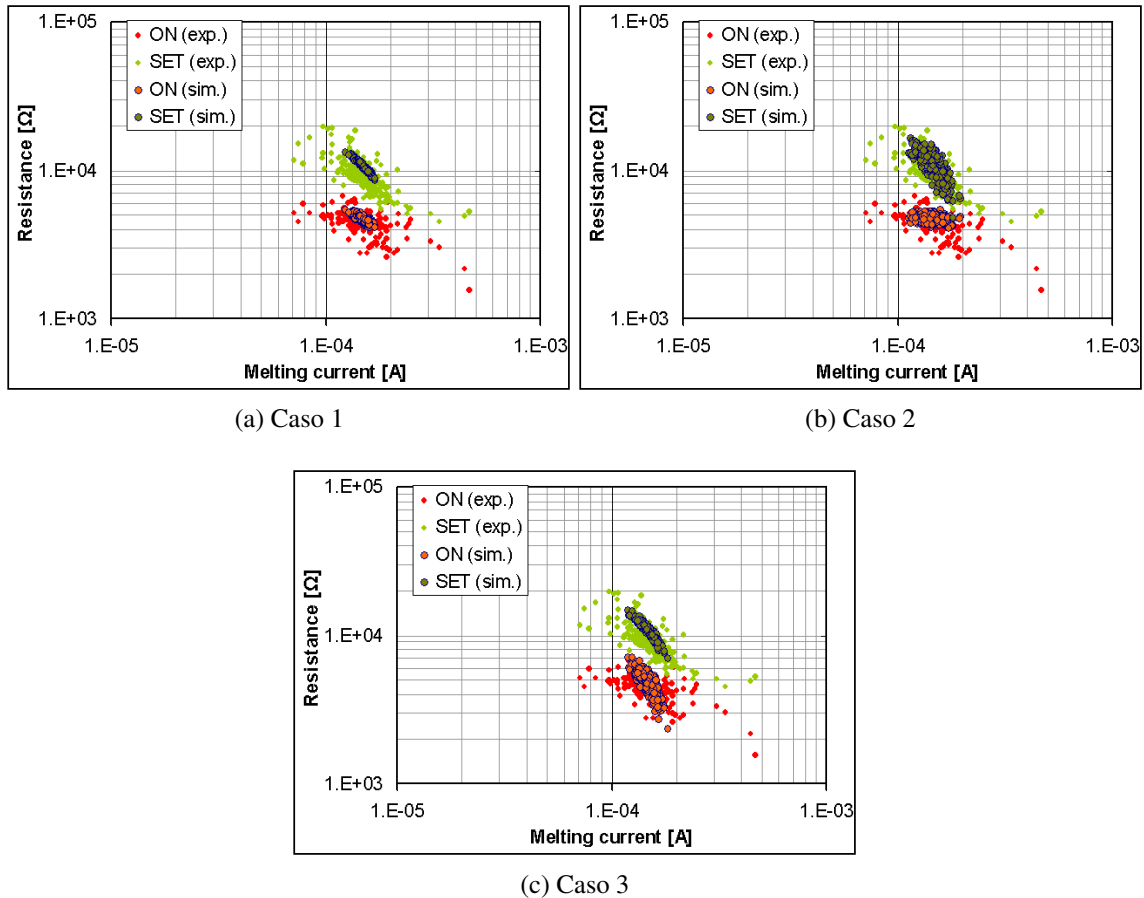


Figura 3.4: Scatter plot $R-I_m$ nei tre casi considerati. *Caso 1*: bassi valori di dispersione per tutte le variabili. Non è possibile in questo modo riprodurre i dati sperimentali. *Caso 2*: elevati valori di dispersione per le variabili relative al GST. È possibile in questo caso riprodurre la correlazione osservata sperimentalmente. *Caso 3*: elevati valori di dispersione per le variabili relative all'heater. Anche in questo caso le misure non vengono ben riprodotte.

	R_{ON}	R_{SET}	I_m
W	-0.53	-1.16	0.93
t	-0.67	-0.72	0.44
L_h	0.77	0.49	-0.31
L_c	-0.02	-0.09	0.15
E_{traps}	-0.021	0.29	-0.12
K_c	0.02	0.03	0.15
μ_h	-0.73	-0.46	0.19
μ_c	0.02	-0.53	0.21
μ_{itf}	-0.07	-0.14	0.11

Tabella 3.2: Tabella riassuntiva dei coefficienti relativi di sensibilità alle variazioni delle proprietà della cella. I coefficienti indicano la frazione di percentuale di variabilità della variabile in ingresso con cui varia il parametro considerato.

dai dati sperimentali: si è infatti sempre osservata una maggiore dispersione di R_{SET} , che comprende i contributi del calcogenuro e dell'heater, rispetto a R_{ON} , che contiene al suo interno solamente informazioni sull'heater. Come ulteriore prova, si è aumentata la deviazione standard dei parametri legati solamente al riscaldatore (L_h , μ_h e μ_{itf}). Ancora una volta, non si riesce a riprodurre la correlazione sperimentale (figura 3.4.c). Come ulteriore conferma del ruolo primario svolto dal GST nel determinare le relazioni tra R e I_m osservate si è deciso di effettuare delle misure *in ciclatura* su un gruppo di circa 50 celle. Queste misure consistono nel raccogliere le informazioni sulle celle dopo aver effettuato un numero piuttosto elevato (diecimila) di operazioni di SET e RESET consecutive, per poi confrontarle con quelle ottenute prima della ciclatura. Confrontando le resistenze di SET e di ON nei due casi si può inferire l'effetto dell'operazione della cella sull'heater e sul GST. I risultati di questo esperimento sono mostrati in figura 3.5. Non vi sono differenze significative tra le celle misurate prima e dopo la ciclatura. Tuttavia, osservando lo scatter plot delle resistenze misurate si osserva che nel caso di R_{ON} la correlazione tra la resistenza prima e dopo la ciclatura ha una pendenza unitaria, mentre nel caso di R_{SET} essa è minore. Questo significa che la ciclatura non ha avuto nessun effetto apprezzabile sul riscaldatore, mentre ha provocato un cambiamento nel GST, confermando che quest'ultimo è maggiormente affetto da variabilità statistica nel corso dell'operazione del dispositivo.

Variabile	Deviazione standard (%)		
	Caso 1	Caso 2	Caso 3
W	5	5	5
t	5	5	5
L_h	2	2	15
L_c	2	5	5
E_{traps}	3	25	3
K_c	3	35	3
μ_h	3	3	15
μ_c	3	30	2
μ_{itf}	3	5	5

Tabella 3.3: Deviazione standard (espressa in percentuale del valore nominale) dei parametri gaussiani in ingresso nei tre casi presi in considerazione. *Caso 1*: basso spread su tutte le variabili. *Caso 2*: fitting della correlazione $R-I_m$ inserendo un'alta variabilità sui parametri relativi al GST. *Caso 3*: alto spread solo sui parametri relativi all'heater.

È possibile quindi concludere l'analisi di sensibilità appena presentata affermando che per ottenere dei dispositivi affetti il meno possibile da variabilità è necessario agire principalmente sull'area attiva della cella e sulla qualità del materiale calcogenuro.

