

Progettazione e Analisi di una Cella SRAM a 6T

Raffaele Petrolo

1 Obiettivo dell'Analisi

Il presente elaborato riguarda la progettazione e la caratterizzazione di una cella SRAM a sei transistor (6T) in tecnologia CMOS 100 nm . L'analisi si focalizza sul compromesso tra robustezza statica (margini di rumore) e dissipazione di potenza, valutando l'impatto della scalabilità della tensione di alimentazione fino al limite del *Data Retention Voltage* (DRV).

2 Design della Cella e Sizing degli Inverter

2.1 Ricerca del punto di ottimo

Inizialmente si è cercato di ottenere inverter interni quanto più simmetrici possibile. Fissata la larghezza dei Pull-down al valore minimo $W_n = 0.12\text{ }\mu\text{m}$, è stata eseguita una simulazione parametrica variando la larghezza W_p dei transistori di Pull-up.

L'obiettivo era centrare la soglia logica della transcaratteristica (VTC) a $V_{DD}/2 = 0.5\text{ V}$.

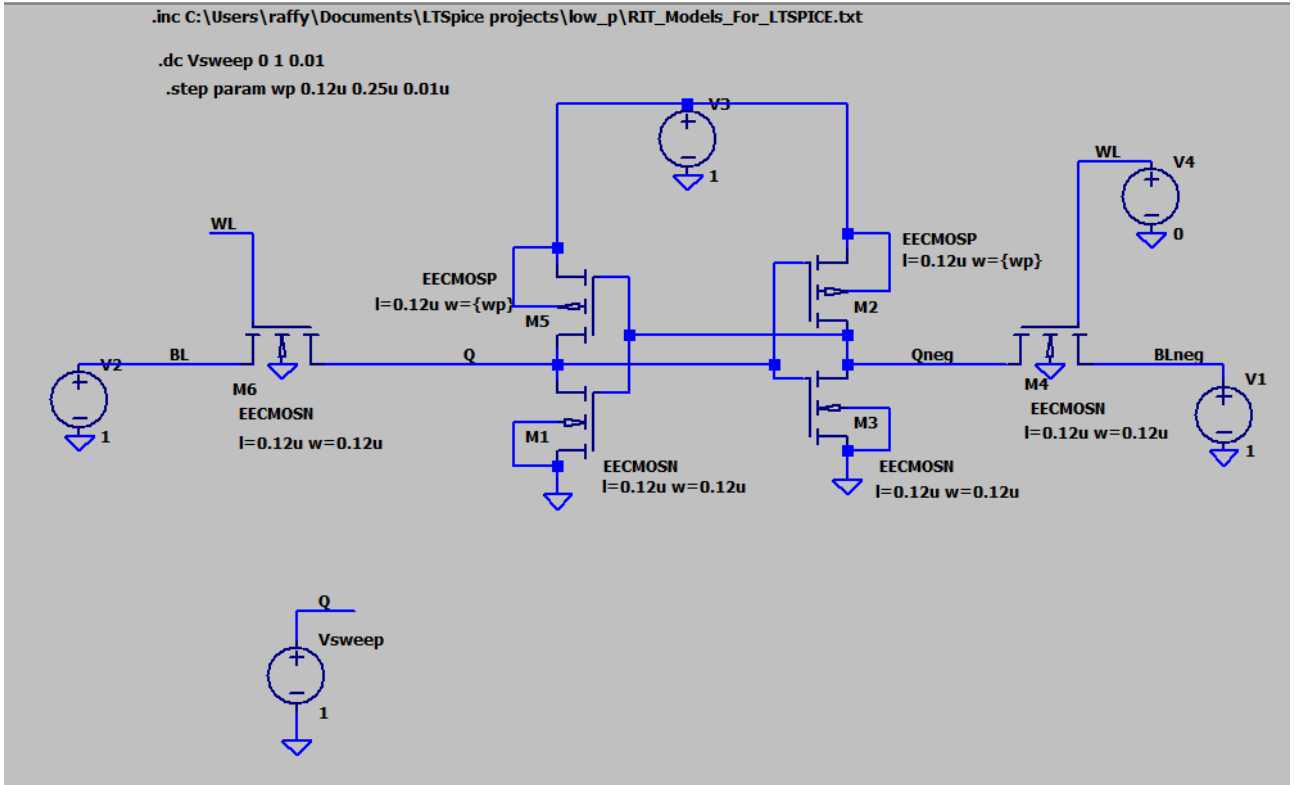


Figura 1: Schematico della cella 6T.

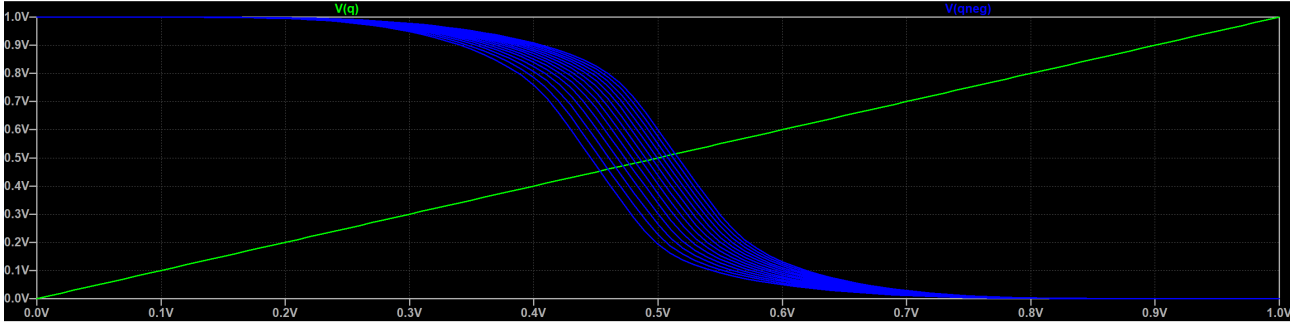


Figura 2: Simulazione parametrica di W_p .

Dall'analisi della Figura 2, il punto di commutazione ottimale è stato individuato per:

- $W_p = 0.21 \mu m$
- $W_n = 0.12 \mu m$

3 Analisi della Stabilità Statica (SNM)

La stabilità della cella è stata valutata tramite lo *Static Noise Margin* (SNM), definito come il lato del massimo quadrato inscrivibile nei lobi delle *Butterfly Curves*.

3.1 Metodologia di calcolo (Script MATLAB)

Per l'estrazione automatica dei margini di rumore, è stato sviluppato uno script MATLAB basato sul metodo grafico della diagonale massima. Il cuore dell'algoritmo risiede nell'identificazione del massimo quadrato inscrivibile nei lobi della *Butterfly Curve*. Il procedimento logico-matematico è il seguente:

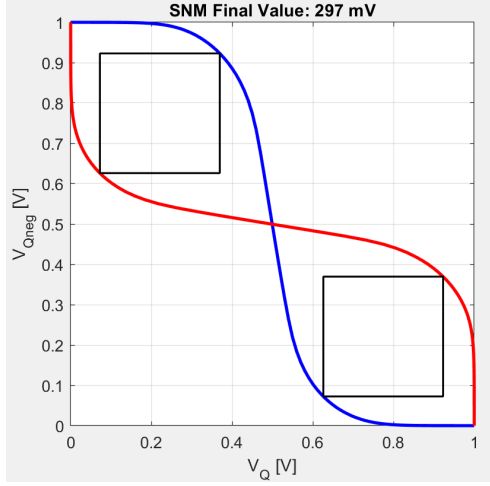
1. **Generazione delle curve:** Partendo dai dati di simulazione (V_{in}, V_{out}) , si definisce la prima curva VTC_1 come $y = f(x)$. La seconda curva VTC_2 , sfruttando la simmetria della cella, è definita come l'inversa matematica $x = f(y)$.
2. **Scansione geometrica:** Si considera il fascio di rette parallele alla bisettrice del primo e terzo quadrante, aventi equazione $y = x + q$. Per ogni valore dell'intercetta q , tale retta interseca le due curve in due punti distinti, P_1 e P_2 .
3. **Ricerca delle intersezioni:** Lo script calcola per ogni punto della VTC_1 il valore $q = y - x$. Successivamente, per un valore target di q , identifica:
 - Il punto $P_1(x_1, y_1)$ sulla VTC_1 tale per cui $y_1 - x_1 = q$.
 - Il punto $P_2(x_2, y_2)$ sulla VTC_2 tale per cui $y_2 - x_2 = -q$ (sfruttando la simmetria speculare delle curve rispetto alla bisettrice principale $y = x$).
4. **Ottimizzazione della diagonale:** La distanza euclidea d tra P_1 e P_2 rappresenta la diagonale di un quadrato con i lati paralleli agli assi coordinati. Lo script ricerca iterativamente il valore massimo di tale distanza (d_{max}) all'interno dei due lobi separatamente.
5. **Calcolo dell'SNM:** Poiché in un quadrato di lato L la diagonale è $d = L\sqrt{2}$, il margine di rumore statico (*Static Noise Margin*) viene estratto come:

$$SNM = \frac{d_{max}}{\sqrt{2}} \quad (1)$$

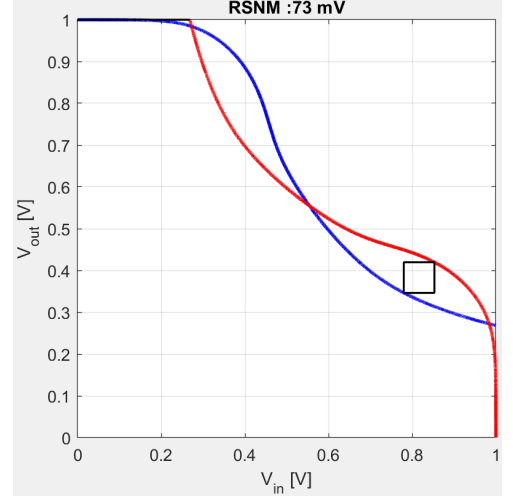
3.2 Analisi con dimensionamento iniziale

Inizialmente, la cella è stata testata con il dimensionamento simmetrico ottenuto nel capitolo precedente ($W_p = 0.21 \mu m$, $W_n = 0.12 \mu m$) e transistor di accesso minimi ($W_{ax} = 0.12 \mu m$).

Le Figure 3a e 3b mostrano i quadrati di rumore ottenuti dallo script per le condizioni di Hold e Read.



(a) HSNM (Hold).



(b) RSNM (Read).

Figura 3: Butterfly curves e quadrati inscritti con dimensionamento iniziale.

Sebbene l'HSNM risulti adeguato ($\approx 297 mV$), l'RSNM è troppo basso, circa $73 mV$. Questa fragilità è dovuta al partitore di tensione formato tra il transistor di accesso e il Pull-down durante la lettura.

3.3 Ottimizzazione del Beta Ratio

Per migliorare il margine di Read, si è inizialmente valutato l'impatto di W_p . Tuttavia, i risultati mostrano che per ottenere un guadagno significativo nel margine di Read sarebbe necessario quadruplicare la larghezza del Pull-up, con un beneficio complessivo molto ridotto (Figura 4).

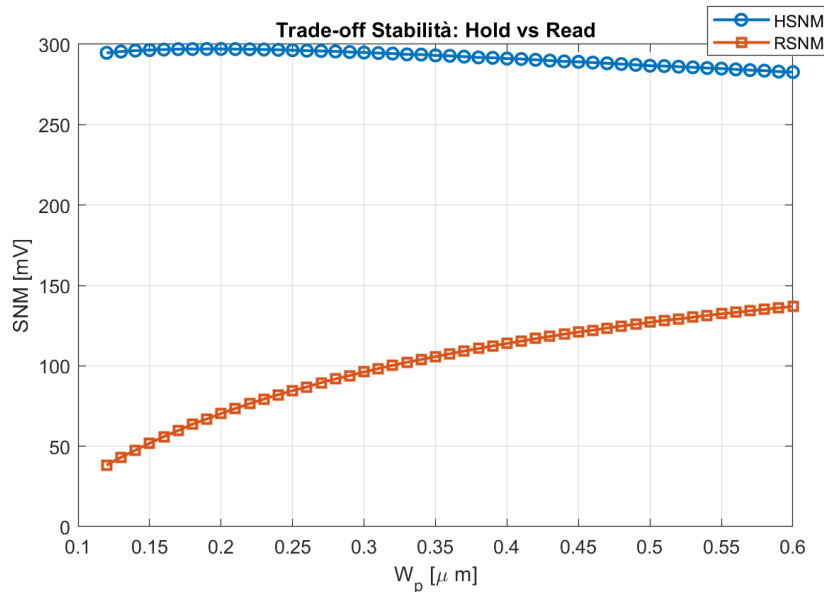


Figura 4: Andamento di HSNM e RSNM al variare di W_p .

Si è quindi deciso di agire sul **Beta Ratio** (β), ovvero il rapporto di forza tra Pull-down e transistor di accesso, definito come:

$$\beta = \frac{W_n}{W_{ax}} \quad (2)$$

Mantenendo W_{ax} al valore minimo, è stata mappata la risposta dei margini al variare di W_n per tre diverse configurazioni di Pull-up.

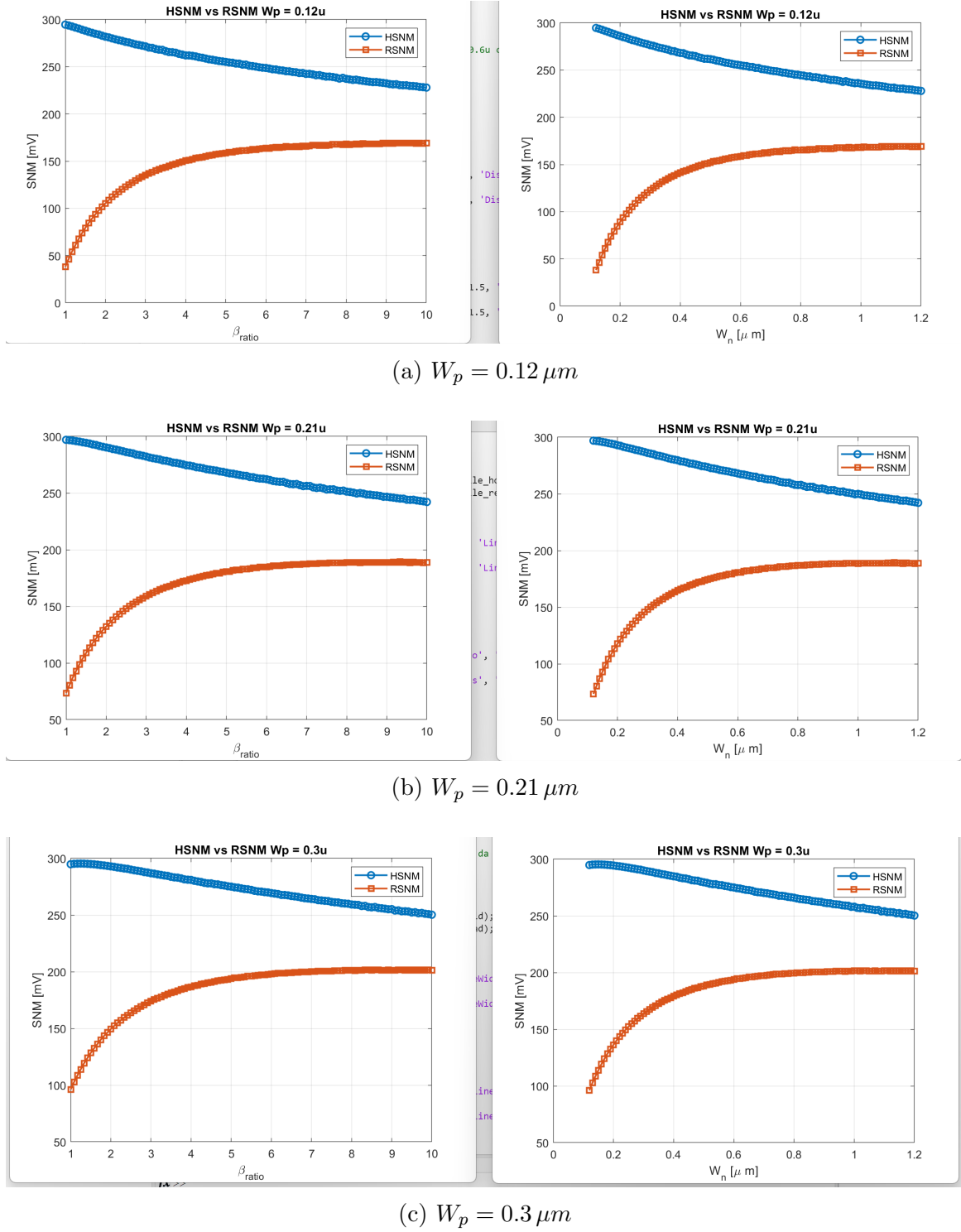


Figura 5: Variazione dei margini di rumore in funzione di W_n per diversi valori di W_p .

3.4 Scelta del dimensionamento ottimale

Dall'analisi comparativa è emerso che l'incremento di W_p produce un offset positivo sull'RSNM estremamente ridotto a fronte di un aumento considerevole dell'area occupata. Si è quindi scelto di mantenere il Pull-up al valore minimo tecnologico, $W_p = 0.12 \mu m$, per minimizzare l'ingombro della cella.

Per quanto riguarda il Pull-down, l'andamento dell'RSNM in funzione di W_n mostra un comportamento di tipo sub-lineare, riconducibile a una funzione radice quadrata. Per individuare il punto di massima efficacia del dimensionamento, è stata calcolata la derivata prima dell'RSNM rispetto a W_n (Figura 6).

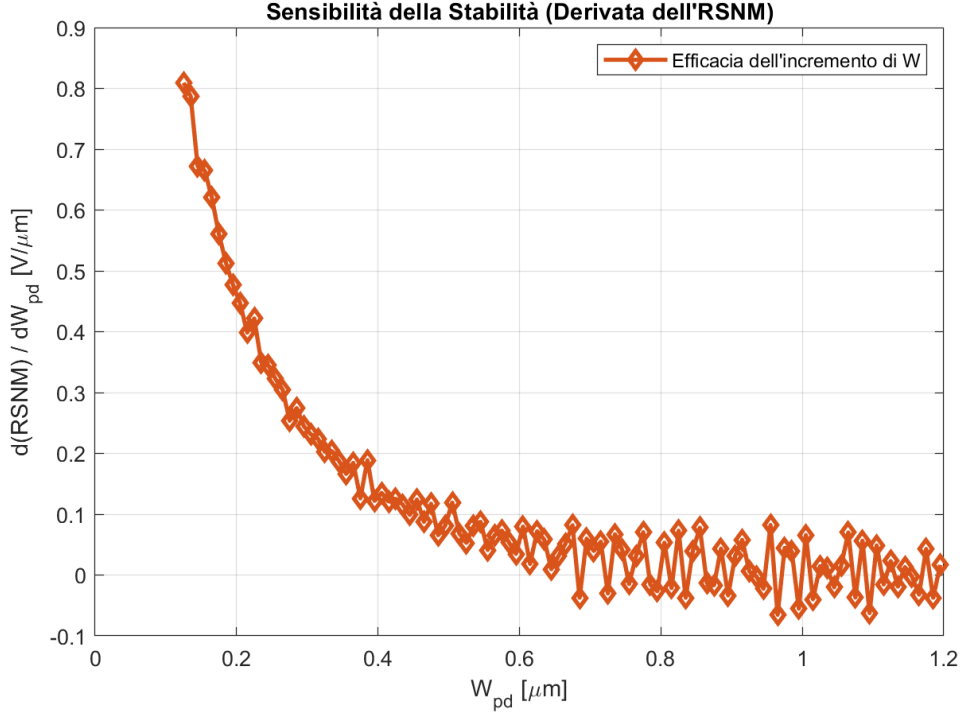


Figura 6: Derivata dell'RSNM rispetto a W_n . L'andamento tipico $1/\sqrt{x}$ conferma che i maggiori benefici in termini di stabilità si ottengono per piccoli incrementi di W_n nel range iniziale.

Il grafico della derivata mostra chiaramente che la sensibilità del margine di rumore cala drasticamente all'aumentare della larghezza del canale. Si è quindi individuato un punto di lavoro a $W_n = 0.375 \mu m$, che rappresenta un buon compromesso tra robustezza statica e densità di integrazione.

3.5 Impatto dei transistor di accesso e validazione del modello

Completata l'analisi sui transistor interni, si è valutato l'impatto della larghezza dei transistor di accesso (W_{ax}). Come previsto dalla teoria del Beta Ratio, un incremento di W_{ax} comporta inizialmente un drastico peggioramento dell'RSNM (Figura 7).

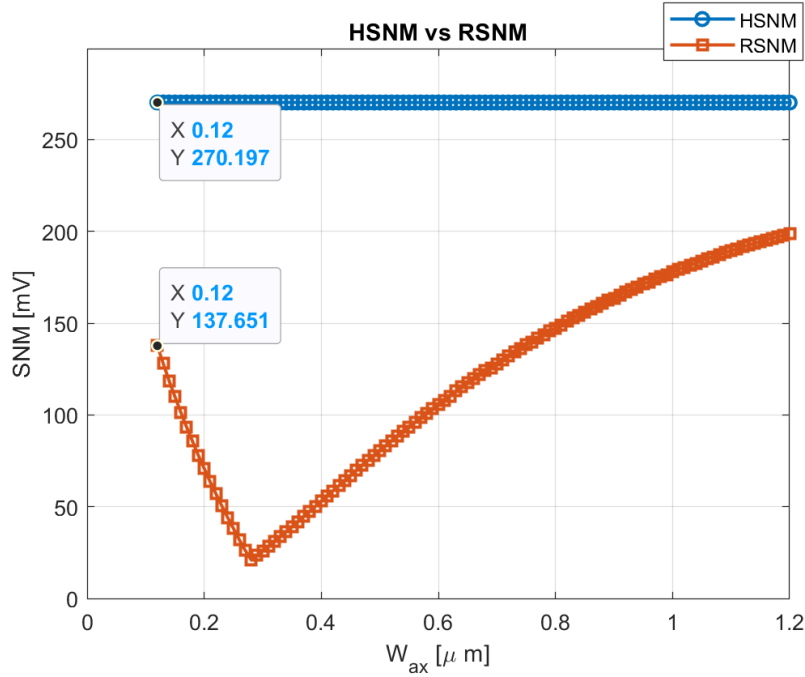


Figura 7: Degradazione dell'RSNM all'aumentare della larghezza dei transistor di accesso W_{ax} .