3.2 Stima delle prestazioni di scaling

La scalabilità di un dispositivo è una delle caratteristiche più importanti per il mercato, in quanto determina per quanto tempo un dispositivo resterà competitivo. Essa ha pesantemente influenzato la decisione di cominciare lo sviluppo delle memorie elettriche a cambiamento di fase. Nei dispositivi tradizionali basati sulla tecnologia CMOS i problemi principali incontrati durante la riduzione delle dimensioni riguardano fenomeni come breakdown degli ossidi, insorgere di effetti di canale corto oppure quantistici che ne degradano le prestazioni. La cella PCM opera a tensioni relativamente basse e il suo funzionamento è dettato da fenomeni elettrotermici ampiamente documentati. Al momento si pensa che lo scaling della memoria PCM sia limitato unicamente dal processo, ovvero dalla dimensione minima della maschera litografica; pertanto, in linea di principio, la cella a cambiamento di fase è liberamente scalabile fino alla dimensione consentita dalla tecnologia di fabbri-

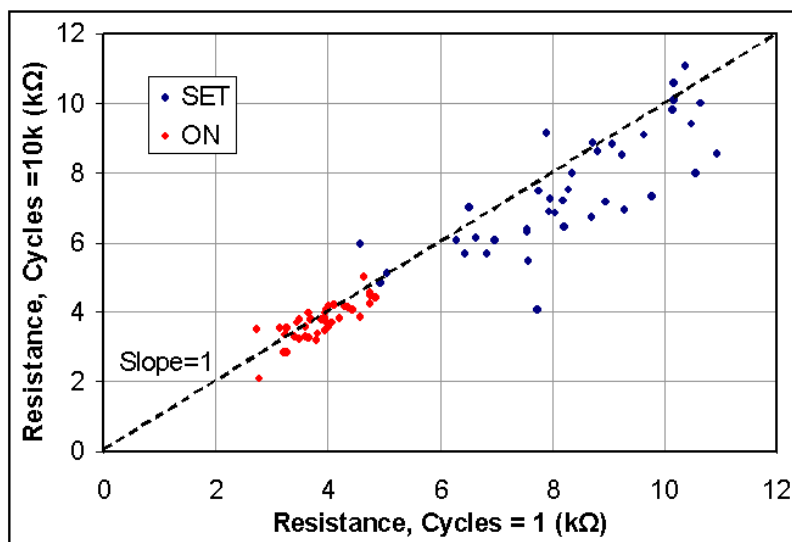


Figura 3.5: Correlazione tra le resistenze di SET e di ON prima e dopo un ciclo di 10.000 operazioni di programmazione e cancellazione. Questa operazione non ha effetto su R_{ON} ma sembra influenzare la resistenza di SET, a ulteriore conferma che il GST è più suscettibile a variazioni statistiche rispetto al riscaldatore.

cazione, e non limitato da fenomeni fisici insormontabili.

Come già mostrato nel capitolo 1, lo scaling delle dimensioni della cella ha l'effetto di aumentare la resistenza di lettura e diminuire la corrente di melting: questo consente di migliorare la densità di celle in un array e di ridurre il consumo di potenza, a scapito di una finestra di lettura più piccola. Si è già proposto un modello analitico di scaling, utile per stimare l'andamento delle caratteristiche della PCM al variare del nodo tecnologico nei due casi di scaling isotropo e anisotropo. La tecnica di scaling adottata nella produzione dei dispositivi reali è fortemente anisotropa: il fattore di scala è diverso per tutte le grandezze, e varia a seconda del nodo da cui si parte per ridurre le dimensioni. Le tabelle 3.4 mostrano l'approccio usato da Numonyx. I dispositivi 90nm e 45nm sono già esistenti, per il nodo 32nm è ad oggi in fase di sviluppo, e sono riportate le dimensioni target per tale nodo. Come si è già avuto modo di dimostrare, il modello analitico non riesce sempre a spiegare il comportamento delle celle, e questo vale anche per i problemi di scaling. Il modello numerico calibrato può quindi essere uno strumento utile per avere un'idea quantitativa sulle prestazioni dei futuri nodi tecnologici. Si è quindi eseguita una

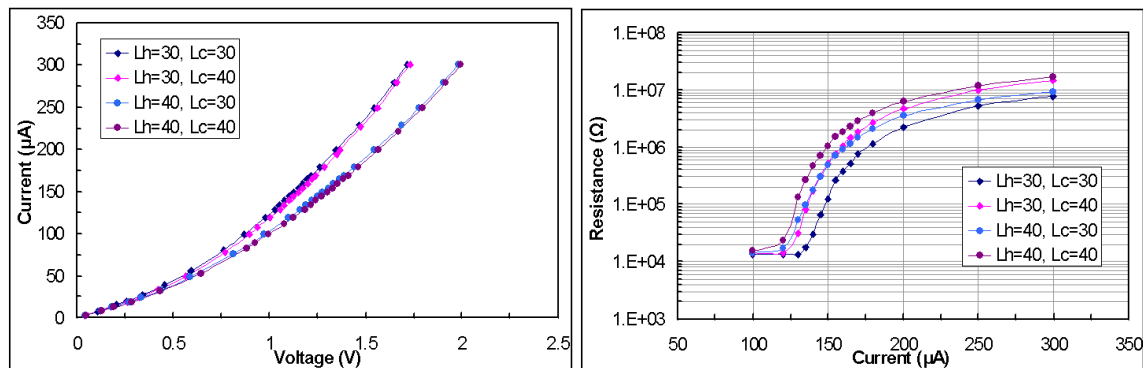
	Dimensione [nm]				Fattore di scala	
	90nm	45nm	32nm		90nm→45nm	45nm→32nm
W_h	160	58	38	W_h	2.76	1.53
W_c	140	50	32	W_c	2.8	1.56
t	5	5	5	t	1	1
L_h	85	50	30	L_h	1.7	1.67
L_c	70	50	30	L_c	1.4	1.67

Tabella 3.4: Dimensioni delle celle nei tre nodi tecnologici presi in considerazione (sinistra) e fattore di scala (destra). La tecnica di scaling è fortemente anisotropa, soprattutto nel passaggio da 90nm a 45nm, in cui si è ridotta molto di più la larghezza della lunghezza. Lo spessore dell'area attiva è sempre costante.

simulazione su una cella PCM 32nm, caratterizzata dalla stessa struttura usata in precedenza ma con le dimensioni riportate in tabella 3.4 e dai parametri numerici calibrati nel capitolo 2.

Sono state esaminate quattro strutture, simili tra loro, che si differenziano per l'altezza dell'heater e del GST: esse sono le quattro combinazioni possibili delle due altezze con i valori di 30 nm e 40 nm. I risultati sono mostrati in figura 3.6. Si osserva subito che lo scaling ha avuto l'effetto atteso di innalzare la resistenza, sia di SET che di ON, e diminuire la corrente di programmazione. Inoltre, anche per il nodo 32nm, l'altezza di heater L_c non ha un effetto rilevante sulla curva I-V, in quanto influenza marginalmente la resistenza di SET (logaritmicamente) e per nulla la resistenza di ON. Invece il suo effetto non è trascurabile sulla corrente di fusione, in quanto altera la resistenza termica del GST e la distanza tra l'interfaccia e il boundary termico superiore.

A questo punto è possibile confrontare i risultati della simulazione con quelli previsti dal modello analitico. La figura 3.7 mostra che, sebbene il modello analitico sia in grado di prevedere qualitativamente l'andamento delle grandezze in gioco all'aumentare dello scaling della cella, il modello numerico fornisce dei valori quantitativamente più vicini a quelli sperimentali. Si può quindi affermare che il modello sia anche più predittivo per quel che riguarda i nodi tecnologici futuri. In particolare, viene messa in luce la maggiore riduzione della corrente di melting rispetto a quanto previsto dal modello analitico. Questa analisi conferma le caratteristiche di scalabilità che hanno reso le PCM un potenziale sostituto



L_h [nm]	L_c [nm]	R_{SET} [Ω]	R_{ON} [Ω]	I_m [μA]
30	30	15043	3960	136
30	40	15763	3981	131
40	30	16296	5460	126
40	40	16588	5415	122

Figura 3.6: Sopra: caratteristiche IV ed RI simulate per le celle del nodo 32nm per alcune combinazioni di altezza di heater e GST. Sotto: riassunto dei parametri caratteristici delle celle simulate.

delle memorie attualmente in commercio.