Osservando la Figura 7, si nota un comportamento singolare: superato il punto critico in cui $W_{ax} \approx W_n$, il margine calcolato dallo script MATLAB sembra apparentemente migliorare. Tuttavia, tale incremento è un *falso positivo* dovuto alla natura puramente geometrica dell'algoritmo.

Quando l'accesso diventa eccessivamente forte rispetto al Pull-down, la cella perde la sua caratteristica bistabile e le curve a farfalla si deformano al punto da non incrociarsi più correttamente per formare i due lobi. Lo script continua a ricercare la massima distanza euclidea tra le curve, identificando un valore numerico elevato che però non rappresenta più una condizione di stabilità, bensì la transizione verso un regime monostabile dove il dato viene sovrascritto dalla Bitline (*read-destructive event*). Per tale ragione, si è confermato il valore minimo tecnologico $W_{ax} = 0.12 \mu m$.

Il design finale della cella per ora presenta quindi i seguenti parametri nominali: $W_p = 0.12 \mu m$, $W_n = 0.375 \mu m$ e $W_{ax} = 0.12 \mu m$. Con tale configurazione, i margini calcolati tramite lo script MATLAB risultano:

- **HSNM:** 270.2 mV
- **RSNM:** 137.65 mV

Per validare l'accuratezza dell'algoritmo MATLAB, è stata eseguita una verifica comparativa utilizzando il metodo di Seevnick in LTspice. Le Figure 8 e 9 mostrano la perfetta coincidenza tra i valori estratti graficamente e quelli simulati, confermando la solidità della procedura.

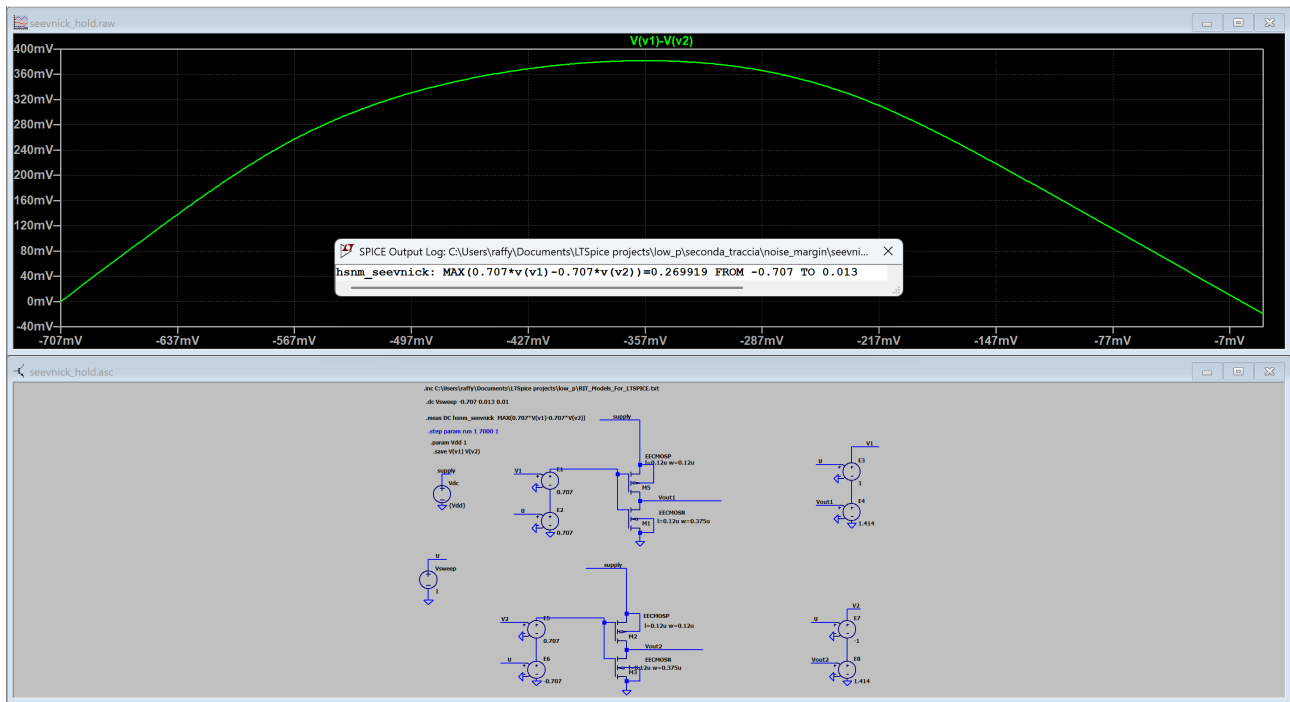


Figura 8: Validazione HSNM tramite metodo di Seevnick.

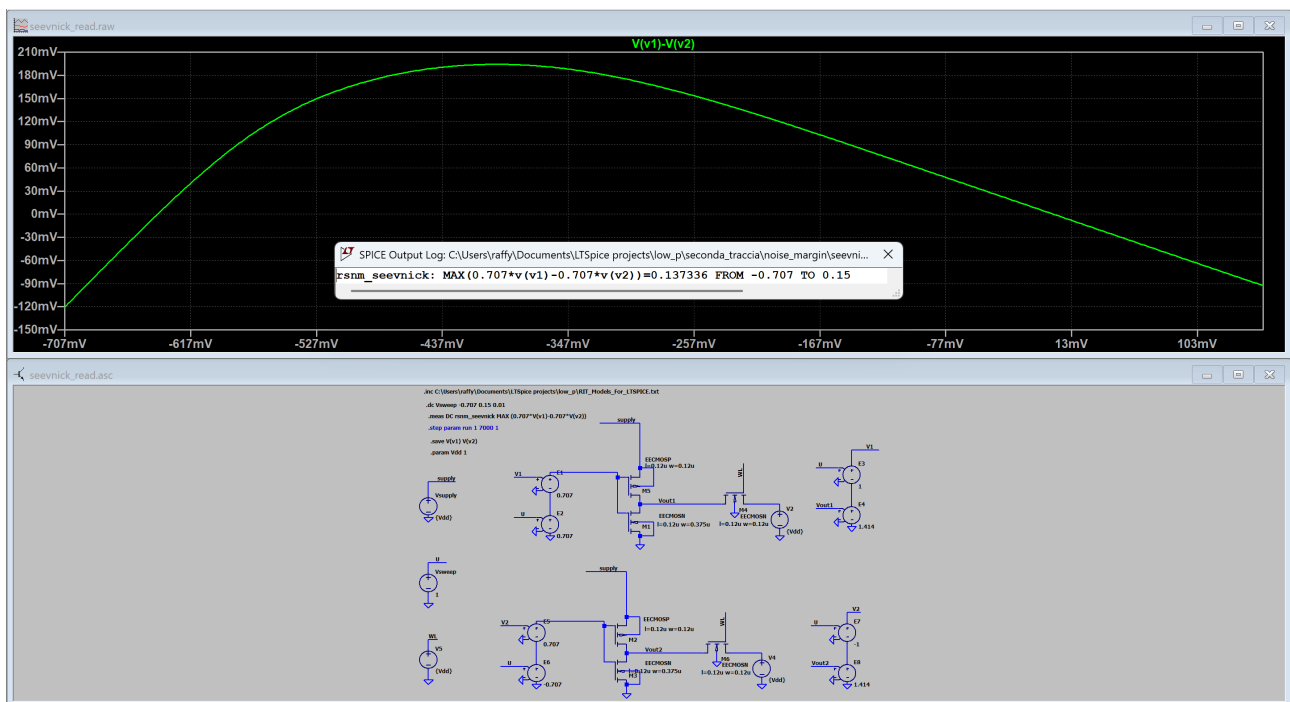


Figura 9: Validazione RSNM tramite metodo di Seevnick.

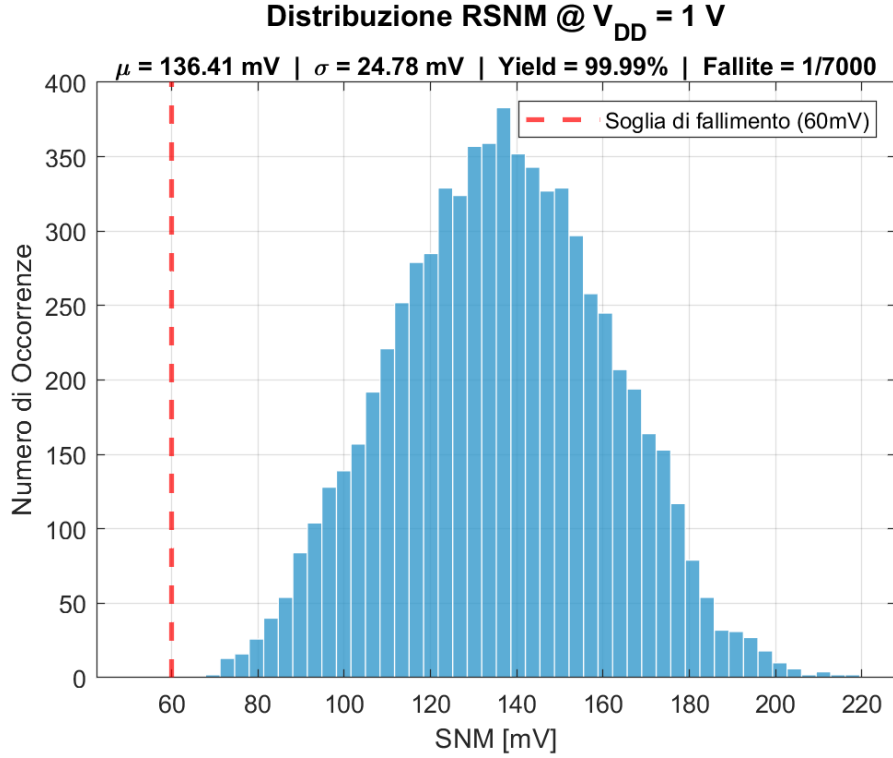
4 Analisi statistica e Scaling della Tensione di Alimentazione

In questo capitolo viene analizzata la robustezza della cella SRAM considerando uno scaling della tensione V_{DD} da 1 V a 0.2 V (con passi di 0.1 V). Per ciascun valore sono state eseguite 7000 simulazioni Monte Carlo, applicando una variazione gaussiana alle tensioni di soglia (V_{th}) dei transistor per modellare le variazioni di processo inter-die.

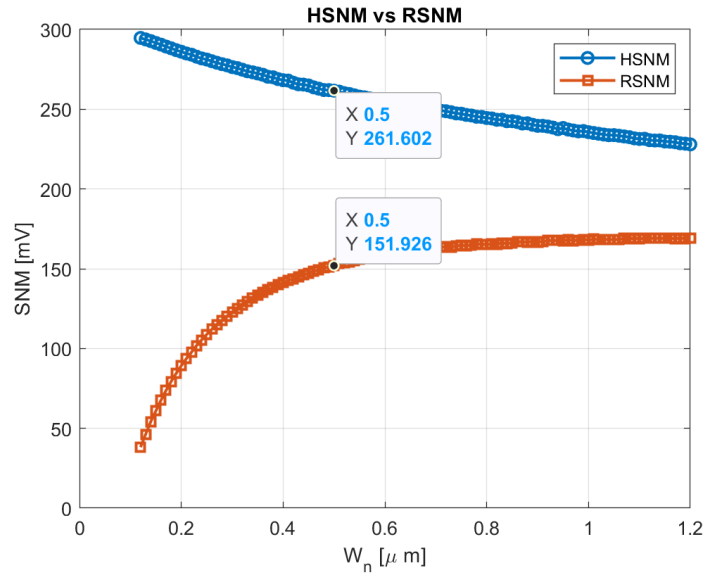
L'obiettivo è identificare il limite di funzionamento della cella, fissando una soglia critica di sicurezza $SNM_{min} = 60\text{ mV}$.

4.1 Ottimizzazione finale del dimensionamento

L'analisi statistica condotta sul dimensionamento precedentemente individuato ($W_n = 0.375\text{ }\mu\text{m}$) ha evidenziato una criticità: già alla tensione nominale di 1 V , si è riscontrato che una cella (quindi 1 su 7000) presentava un RSNM inferiore alla soglia di sicurezza di 60 mV .



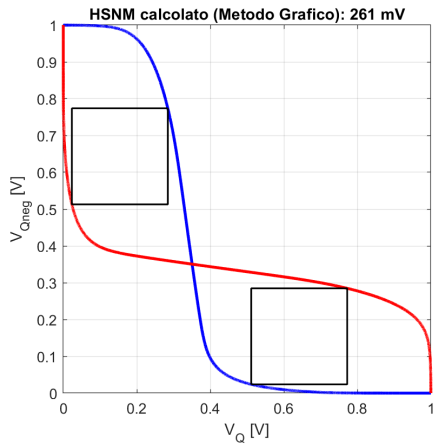
Per garantire quindi un *yield* elevato, si è deciso di incrementare ulteriormente il Beta Ratio, spostandosi sulla curva di sensibilità precedentemente calcolata verso un valore di $W_n = 0.5\text{ }\mu\text{m}$.



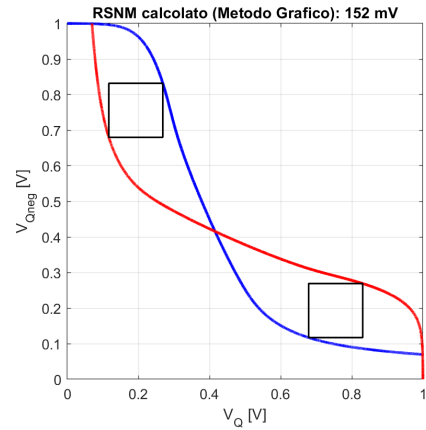
Con questo nuovo dimensionamento, i margini nominali calcolati a 1 V risultano appunto come in figura:

- **HSNM:** 261 mV
- **RSNM:** 152 mV

Anche in questo caso, la coerenza tra il metodo grafico MATLAB (Figure 10a e 10b) e il metodo di Seevinck (Figure 11a e 11b) è totale.

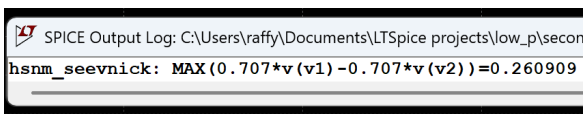


(a) HSNM (Metodo Grafico).

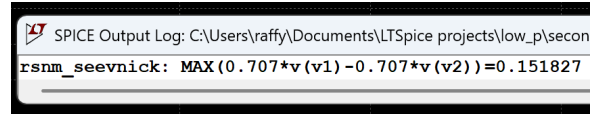


(b) RSNM (Metodo Grafico).

Figura 10: Verifica dei nuovi margini con $W_n = 0.5 \mu m$ tramite script MATLAB.



(a) HSNM (Seevinck).



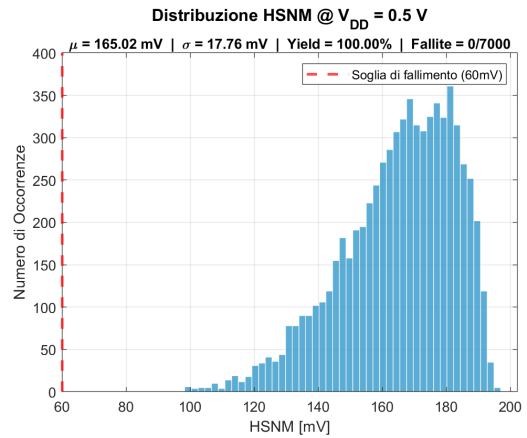
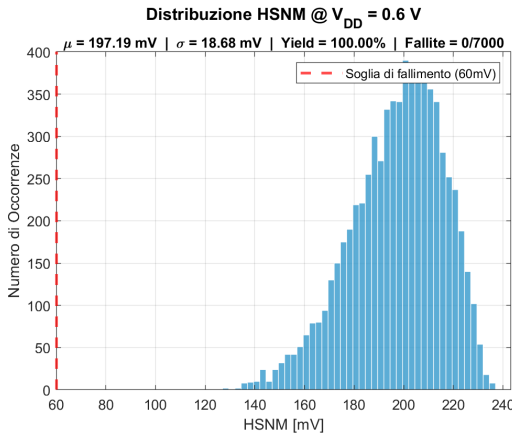
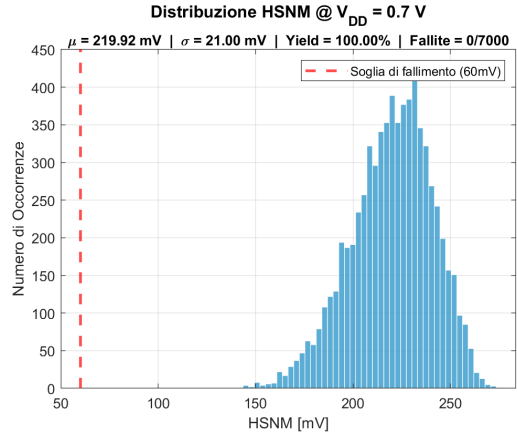
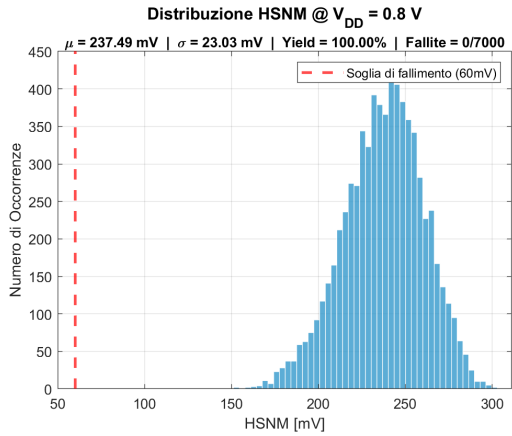
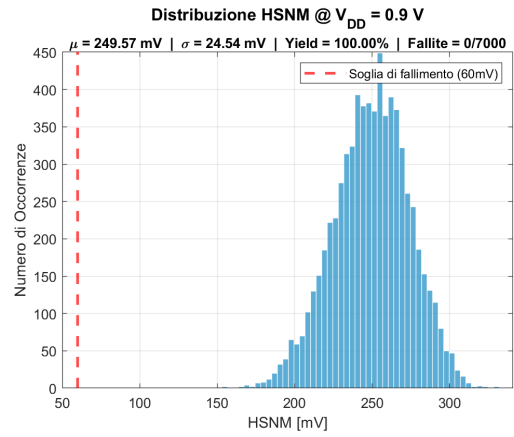
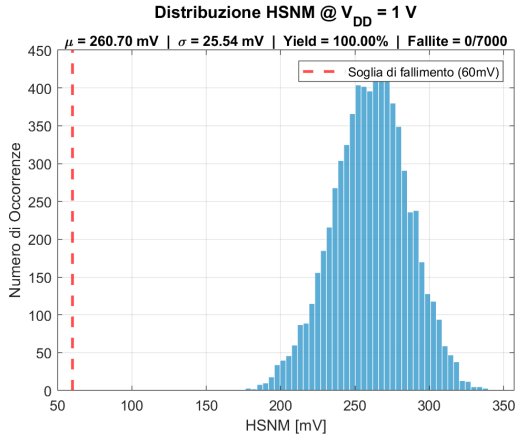
(b) RSNM (Seevinck).

Figura 11: Validazione dei margini tramite simulazione Seevinck.

4.2 Evoluzione statistica dei margini

Di seguito vengono riportate le distribuzioni statistiche ottenute dalle simulazioni Monte Carlo per l'intero range di scaling. Si osserva un fenomeno fisico rilevante: mentre a tensioni elevate le distribuzioni seguono un andamento prettamente gaussiano, al ridursi della V_{DD} la forma tende verso una distribuzione log-normale.