3.3 Studio di geometrie alternative

Oltre alla cella di tipo *Wall* trattata fino a questo momento, esistono numerose proposte alternative per la realizzazione delle memorie PCM. Alcuni esempi sono la cella *microtrench*[17], la cella *dash-confined*[18] e la cella *damascata*. La cella **damascata**[19] è caratterizzata da un foro a pianta cilindrica dentro al quale viene depositato il GST, che si troverà direttamente a contatto con gli elettrodi. Con questa geometria non viene fatto uso dell'heater, e l'intero funzionamento della cella è dovuto al self-heating del calcogenuro. L'intero cilindro, di diametro corrispondente al nodo tecnologico, è confinato da materiali isolanti, aumentando notevolmente l'efficienza termica della cella rispetto all'architettura *Wall*. È stato deciso di simulare quest'ultima struttura, in quanto caratterizzata da un principio di fun-

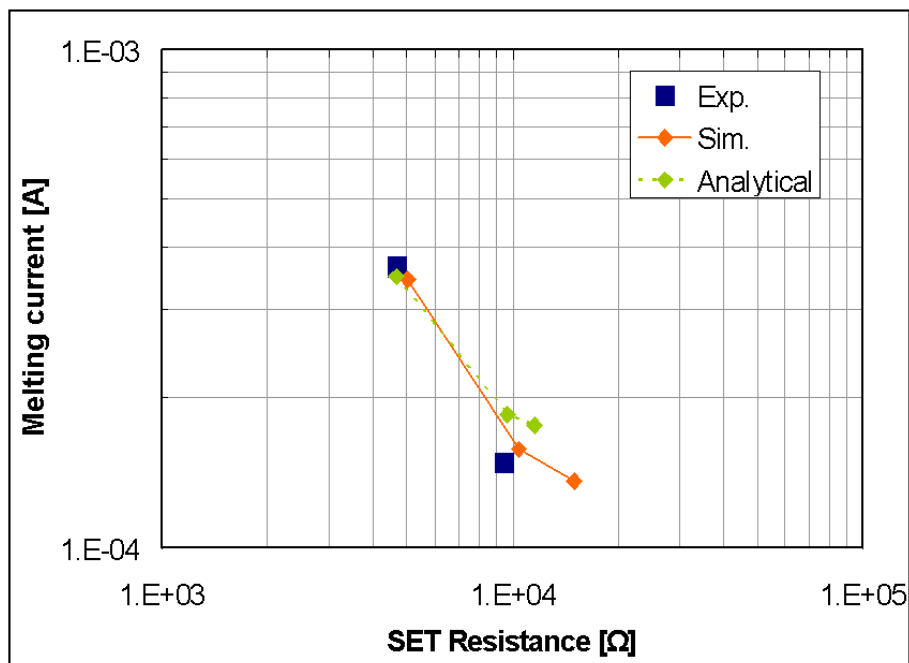


Figura 3.7: Scaling di R_{SET} e I_m come previsto dal modello analitico e dalla simulazione numerica, sovrapposto ai dati sperimentali per i nodi 90nm e 45nm. Il modello a elementi finiti risulta quantitativamente migliore di quello analitico.

zionamento diverso rispetto alle celle di tipo Wall, e di maggior interesse industriale per Numonyx. I risultati saranno confrontati con dati reperiti in letteratura, in quanto questo tipo di cella al momento non viene prodotto localmente.

La struttura simulata numericamente è piuttosto semplice: consiste in un cilindro di GST di diametro 32nm, con altezze L_c variabili (32, 64 e 96nm), collegato ai contatti esterni mediante due plug fatti di un materiale perfettamente conduttivo, sia termicamente che elettricamente. Il cilindro è interamente circondato dall'ossido, e viene inclusa la resistenza termica interfacciale calibrata in precedenza. Per prima cosa è bene avere un'idea qualitativa di ciò che ci si attende dalla simulazione. La resistenza di SET è facilmente calcolabile:

$$R_{SET} = \rho_c \cdot \frac{L_c}{\pi r^2} \quad (3.4)$$

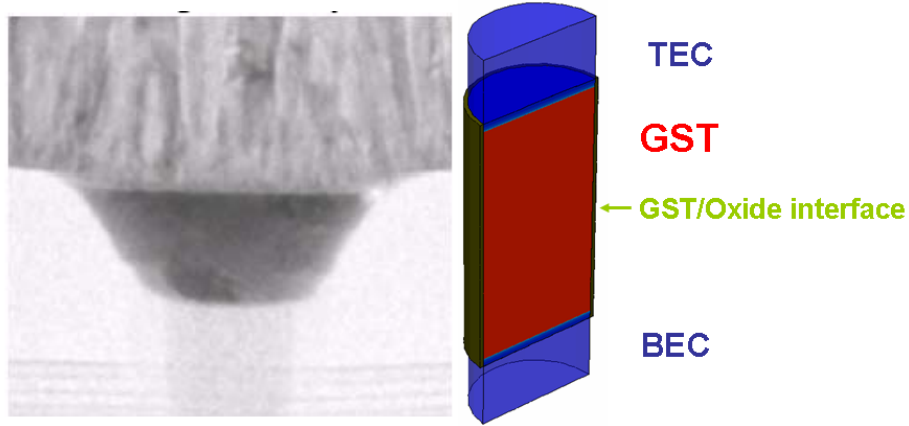


Figura 3.8: Immagine TEM di una cella damascata (sinistra) e schema della struttura usato nella simulazione [15]. La pianta della colonna di GST è circolare.

in cui r indica il raggio del cilindro di GST. Ci si attende quindi una resistenza di SET nell'ordine della decina dei $k\Omega$. Per quanto riguarda la R_{ON} , la resistenza ad alti campi non è più limitata dall'heater e dall'interfaccia GST/heater, e dipenderà solo dal GST. Per via dell'attivazione della conducibilità si avrà quindi una R_{ON} molto ridotta. Per avere una stima del prim'ordine della corrente di melting si può risolvere l'equazione del calore nel caso monodimensionale, in cui si tiene conto della struttura imponendo una sezione circolare nel calcolo della potenza generata e impiegando un parametro α che descrive l'efficienza del riscaldamento Joule. Si ha:

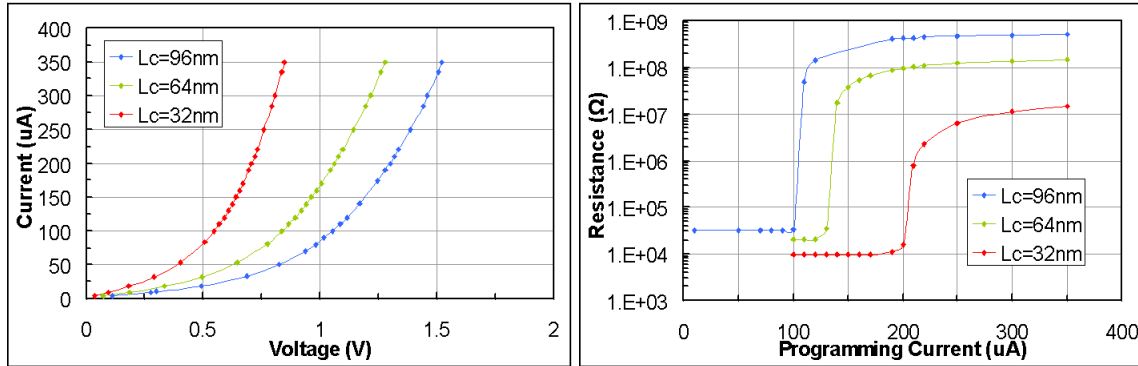
$$\frac{d^2T}{dx^2} = -Q \quad (3.5)$$

in cui $Q = \frac{I^2 \rho_c \alpha}{A_c}$ e $A_c = \pi r^2$. Imponendo la condizione $T = T_0$ ai bordi superiore e inferiore si ottiene:

$$T(x) = T_0 + \frac{Q}{2} \left(\frac{Lc^2}{4} - x^2 \right) \quad (3.6)$$

da cui si trova che nel punto di massima temperatura, ovvero al centro del cilindro di GST ($x = 0$) si ha:

$$T(0) = T_0 + \frac{QLc^2}{8} \quad (3.7)$$



L_c [nm]	R_{SET} [Ω]	R_{ON} [Ω]	I_m [μ A]
32	10901	683	206
64	22749	914	148
96	35363	1009	113

Figura 3.9: Sopra: caratteristiche IV ed RI delle celle damascate simulate. Sotto: riassunto dei parametri caratteristici.

Facendo uso del modello elettrico per il riscaldamento del dispositivo si può definire la resistenza termica come $R_{th} = \frac{\alpha L_c}{8A_c K_c}$. Dalla relazione

$$I_m = \frac{A_c}{L_c} \sqrt{\frac{8(T_m - T_0) K_c}{\rho_c \alpha}} \quad (3.8)$$

Il parametro α deve tenere conto sia della perdita di calore dovuta agli isolanti non ideali sia della diminuzione di resistività del GST a causa della generazione termica dei portatori. Inoltre, anche la conducibilità elettrica del GST non è pari a quella ideale per via della relazione di Wiedemann-Franz. Ci si attende qualitativamente un'andamento iperbolico di I_c con L_c e dei valori compresi circa tra 100 e 200 μ A. Per via del miglior confinamento termico (il GST è interamente circondato da ossidi a una distanza pari al suo raggio) ci si aspetta inoltre che la curva di programmazione RI sia più ripida che nel caso della cella Wall. La figura 3.9 mostra le curve IV ed RI delle celle damascate così simulate e il valore numerico dei loro parametri caratteristici. I valori numerici sono concordi con quanto atteso dal modello analitico. Quest'ultimo è più preciso rispetto a quello elaborato

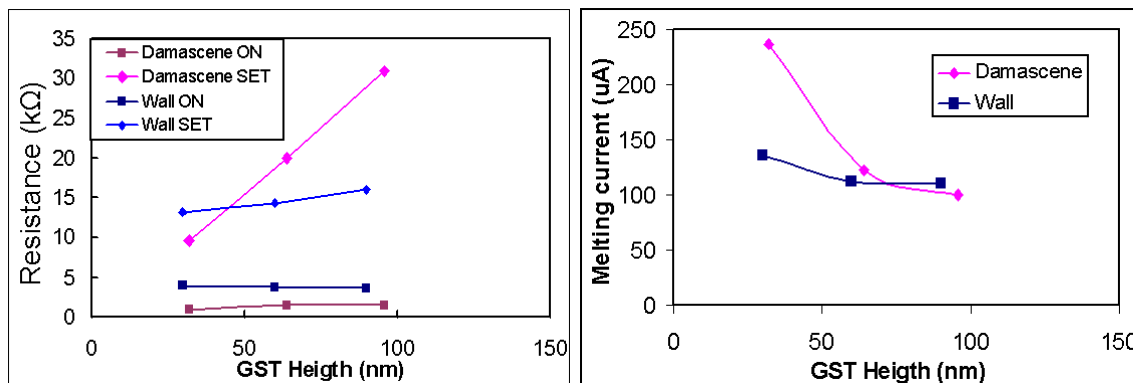


Figura 3.10: Confronto tra le caratteristiche salienti della cella damascata e della cella Wall 32nm.

per la cella Wall in quanto la geometria, nel caso damascato, è più semplice e l'assenza dell'heater rimuove i problemi relativi alle interfacce elettriche. Osservando la curva di programmazione si osserva che la cella passa dallo stato ad alta conducibilità allo stato altamente resistivo in un range di correnti di programmazione decisamente ridotto, segno che il confinamento termico di questo tipo di cella è molto efficace. La figura 3.10 mostra l'andamento di R_{SET} , R_{ON} e I_m in funzione dell'altezza del GST, confrontato con la cella 32nm simulata nella sezione precedente.

Si nota la dipendenza perfettamente lineare della resistenza di lettura da L_c , molto differente da quella logaritmica presente nella cella Wall. Inoltre, la dipendenza iperbolica della corrente di fusione dalla lunghezza del GST fa sì che la I_m per L_c minori di 64nm sia assai elevata: oltre 200μA, da confrontare con i 126μA presentati dalla cella Wall. Questo è dovuto all'avvicinarsi dei boundary termici al centro del cilindro, dove si ha il maggiore aumento di temperatura: risulta infatti molto più difficile scaldare il GST quando esso è così vicino ai contatti idealmente conduttivi. Per via della geometria della cella, risulta sempre verificata la relazione $I_m \propto R_{SET}^{-1}$, per cui la pendenza di scaling sul grafico R_{SET} - I_m è sempre -1. In base a queste considerazioni, lo scaling delle celle damascate può rappresentare un problema, in quanto impone un considerevole trade-off tra la corrente di programmazione e la corrente di lettura. Un'ulteriore considerazione riguarda la variabilità statistica dei parametri. Come osservato nella sezione precedente, i parametri relativi al GST sono caratterizzati da un forte spread statistico. Nella cella Wall essi con-

tribuiscono solo in parte a determinare la variabilità finale dei parametri della cella, che sono influenzati anche dalle proprietà dell'heater. In questo caso, invece, non essendoci il riscaldatore la variabilità della cella coincide completamente con quella del GST, e risulterà essere conseguentemente più alta.

3.4 Analisi della ricristallizzazione parziale

Come accennato discutendo dei principi di funzionamento della cella PCM, la programmazione della cella nello stato di RESET consiste nel fondere il materiale attivo portandolo sopra alla sua temperatura di melting T_m e raffreddandolo (quenching) rapidamente in modo che non vi sia il tempo sufficiente a formare i legami cristallini. Se il quenching fosse idealmente rapido, tutta la parte di GST che è stata fusa dovrebbe trovarsi, una volta raggiunta la temperatura ambiente, nella fase amorfa. Il raffreddamento non è tuttavia istantaneo, e ci sarà un certo lasso di tempo in cui la temperatura del calcogenuro è sufficiente ad attivare un processo di crescita reticolare a partire dal materiale cristallino circostante. Si avrà quindi che la dimensione della calotta amorfa risultante dall'operazione di RESET non coinciderà con la parte di GST che è stata portata a fusione. Nell'uso del modello numerico fino a questo momento, per via dell'approssimazione di quasi-stazionarietà, non è stato mai indagato il transitorio di raffreddamento per studiare questo fenomeno, in quanto si ha sempre inizializzato come "amorfo" ogni punto del materiale che supera la temperatura T_m . Presso i laboratori di fisica Numonyx sono stati condotti alcuni esperimenti, basati sulle immagini TEM, per caratterizzare la ricristallizzazione parziale. In queste immagini è possibile riconoscere sia un "fronte di fusione", individuabile in base alla presenza di bolle di gas Argon formatesi nel corso della fusione, sia la calotta amorfa risultante, chiaramente distinguibile dal GST cristallino che la circonda. La figura 3.11 mostra un esempio di queste immagini. I dati sperimentali riguardanti queste osservazioni si trovano in figura 3.12. Essi mostrano chiaramente che la parte di GST che è stata portata a fusione non si è interamente amorfizzata. Il fatto che l'altezza della parte di materiale fuso non sia monotona con la corrente di programmazione può essere indice di un qualche tipo di saturazione. Infatti, si potrebbe avere che tutto il GST è stato fuso nella direzione verticale, per cui ulteriori aumenti di temperatura non provocano alcun effetto.