4.2.1 Istogrammi HSNM ($1\text{ V} \rightarrow 0.2\text{ V}$)



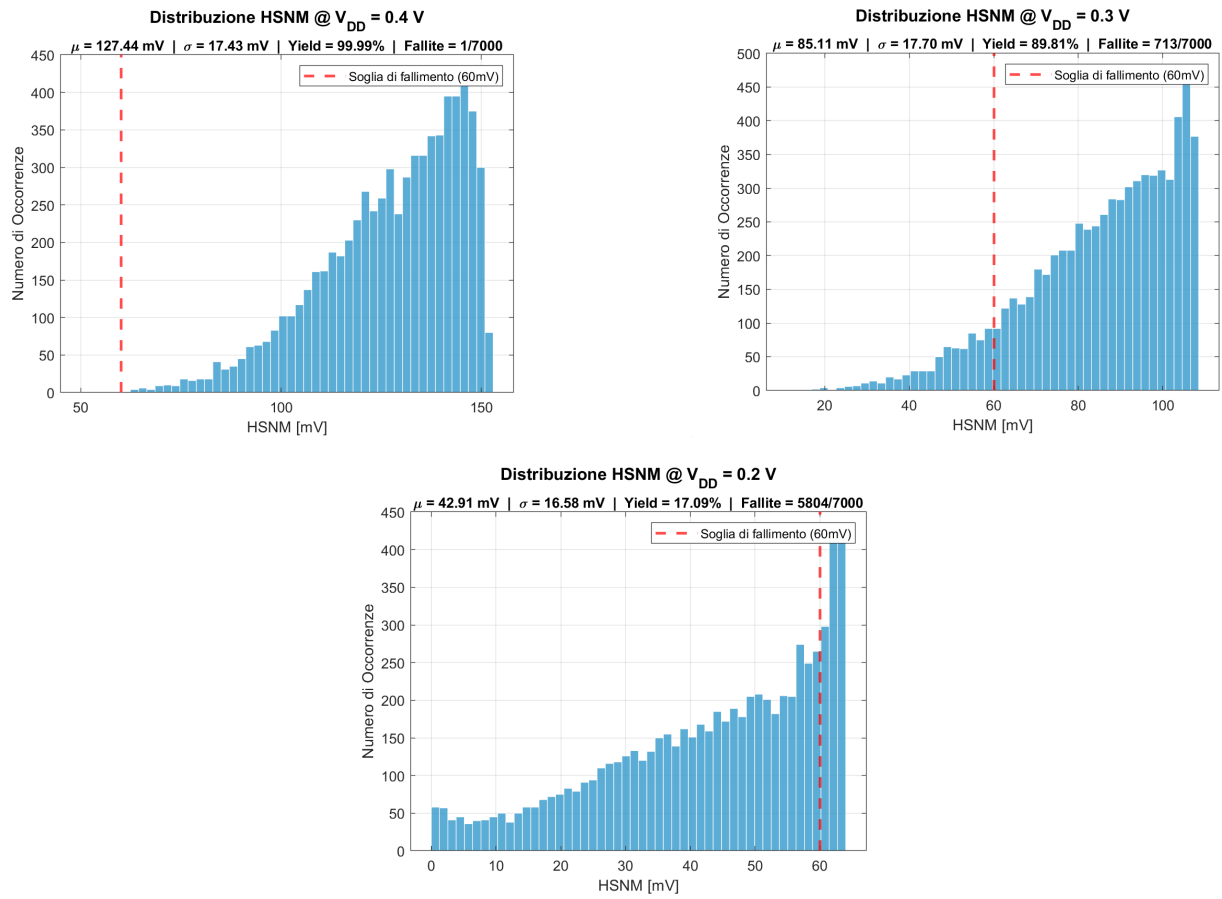
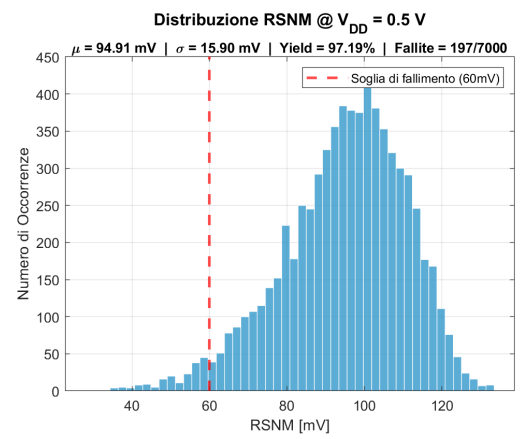
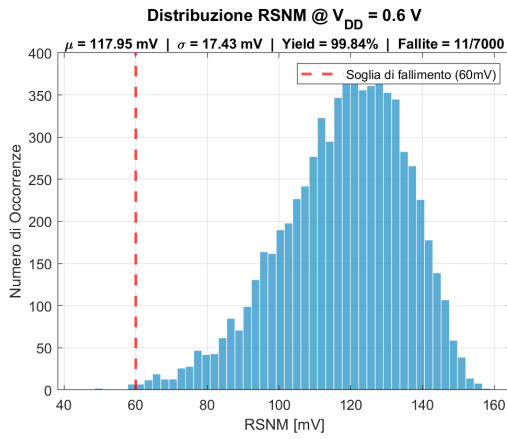
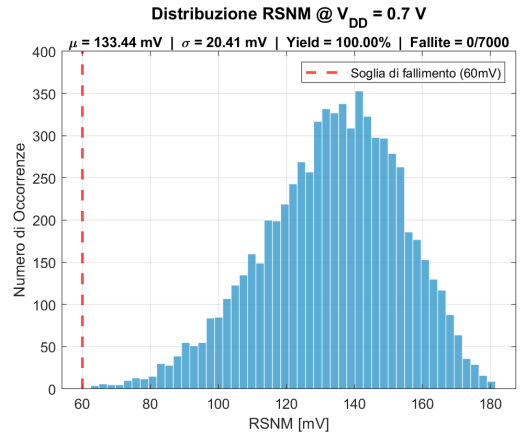
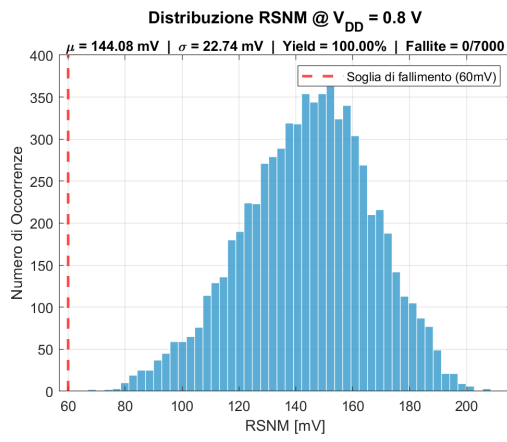
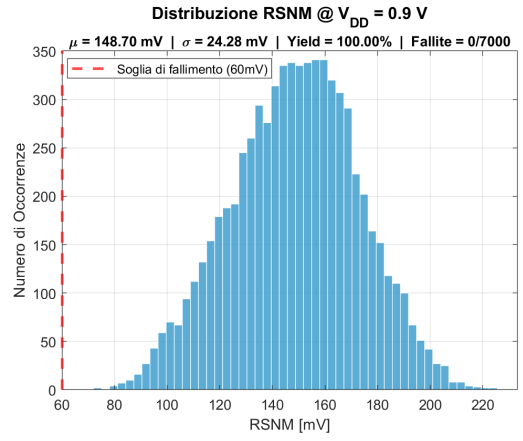
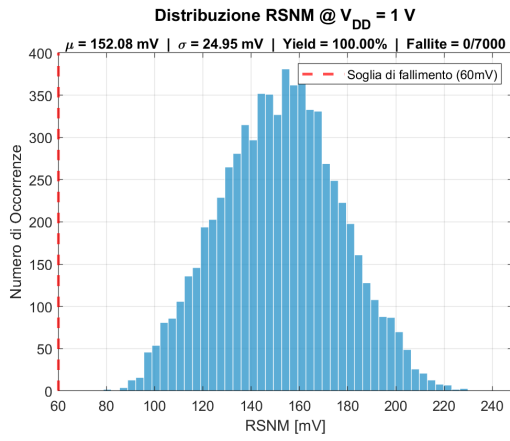


Figura 13: Distribuzioni Monte Carlo HSNM (1V - 0.2V).

4.2.2 Istogrammi RSNM ($1V \rightarrow 0.2V$)



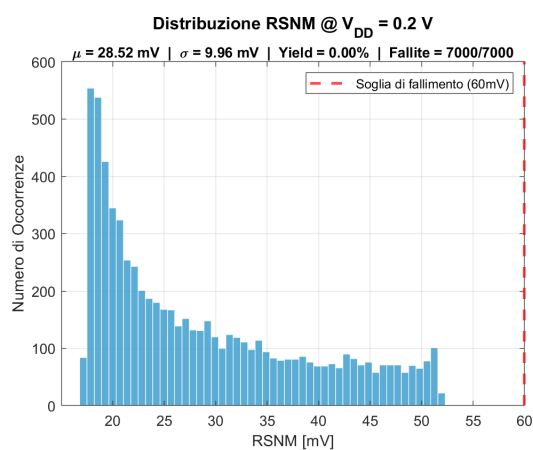
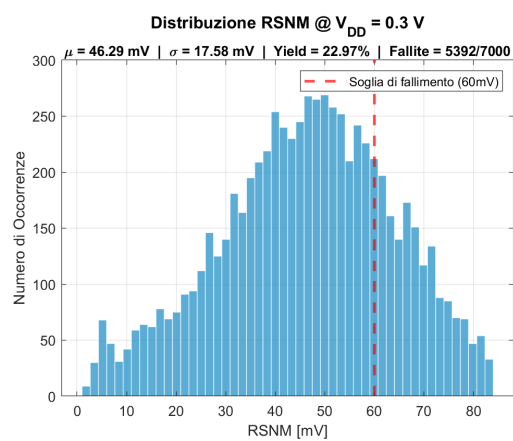
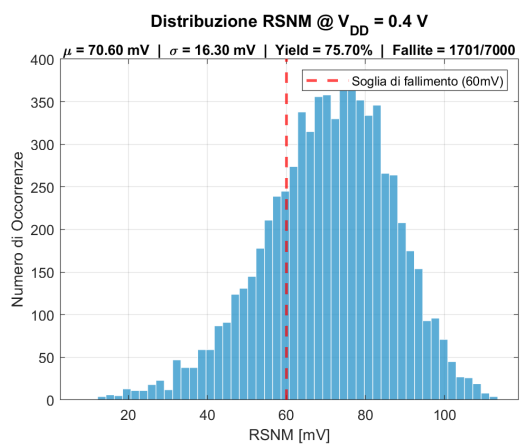


Figura 15: Distribuzioni Monte Carlo RSNM (1V - 0.2V).

4.3 Analisi delle correnti e della potenza di leakage

Per analizzare il consumo statico della cella, è stata condotta una simulazione in transitorio articolata in due fasi. Inizialmente è stato forzato il caricamento di un dato logico all'interno della cella; successivamente, il circuito è stato portato in fase di *hold* (Wordline a 0 V e Bitline a V_{DD}). La lettura delle correnti di leakage è stata effettuata al termine del transitorio, garantendo che i nodi fossero in regime stazionario.

La corrente di leakage totale della cella è stata calcolata come somma dei contributi dei singoli transistor interdetti, come mostrato nello schematico di Figura 16:

- I_{PU} : corrente sul PMOS di Pull-up spento (M_2).
- I_{PD} : corrente sull'NMOS di Pull-down spento (M_1).
- I_{AX} : corrente sull'NMOS di accesso spento (M_4).