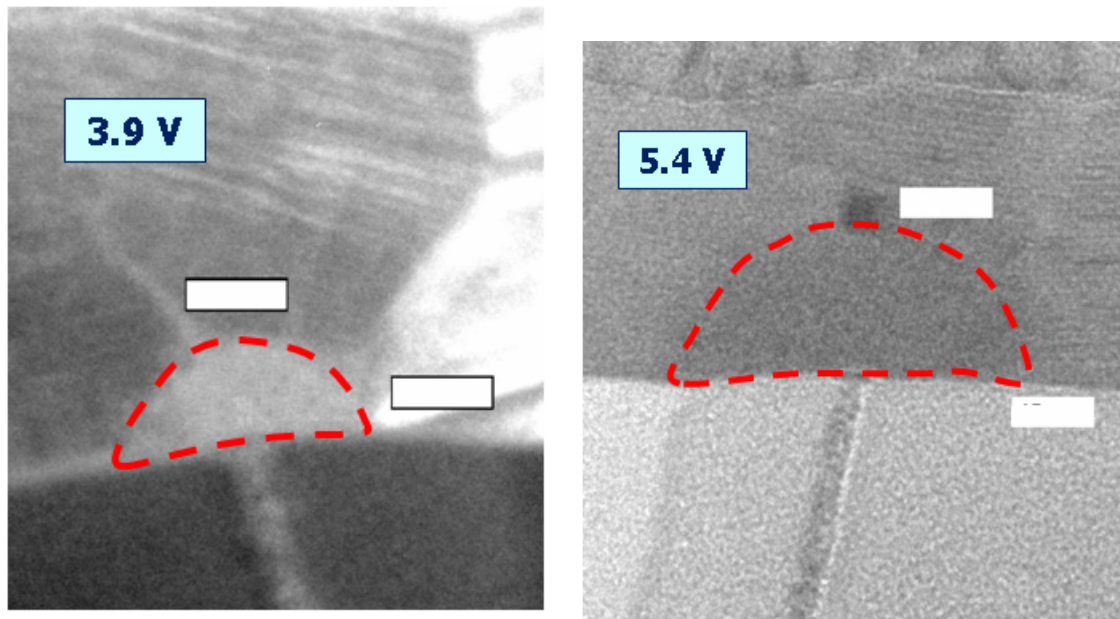


Figura 3.11: Immagini TEM della cella PCM 90nm in prossimità dell'interfaccia, per diverse tensioni di programmazione. È chiaramente visibile la calotta di materiale amorfo.

Il modello numerico calibrato può essere utilizzato per indagare la ricristallizzazione. Per farlo è necessario eseguire una simulazione del dispositivo in transitorio, durante il quale viene applicato un impulso di RESET, descritto nel capitolo 1. Si è interessati in particolare al fronte di discesa, in cui avviene il quenching. Il simulatore utilizza per la cristallizzazione il modello di transizione di fase illustrato nel capitolo 1, i cui relativi parametri sono stati inseriti in base a calibrazioni precedenti[1], lasciando però come parametro di fitting la velocità di cristallizzazione.

La struttura della cella standard 90nm utilizzata fino ad ora viene simulata come parte di un semplice circuito, illustrato in figura 3.13.a, costituito da un generatore di tensione che applica un segnale a rettangolo a una rete RC formata da una resistenza di 100Ω in serie al parallelo di un condensatore da 300pF con la cella PCM. La costante di tempo risultante, 30ns , è stata scelta in base alle RC reali dei circuiti di programmazione delle

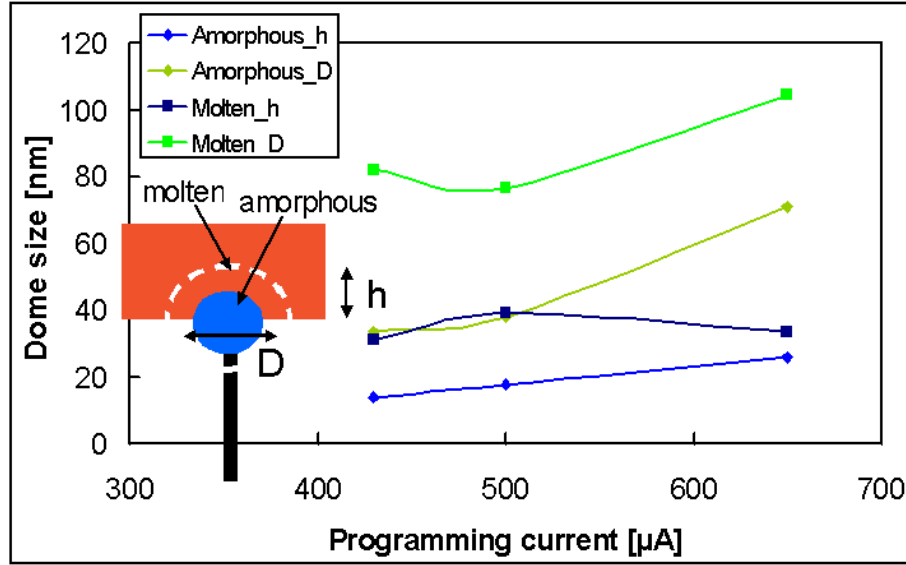


Figura 3.12: Misure sperimentali tratte dalle immagini TEM della dimensione della zona di GST fusa e della calotta di materiale amorfo. I dati si riferiscono a cella a 90nm.

celle analitiche. La velocità di cristallizzazione è data dalla formula

$$V_x = p_0 \alpha \cdot \exp\left(-\frac{0.8}{1 - T/T_m}\right) \cdot \left(1 - \exp\left(-\frac{\Delta E}{kT}\right)\right) \cdot \left(-\frac{E_a}{kT}\right) \quad (3.9)$$

il cui andamento in funzione della temperatura è mostrato in figura 3.13.b. In questa formula ΔE ed E_a sono parametri il cui valore è stato preso da [1], mentre il prodotto $p_0 \alpha$ rappresenta la velocità di cristallizzazione. La simulazione verrà confrontata sia con i dati sperimentali sia con un semplice modello analitico, utile per avere un'idea di massima dei risultati da attendersi, che verrà ora illustrato. Per semplicità, il modello tratterà un caso monodimensionale, concentrandosi sull'altezza della calotta amorfa e trascurandone le dimensioni laterali. Una prima ipotesi semplificativa consiste nel considerare la generazione di calore all'interno del GST pari a 0, imputandola completamente all'heater. In questo caso il profilo spaziale di temperatura all'interno del GST nell'istante $t=t_0$, coincidente con l'inizio del fronte di discesa, ha forma lineare, come stabilito dall'equazione di trasporto