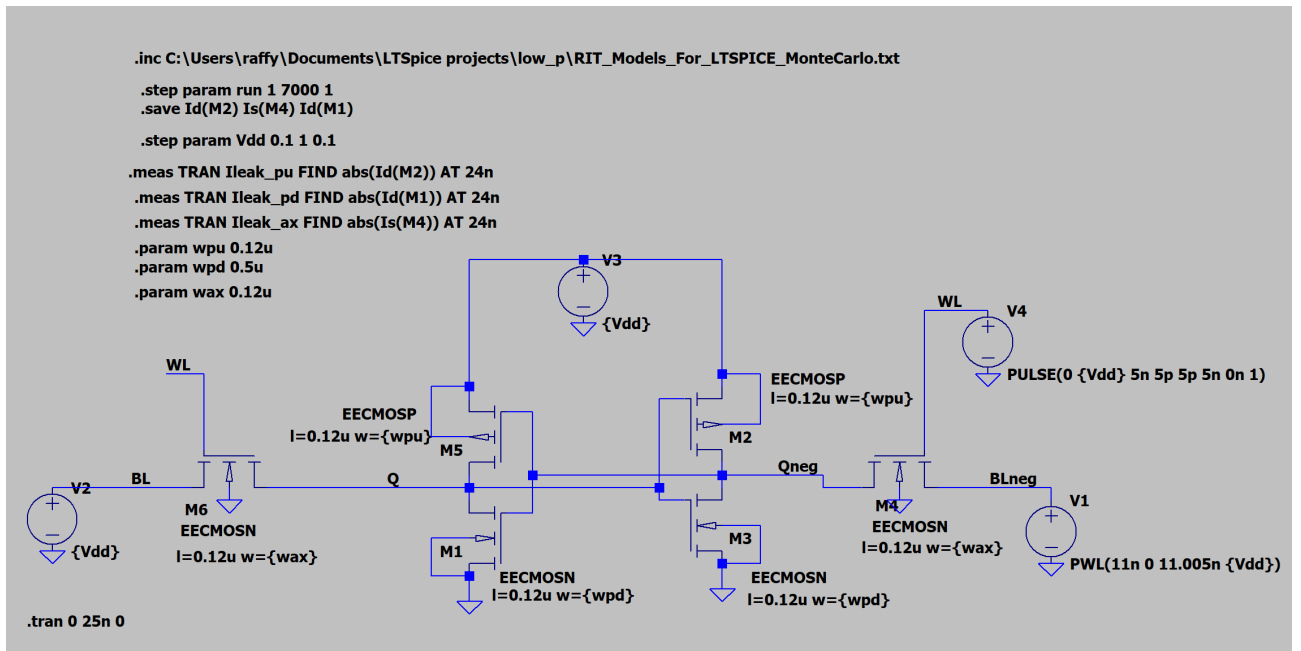
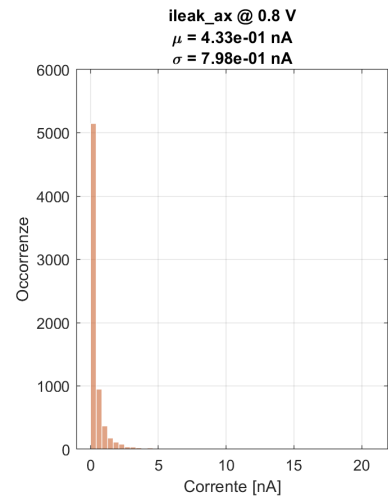
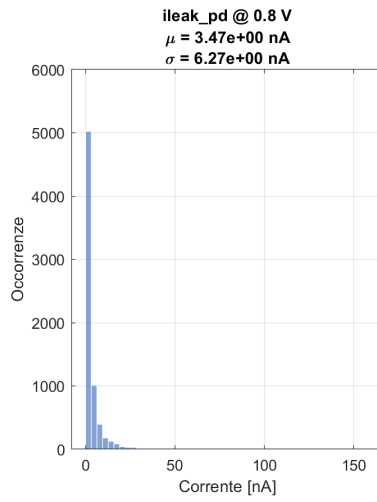
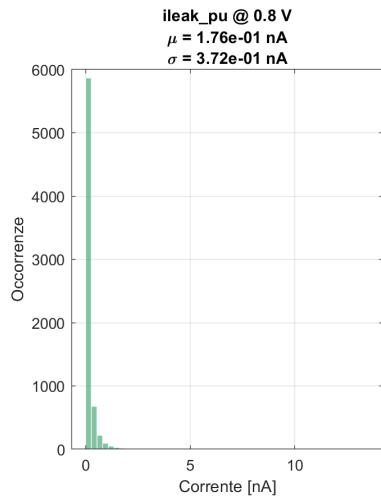
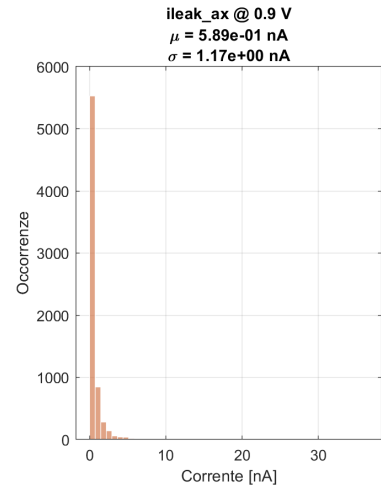
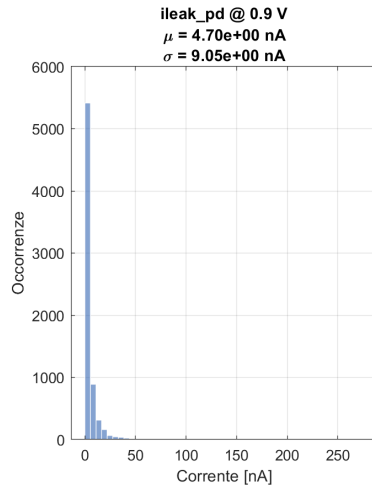
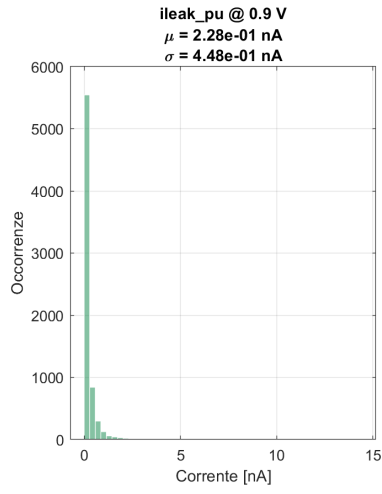
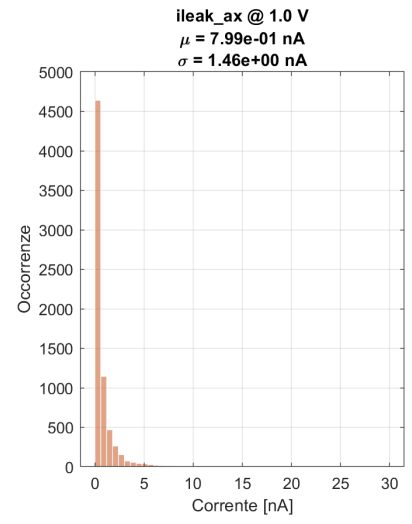
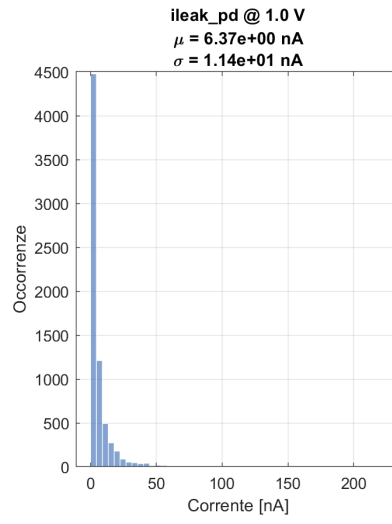
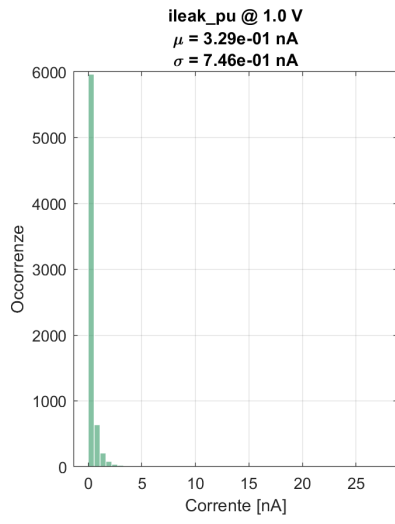
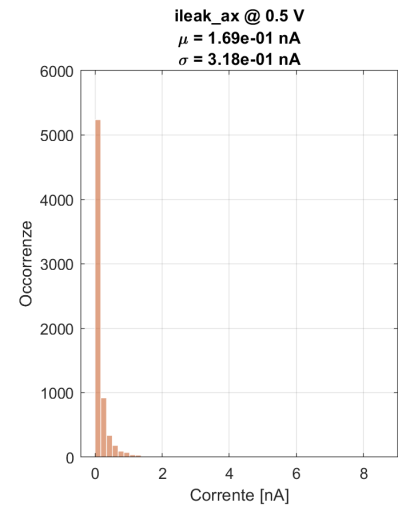
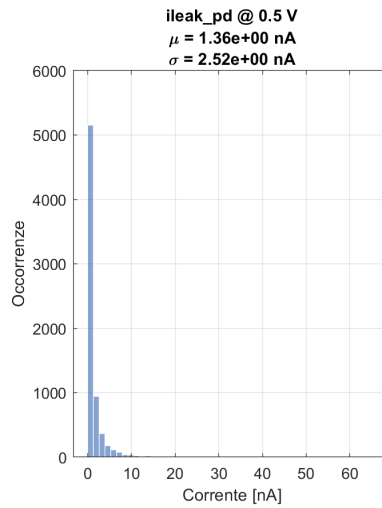
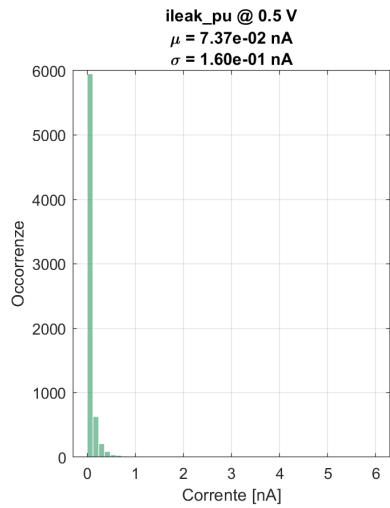
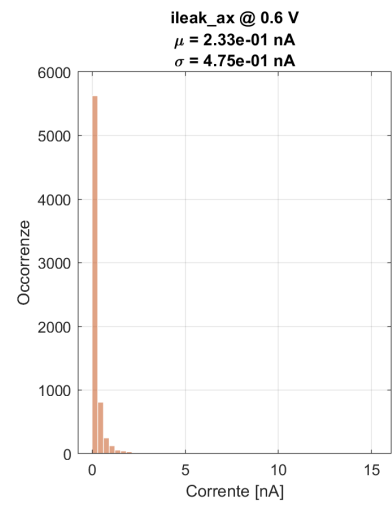
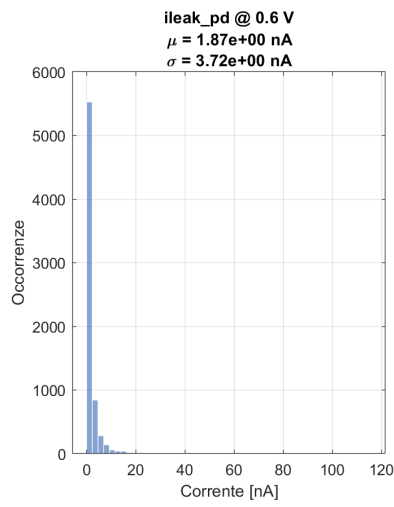
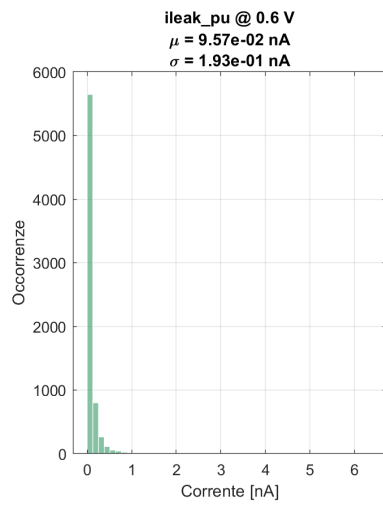
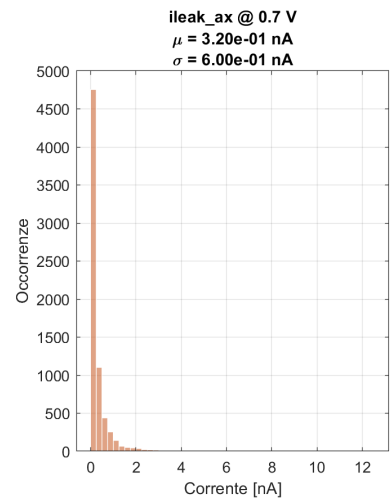
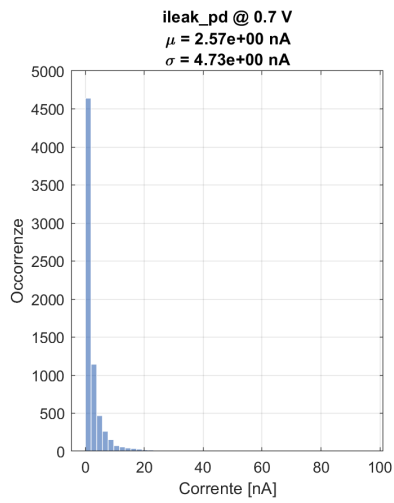
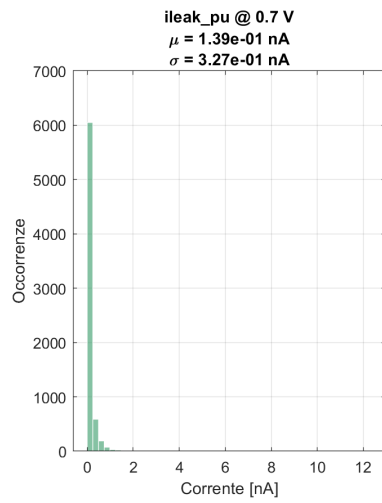


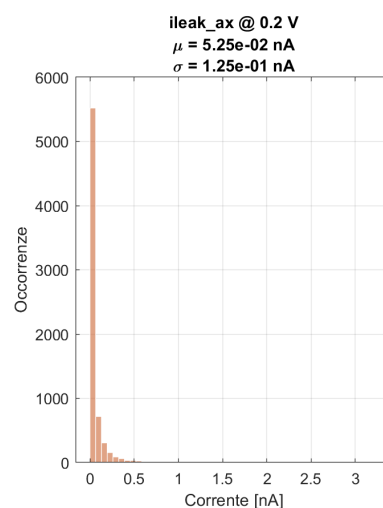
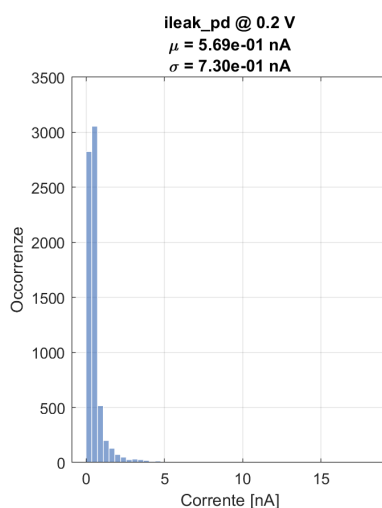
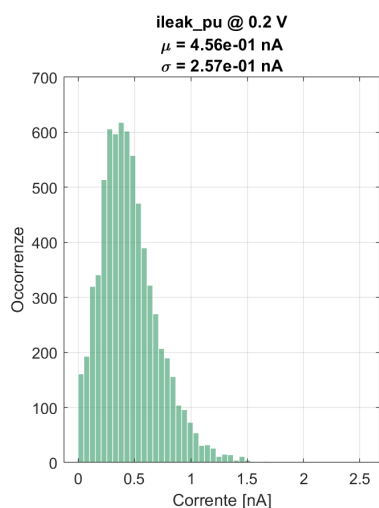
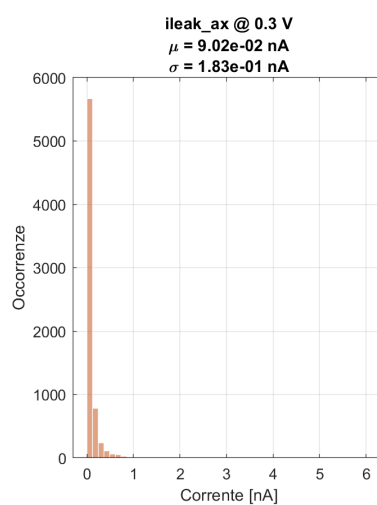
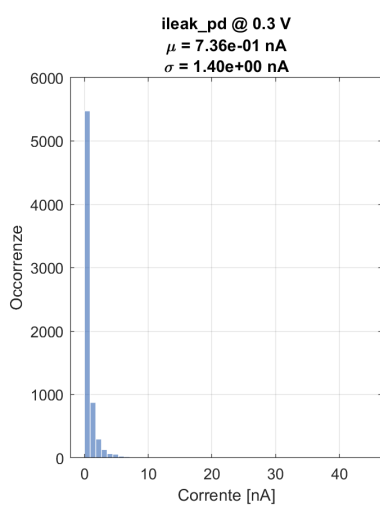
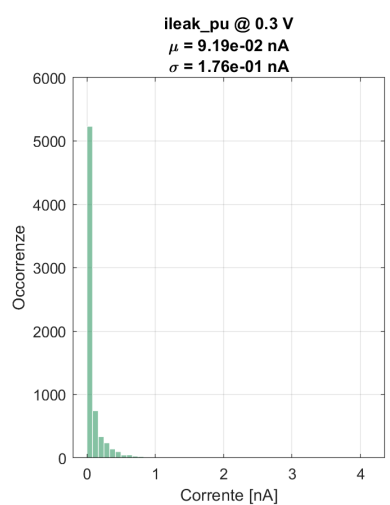
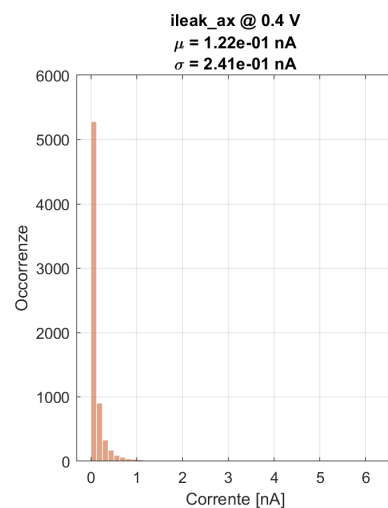
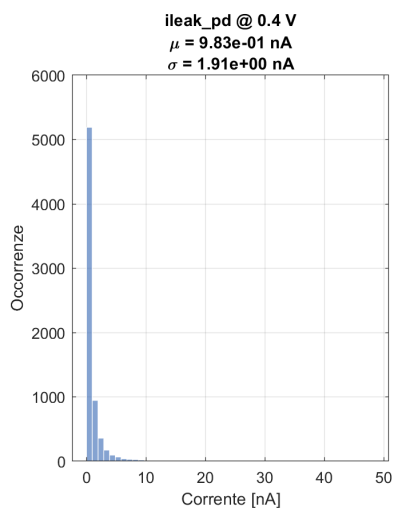
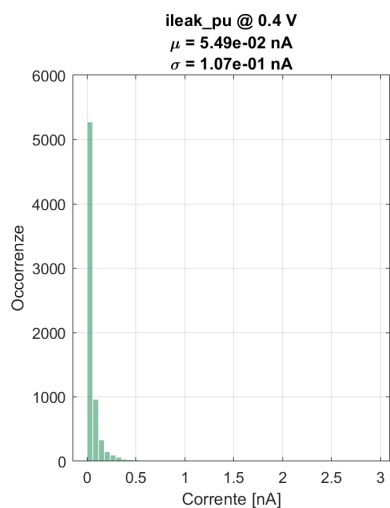
Figura 16: Configurazione circuitale per l'estrazione delle componenti di leakage.

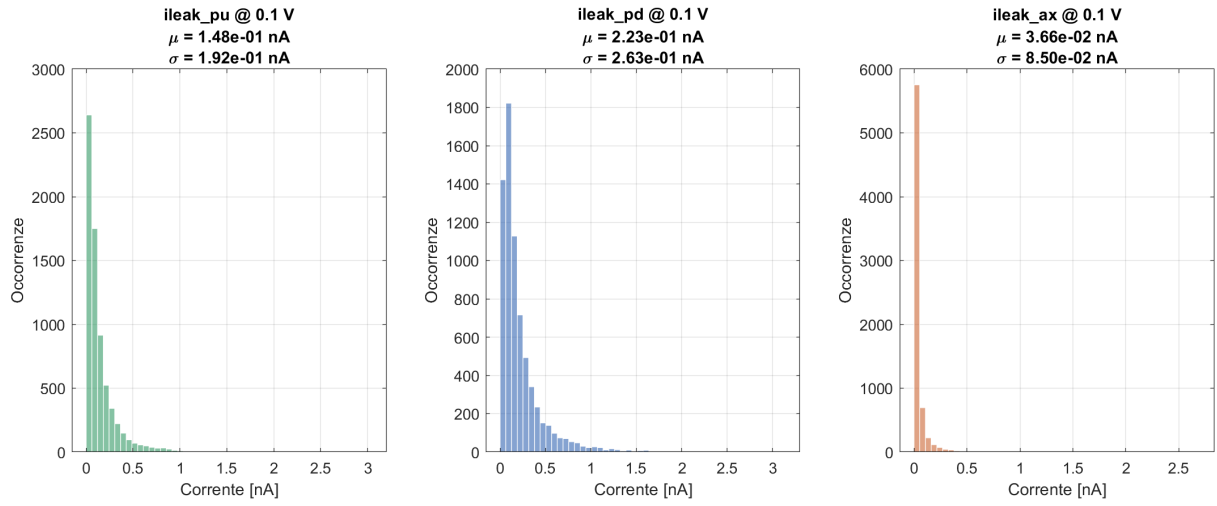
4.3.1 Distribuzioni statistiche delle correnti ($1V \rightarrow 0.1V$)

Le simulazioni Monte Carlo rivelano come il leakage sia fortemente influenzato dalle variazioni della soglia (V_{th}). Di seguito si riportano le distribuzioni statistiche scendendo nel regime di sottosoglia; si nota come, con il diminuire della tensione, la dispersione statistica aumenti sensibilmente.









4.3.2 Andamento globale e potenza dissipata

L'analisi complessiva delle medie statiche è mostrata in Figura 21. Come previsto, all'abbassarsi della tensione le correnti calano drasticamente; si nota inoltre come la componente di accesso sia trascurabile rispetto alle altre, grazie all'effetto body e alla tensione V_{DS} nulla in fase di hold.