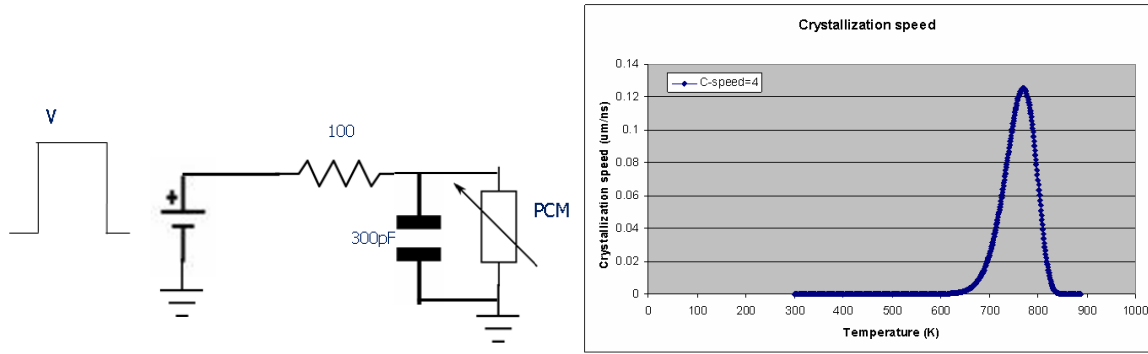


Figura 3.13: Sinistra: schema del circuito usato per la simulazione in transitorio. Destra: andamento della velocità di cristallizzazione in funzione della temperatura.

del calore. Considerando la differenza di temperatura rispetto a T_0 si ha:

$$\Delta T(x, t) = \Delta T_{\max} \left(1 - \frac{x}{L_c} \right) \exp\left(-\frac{t}{\tau}\right) \quad (3.10)$$

in cui ΔT_{\max} è la temperatura raggiunta dall'interfaccia, τ la costante di tempo termica, che si ipotizza essere uguale a quella elettrica. L'origine della coordinata x è posta all'interfaccia GST/heater. Il fatto che la temperatura dipenda dallo spazio e dal tempo in una forma a variabili separate è un'ipotesi derivata dalle simulazioni, in cui si osserva che la temperatura di ogni punto del dispositivo tende a tornare alla temperatura ambiente con un andamento esponenziale indipendente dalla posizione (ovvero dalla temperatura iniziale). Spegnendo l'impulso di corrente, la temperatura tende a tornare, esponenzialmente in ogni punto del profilo, al valore T_0 . Durante il raffreddamento verrà attraversata la temperatura T_x di cristallizzazione, in cui si ha il massimo della velocità di crescita (figura 3.13.b). Ora si effettua un'ulteriore approssimazione: si ipotizza che la crescita del cristallo avvenga unicamente alla temperatura T_x (approssimando quindi la funzione descritta dalla 3.9 con una delta di Dirac) con una velocità efficace V , che tiene conto della crescita alle temperature prossime a quella di massimo. Il punto del profilo che si trova alla temperatura T_x si

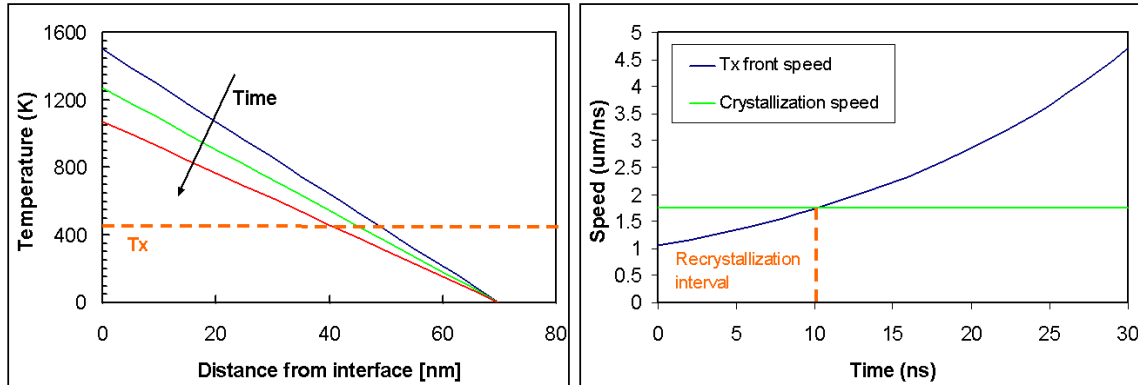


Figura 3.14: Sinistra: profilo di temperatura nel GST secondo il modello lineare in vari istanti di tempo. Il fronte a temperatura pari a T_x si sposta verso l'interfaccia. Destra: velocità di spostamento del fronte confrontata con la velocità di crescita del cristallo. L'intervallo di ricristallizzazione si ha per $V_c > V_t$.

sposta verso l'interfaccia, in base alla 3.10, con la velocità:

$$V_t = -\frac{T_x L_c}{\tau T_{\max}} \exp\left(\frac{t}{\tau}\right) \quad (3.11)$$

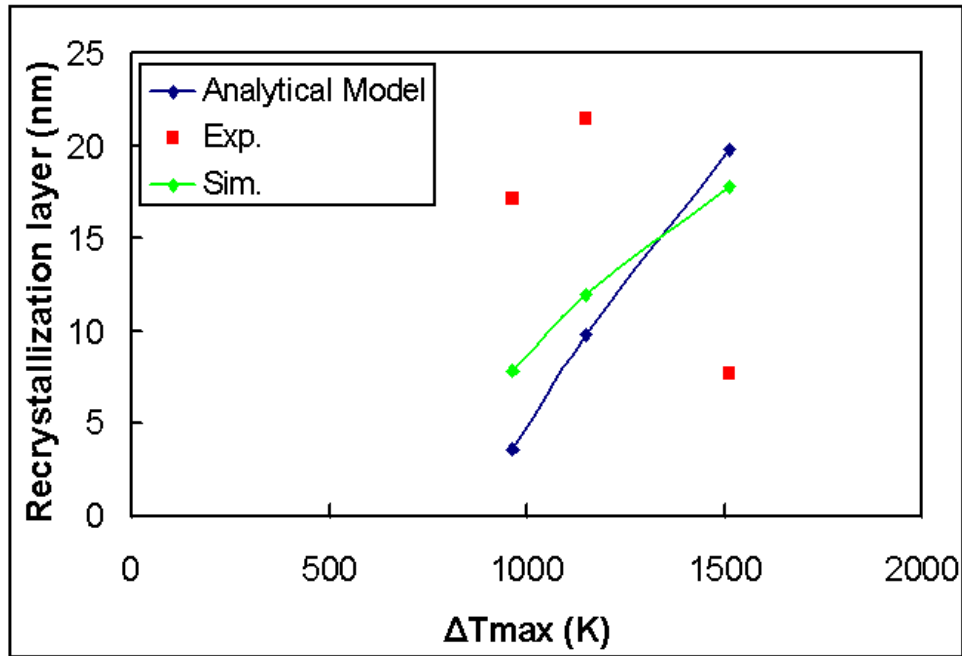
Fintanto che la velocità del fronte a temperatura T_x sarà inferiore a quella di cristallizzazione, si avrà crescita del reticolo, e ricristallizzazione parziale. Nel momento in cui la crescita non è in grado di seguire il calo di temperatura, essa si ferma, e il GST fuso rimanente amorfizzerà. Si ha quindi che la parte di GST fuso ma ricristallizzato ha dimensione:

$$L_x = V\tau \ln\left(\frac{V\tau T_{\max}}{T_x L_c}\right) \quad (3.12)$$

La figura 3.14 mostra graficamente il modello appena illustrato. L'equazione 3.12 mette in relazione la lunghezza del tratto di GST ricristallizzato alla massima temperatura raggiunta dall'interfaccia durante la programmazione. Ottenere la dipendenza di questa temperatura in modo preciso non è semplice: si è visto (ad esempio in figura 2.5) che il semplice modello termico equivalente non è adatto a una descrizione quantitativa del riscaldamento. Un'analisi accurata prevederebbe la risoluzione completa dell'equazione di trasporto del calore, troppo onerosa. Si è pertanto usato il simulatore per ottenere l'intervallo di tem-

perature T_{max} corrispondenti alle correnti di programmazione usate nei dati sperimentali. Per correnti tra $430\mu\text{A}$ e $650\mu\text{A}$ la temperatura va da 1266 K a 1815 K. A questo punto si è lanciata una serie di simulazioni per cercare di ottenere un fitting dei dati sperimentali. Con una velocità di cristallizzazione pari a 1.75nm/ns si sono ottenuti i risultati mostrati in figura 3.15. Il modello analitico e la simulazione non concordano quantitativamente ma hanno la stessa monotonia, che però non è in grado di spiegare il comportamento sperimentale, che, a causa della saturazione dell'altezza di GST fuso a L_c , prevede una diminuzione equivalente della dimensione della zona ricristallizzata. Infatti si nota dalla minore dimensione della zona fusa che le temperature raggiunte nella simulazione sono più basse che nel dispositivo reale.

Quanto osservato finora suggerisce sia che il modello calibrato può ancora essere migliorato per tenere conto di questi fenomeni, ma che è uno strumento estremamente utile per analizzare comportamenti della cella PCM difficilmente osservabili direttamente, nonché per confermare o per perfezionare i modelli analitici. Non è stato possibile per ragioni di tempo proseguire l'analisi di questo aspetto, che viene quindi lasciato come possibilità di proseguimento futuro di questo lavoro.



I [μ A]	Amorfo		Fuso	
	h [nm]	D [nm]	h [nm]	D [nm]
430	15	29	22.8	53
500	19.37	41.8	31.3	65.7
650	26	76.8	43.8	114

Figura 3.15: Sopra: simulazione, dati sperimentali e modello analitico della dimensione della zona ricristallizzata. Sull'asse delle ascisse è riportato lo scostamento dal valore $T_0=300K$. Sotto: dimensioni della zona fusa e amorfizzata al termine della simulazione di transitorio.

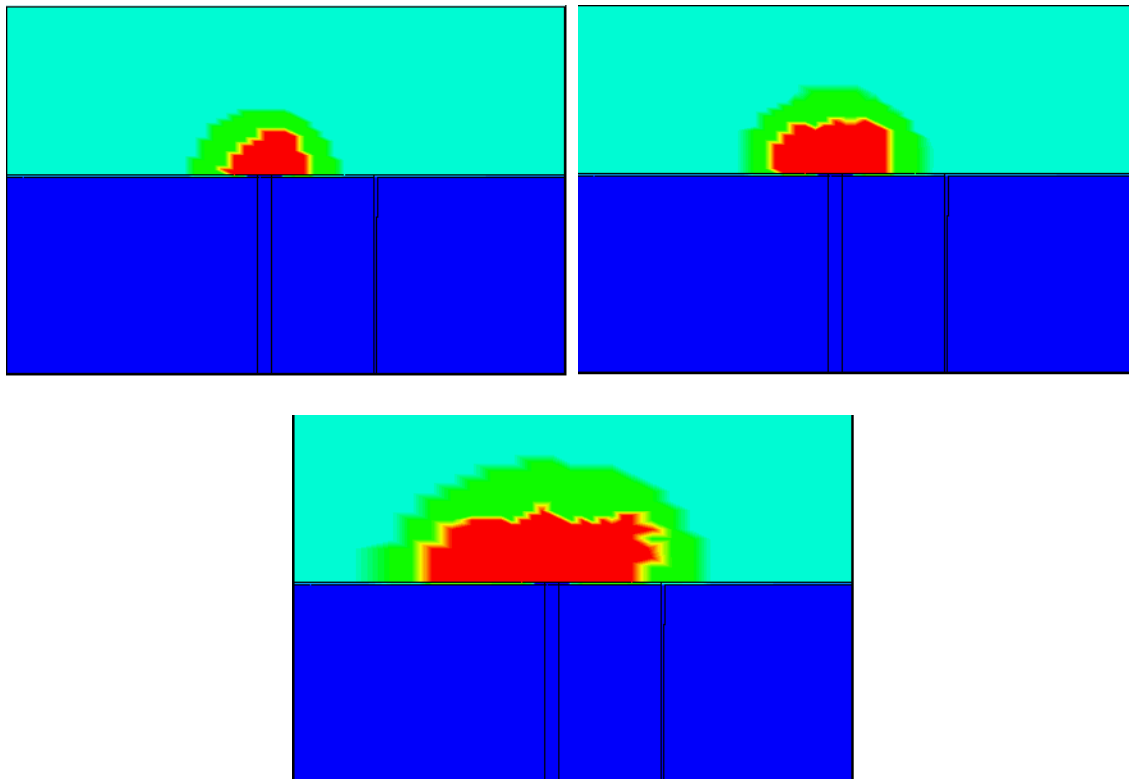


Figura 3.16: Simulazione della zona ricristallizzata e amorfizzata per diversi valori della corrente di programmazione: $430\mu\text{A}$, $500\mu\text{A}$, $650\mu\text{A}$. I colori rappresentano le fasi del GST: in celeste viene raffigurata la fase cristallina che non ha raggiunto la temperatura di fusione. Il verde rappresenta la parte di GST che si è fusa ma si è poi ricristallizzata durante il quenching. Il rosso rappresenta la calotta amorfa al termine del transitorio.

Conclusioni

In questo lavoro di tesi è stato presentato e calibrato un modello numerico a elementi finiti per la descrizione delle memorie a cambiamento di fase. La descrizione per via analitica della cella PCM non consente di riprodurre quantitativamente tutte le sue caratteristiche salienti, a causa della complessa interazione tra fenomeni elettrici e termici che avvengono nel corso dell'operazione del dispositivo. Il modello numerico è pertanto necessario per risolvere tutte le equazioni differenziali che regano i suddetti fenomeni. Queste equazioni sono caratterizzate da un vasto numero di parametri, non noti a priori, che descrivono il comportamento dei materiali presenti. Ad esempio, l'elemento di memoria, un materiale a base calcogenura in grado di cambiare il suo stato fisico da amorfo a cristallino e viceversa in tempi molto rapidi, viene descritto come un semiconduttore molto difettoso, caratterizzato da un'elevata concentrazione di stati trappola all'interno del gap energetico. Per descriverne correttamente il funzionamento è necessario specificare al simulatore una serie di parametri come la posizione energetica del livello trappola, la mobilità dei portatori, la conducibilità termica, ecc. Per calibrarli correttamente si è seguito un procedimento basato sull'estrazione diretta di più parametri possibile dai dati sperimentali, per poi dedurre i rimanenti in base a esperimenti reperiti in letteratura o grazie al fitting numerico delle curve caratteristiche delle celle. In particolare, si è osservato l'andamento della resistenza di lettura, della resistenza differenziale ad alti campi e della corrente di fusione in funzione di alcuni parametri geometrici, come l'altezza dello strato di GST o del riscaldatore, per estrarre la resistività e la conducibilità termica dei materiali in gioco. Un'importante caratteristica innovativa di questo modello è la presenza delle resistenze termiche interfacciali tra il calcogenuro e i materiali isolanti circostanti. La presenza e la rilevanza di queste interfacce è stata dimostrata da esperimenti svolti presso i laboratori

Numonyx, e in questo lavoro è stato dimostrato che il loro contributo al confinamento del calore durante l'operazione della cella ha un'influenza importante sulle sue prestazioni, osservabile dall'effetto ottenuto sulla curva di programmazione RI. Il valore numerico delle resistenze interfacciale è stato anch'esso calibrato sulla base di dati reperiti da esperimenti effettuati localmente, dalla letteratura e dal fitting numerico delle curve caratteristiche. Il modello così ottenuto è stato validato confrontandolo con dati sperimentali non utilizzati per la calibrazione, e si è mostrato che esso è in grado di riprodurre adeguatamente il comportamento delle celle appartenenti a due nodi tecnologici differenti, 90nm e 45nm.