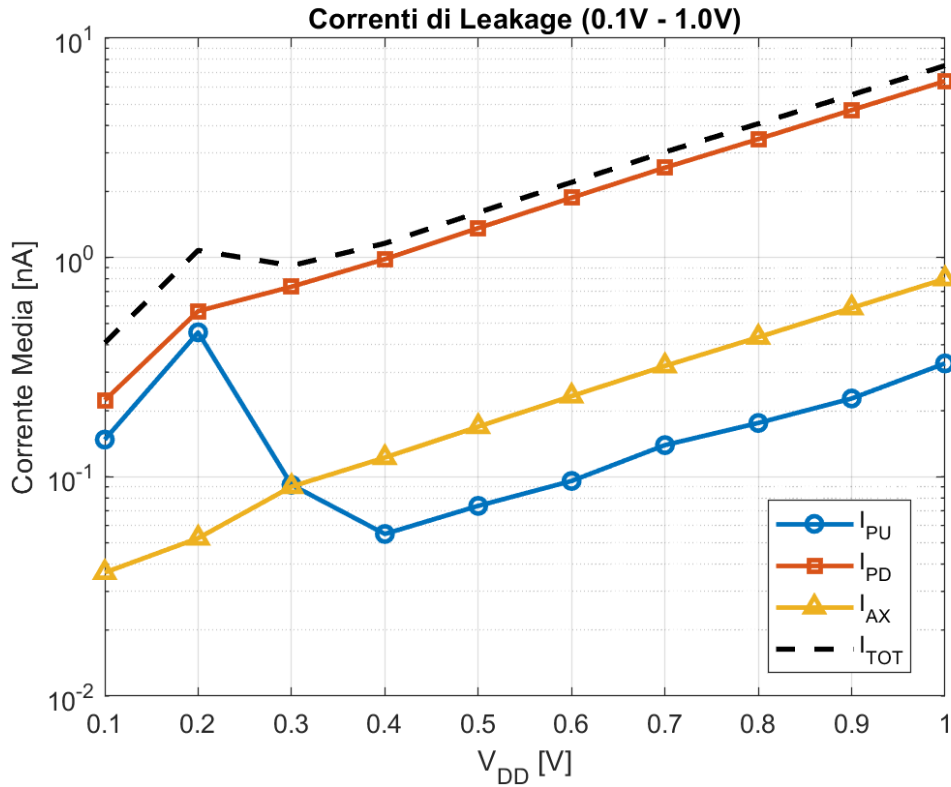


Figura 21: Andamento delle correnti di leakage medie in scala logaritmica.

Infine, lo scaling della potenza di leakage media totale ($P_{leak} = I_{leak_{tot}} \cdot V_{DD}$) è riportato in Figura 22. Si nota come il risparmio di potenza sia di diversi ordini di grandezza scendendo verso il limite di operatività della cella.

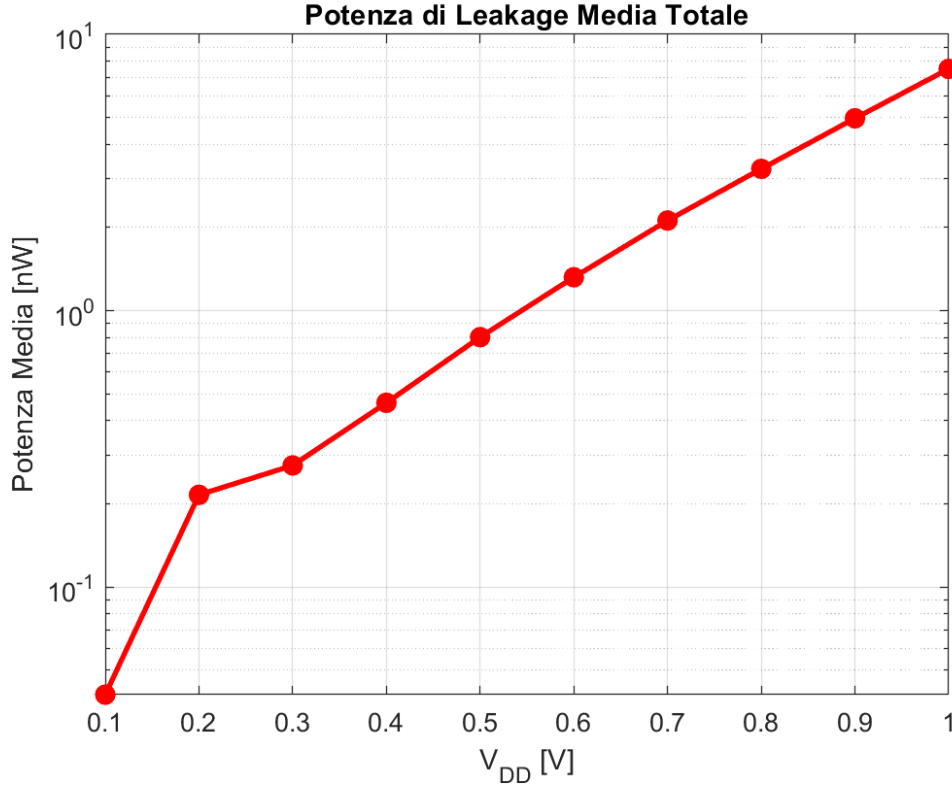


Figura 22: Scaling della potenza di leakage media totale al variare di V_{DD} .

Ecco una tabella riassuntiva di tutti i valori delle correnti di leakage trovati:

V_{DD} [V]	I_{PU} [nA]	I_{PD} [nA]	I_{AX} [nA]	I_{TOT} [nA]	P_{TOT} [nW]
1.0	0.33 ± 0.75	6.37 ± 11.39	0.80 ± 1.46	7.50	7.50
0.9	0.23 ± 0.45	4.70 ± 9.05	0.59 ± 1.17	5.52	4.97
0.8	0.18 ± 0.37	3.47 ± 6.27	0.43 ± 0.80	4.08	3.26
0.7	0.14 ± 0.33	2.57 ± 4.73	0.32 ± 0.60	3.03	2.12
0.6	0.10 ± 0.19	1.87 ± 3.72	0.23 ± 0.47	2.20	1.32
0.5	0.07 ± 0.16	1.36 ± 2.52	0.17 ± 0.32	1.60	0.80
0.4	0.05 ± 0.11	0.98 ± 1.91	0.12 ± 0.24	1.16	0.46
0.3	0.09 ± 0.18	0.74 ± 1.40	0.09 ± 0.18	0.92	0.28
0.2	0.46 ± 0.26	0.57 ± 0.73	0.05 ± 0.13	1.08	0.22
0.1	0.15 ± 0.19	0.22 ± 0.26	0.04 ± 0.09	0.41	0.04

Tabella 1: Sintesi statistica del leakage al variare della tensione.

Dall'analisi sistematica delle componenti di corrente, emerge che il contributo predominante al leakage totale è riconducibile al transistor di Pull-down (M_1). Al contrario, i contributi derivanti dai transistor di accesso (M_4) e di Pull-up (M_2) risultano inferiori di almeno un ordine di grandezza.

Per quanto concerne i transistor di Pull-down e di accesso, si osserva una riduzione esponenziale della corrente al decrescere della tensione di alimentazione V_{DD} , in linea con la teoria della conduzione di sottosoglia. Un comportamento peculiare si riscontra invece nel transistor di Pull-up: in seguito allo scaling nel range $[0.2\text{ V} - 0.4\text{ V}]$, si rileva un incremento esponenziale della corrente di leakage. Tale fenomeno è fisicamente plausibile e imputabile alle ridotte dimensioni del dispositivo e al regime di sottosoglia profonda in cui opera; in tali condizioni, le

variazioni della tensione di soglia e gli effetti di secondo ordine (come il *Drain-Induced Barrier Lowering*) possono portare a un incremento dei parassiti statici.

Tuttavia, questo aumento marginale della corrente di Pull-up non altera l'andamento globale della dissipazione: la potenza di leakage complessiva continua a diminuire drasticamente al decrescere di V_{DD} , come mostrato in Figura 22. Si può dunque concludere che la riduzione della tensione di alimentazione rappresenti la strategia più efficace per il contenimento della potenza statica, garantendo un risparmio energetico di diversi ordini di grandezza.

5 Analisi del Data Retention Voltage (DRV)

Il *Data Retention Voltage* (DRV) è definito come la tensione minima di alimentazione alla quale una cella SRAM riesce a mantenere correttamente lo stato logico memorizzato. Il superamento di questo limite durante lo scaling della V_{DD} comporta la perdita del dato (*flipping* della cella), rendendo il DRV il parametro critico per definire il limite operativo in modalità *standby*.

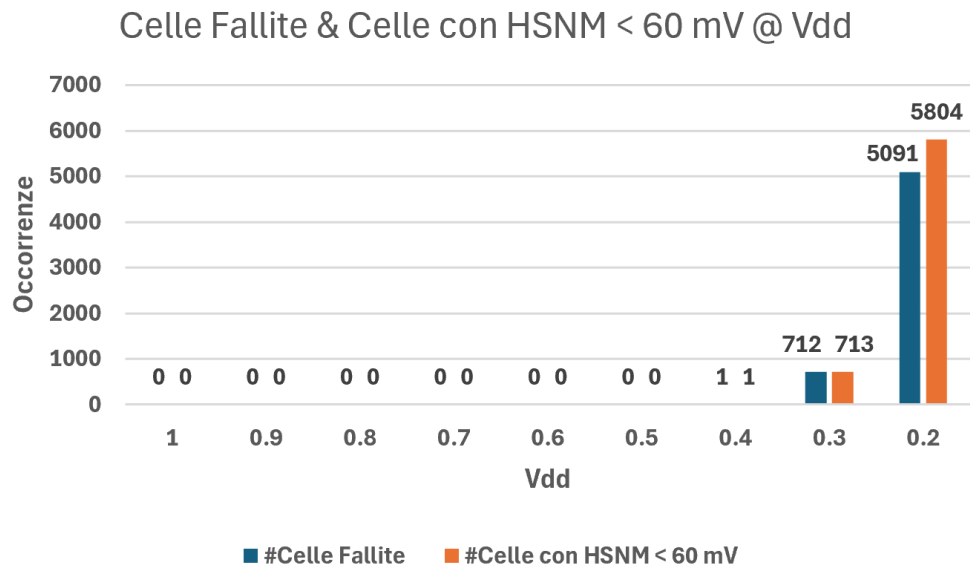
Per garantire la robustezza del sistema, è necessario che la condizione $V_{DD} > DRV$ sia sempre soddisfatta. Nella pratica industriale, si aggiunge solitamente un margine di sicurezza (*safety margin*) di circa 100 mV per compensare variazioni termiche, rumore di alimentazione o fluttuazioni locali del processo non modellate dalla simulazione.

5.1 Risultati statistici e Yield

L'analisi è stata condotta monitorando il numero di celle fallite su un campione di 7000 run Monte Carlo per ogni step di tensione. Un fallimento viene registrato quando l'HSNM scende al di sotto della soglia di sicurezza di 60 mV . I risultati sono sintetizzati nella Tabella 2.

Tabella 2: Analisi dei fallimenti e stima dello Yield al variare della V_{DD} .

V_{DD} [V]	# Celle Fallite (Flipped)	# Celle con HSNM < 60 mV
1.0	0	0
0.9	0	0
0.8	0	0
0.7	0	0
0.6	0	0
0.5	0	0
0.4	1	1
0.3	712	713
0.2	5091	5804



Come si evince dai dati, la cella presenta uno Yield del 100% fino a 0.5 V. Il primo cedimento statistico si verifica a 0.4 V, mentre a tensioni inferiori il tasso di fallimento diventa inaccettabile (> 10% a 0.3 V e > 70% a 0.2 V). Sulla base di queste osservazioni, il DRV statistico della cella può essere fissato a 0.4 V. Applicando il margine di sicurezza di 100 mV, si determina che la **tensione di alimentazione minima sicura** per questa specifica implementazione SRAM è pari a **0.5 V**. Questo punto di lavoro garantisce la massima efficienza energetica (riducendo drasticamente il leakage, come visto nel capitolo precedente) senza compromettere l'integrità dei dati memorizzati.