È stato mostrato che le proprietà salienti delle celle PCM sono affette da una grande dispersione statistica, che non è possibile riprodurre qualitativamente mediante il modello analitico. In particolare, la correlazione $R_{SET}-I_m$ è più accentuata della correlazione $R_{ON}-I_m$, mentre il modello prevede l'andamento opposto. Per investigare questo fenomeno è stata sfruttata la simulazione numerica calibrata per ricavare dei coefficienti che mettono in relazione le variazioni fisiche e geometriche della cella con le variazioni dei parametri caratteristici, mediante i quali si è poi effettuata un'analisi statistica. Si è mostrato che in questo modo è possibile riprodurre le correlazioni sperimentali, dimostrando che esse sono dovute a interazioni tra fenomeni elettrici e termici non tenute in conto nelle formule analitiche. Si è anche verificato che la dispersione statistica del GST è notevolmente più alta di quella dell'heater, e pertanto la sua ottimizzazione è cruciale per le prestazioni della memoria.

Il modello calibrato è stato anche utilizzato per stimare le prestazioni di due dispositivi non ancora disponibili per un'analisi sperimentale: una cella scalata al nodo 32nm e una cella *damascata*, costituita da uno strato di GST a base cilindrica interamente racchiuso da isolanti e che non prevede l'utilizzo dell'heater. Nel caso della cella scalata è stato mostrato che lo scaling delle memorie PCM segue qualitativamente l'andamento previsto dai modelli analitici, ma per una descrizione quantitativamente precisa l'utilizzo del modello numerico è fondamentale. Si è verificato che è possibile, mediante la riduzione delle dimensioni geometriche, continuare a ridurre le correnti di programmazione richieste senza sacrificare l'ampiezza della finestra di lettura. Per quanto riguarda la cella damascata, si è osservato che il confinamento termico gioca un ruolo cruciale nel suo funzionamento: grazie ad esso la curva di programmazione di questa cella è estremamente ripida, a scapito

di un'elevata sensibilità della corrente di melting alle variazioni di altezza dello strato di GST.

È stato anche avviato lo studio del fenomeno di ricristallizzazione del GST durante il quenching in un'operazione di RESET. I dati sperimentali mostrano che non tutto il GST che viene fuso amorfizza: le parti più lontane dall'interfaccia risentono del meccanismo di crescita cristallina dovuto al reticolo che le circonda, e non raggiungono la fase amorfa alla fine del transitorio. Grazie all'uso del modello numerico è stata eseguita una simulazione in transitorio per cercare di caratterizzare questo effetto. È stato mostrato che è possibile riprodurre qualitativamente la ricristallizzazione mediante un'opportuna calibrazione della velocità di cristallizzazione. Non è stato possibile riprodurre quantitativamente le misure sperimentali, per cui questo aspetto viene lasciato come possibile futuro sviluppo del modello a partire dal presente lavoro.

Bibliografia

- [1] A. Redaelli *et al.*, “Threshold switching and phase transition numerical models for phase change memory simulations”, *Journal of Applied Physics* 103, 111101 (2008)
- [2] S. R. Ovshinsky, *Physical Review Letters* 21, 1450 (1968)
- [3] A. Pirovano *et al.*, “Electronic Switching in Phase-Change Memories”, *IEEE Transactions on Electron Devices* Vol.51, n.3 (2004)
- [4] D.Adler, “Defects in amorphous semiconductors”, *Journal of Non-Crystalline Solids*, vol. 35-36, pp. 819-824 (1980)
- [5] D.Ielmini, Y.Zhang, “Analytical model for subthreshold conduction and threshold switching in chalcogenide-based memory devices”, *Journal of Applied Physics* 102, 054517 (2007)
- [6] F. Pellizzer *et al.*, *Proceeding of VLSI Technology Symposium* (2006)
- [7] G. Servalli, *IEDM Technical Digest*, p. 113 (2009)
- [8] U. Russo *et al.*, “Modeling of Programming and Read Performance in Phase Change Memories - Part I: Cell Optimization and Scaling”, *IEEE Transactions on Electron Devices* Vol. 55, n.2 (2008)
- [9] J. A. Kalb, “Stresses, viscous flows and crystallization kinetics in thin films of amorphous chalcogenides used for optical data storage”, *Rheinisch-Westphalischen Technischen Hochschule Aachen* (2002)

- [10] Ho-Ki Lyeo *et al.*, “Thermal Conductivity of Phase Change Material $\text{Ge}_2\text{Sb}_2\text{Te}_5$ ”, Applied Physics Letters 89, 151904 (2006)
- [11] Giraud *et al.*, “Thermal Characterization and Analysis of Phase Change Random Access Memory”, Journal of Applied Physics 98, 013520 (2005)
- [12] E.K. Kim *et al.*, Applied Physics Letters 76, 3864 (2000)
- [13] C.Peng *et al.*, Journal of Applied Physics 82, 4183 (1997)
- [14] J.-L. Battaglia *et al.*, “Thermal Characterization of the $\text{SiO}_2\text{-Ge}_2\text{Sb}_2\text{Te}_5$ interface from room temperature up to 400°C ”, Journal of Applied Physics 107, 044314 (2010)
- [15] D. Kencke *et al.*, “The Role of Interfaces in Damascene Phase-Change Memory” IEEE Technical Digest (2007)
- [16] S. Savransky, I. Karpov, “Investigation of SET and RESET states resistance in Ohmic regime for Phase-Change Memory”, Materials Research Society Symposium Procedures Vol. 1072 (2008)
- [17] A. Pirovano *et al.*, “Self-Aligned μ trench Phase-Change Memory Cell Architecture for 90nm Technology and Beyond”, VLSI Technology Symposium (2007)
- [18] I. S. Kim *et al.*, “High Performance PRAM Cell Scalable to sub-20nm Technology with below $4F^2$ Cell Size, Extendable to DRAM Applications”, VLSI Tech. (2010)
- [19] I. V. Karpov, “Damascene phase change memory”, U.S. Patent #7211819 (2003)
- [20] T. Kato, K. Tanaka, “Electronic properties of amorphous and crystalline $\text{Ge}_2\text{Sb}_2\text{Te}_5$ Films”, Japanese Journal of Applied Physics Vol. 44, n.10 (2005)
- [21] J. H. Lienhard, “A Heat Transfer Textbook”, Phlogiston Press (2008)
- [22] Bong-Sub Lee *et al.*, “Investigation of the optical and electronic properties of $\text{Ge}_2\text{Sb}_2\text{Te}_5$ phase change material in its amorphous, cubic and hexagonal phases”, Journal of Applied Physics 97, 093509 (2005)