SIMULAZIONI MEDIANTE SOFTWARE SIMETRIX

Riccardo Fontanini

A. A. 2018 - 2019

UNIVERSITÀ DEGLI STUDI DI UDINE

Corso di DEVICES AND CIRCUITS WITH HIGH ENERGY EFFICIENCY FOR IOT APPLICATIONS

1 Introduzione

1.1 Il modello

Il seguente documento raccoglie i risultati di diverse simulazioni di circuiti basati su transistori NMOS e PMOS che sfruttano il modello semi-empirico che esprime la corrente I_{DS} come:

NMOS

$$I_{DS} = W\ Q_{SUB} F_{SD} v_{th} = W\ I_0 \ln\left(1 + exp\left(\frac{V_{gs} - V_t}{n\phi_t}\right)\right)\left(1 - \exp\left(-\frac{V_{DS}}{\phi_t}\right)\right)$$

PMOS

$$I_{DS} = W\ Q_{SUB} F_{DS} v_{th} = W\ I_0 \ln\left(1 + exp\left(-\frac{V_{gs} - V_t}{n\phi_t}\right)\right)\left(1 - \exp\left(\frac{V_{DS}}{\phi_t}\right)\right)$$

In entrambi i casi $V_t = V_{t0p} - \lambda_{DS} V_{DS} + \lambda_{SB} V_{SB}$

Mentre $I_0 = \frac{t_{CH}}{2 - t_{CH}}\ C_{inv}\ n\ \phi_t\ v_{th}$

Il modello del NMOS (PMOS omesso perché equivalente) è quindi stato scritto come segue:

```
**********************************
*UNIVERSITA' DEGLI STUDI DI UDINE
*GRADUATION THESIS
*
*Grad Student:BERTONI LUCA, RICCARDO FONTANINI
*Supervisor:Prof. ESSENI DAVID
*Assistant Supervisor: Dott. Ing. ROLLO TOMMASO*
*
*MODELLO NMOS
*14/02/2019
**********#PARAMETERS#*****************
*dimensions.........W,L=[m];
*thermal voltage....Vth=[V];
*non-ideal factor...n=[a.u.];
*thermal velocity...Term_V=[m/s];
*oxide cap..........Cinv=[F/m2];
**************************************
.subckt SUB_NMOS drainin gatein sourcein
.param Vt0 = {{(Vt0n)}}
.param Vth= {BOLTZ * {(T)} / ECHARGE}
.param Vt= {{(Vt0)}-{Kds}*V(drain,source)}
.param I0= {{(Cinv)} * {(n)} * Vth * Term_V * ( {(tch)} / ( 2 - {(tch)} )) }
bswitch drain source I=  {{(W)} * I0 * LN(1+exp((V(gate,source)-Vt)/( {(n)} * Vth )))*(1-
exp(-(V(drain,source))/ Vth )) }
*Parasitic drain, source, and gate resistances*
rd drainin drain {(rpara)}
rs sourcein source {(rpara)}
rg gatein gate {(.6)}
* Cg capacitance*
C_GD gate drain  {1e-16 * {(W)} * 1e6}
C_GS gate source  {1e-16 * {(W)} * 1e6}
BC_GS gatein sourcein Q= {0.5 * {(Cinv)} * {(W)} * {(L)} * {(n)} * Vth *LN(1+exp(
(V(gate,source)-Vt)/( {(n)} * Vth )  )) }
BC_GD gatein drainin  Q= {0.5 * {(Cinv)} * {(W)} * {(L)} * {(n)} * Vth *LN(1+exp(
(V(gate,drain)-Vt)/( {(n)} * Vth )  )) }
.ends
```

Al fine di simulare il comportamento di tale modello tramite l'ausilio del software Simetrix[1], si è deciso di parametrizzare il file descrittore, ed applicare i parametri dei transistori direttamente nell'ambiente simulativo. Questo ha permesso di velocizzare le operazioni di creazione del circuito da simulare ed una migliore elasticità dei modelli implementati.

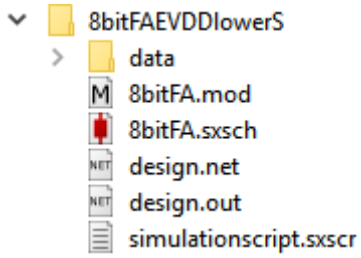
I valori standard presi come riferimento sono:

T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,35	-0,35	1	0,7	1,23e5	0,02	0,04	60

Una simulazione può avere parametri differenti dai precedenti al fine di enfatizzare un determinato comportamento del modello impiegato.

1.2 Progetto di simulazione

Ogni progetto di simulazione è creato per rappresentare una determinata caratteristica che si vuole simulare. Ad ogni progetto è assegnata una cartella con un nome specifico. All'interno di essa sono presenti diversi file, i cui principali sono:

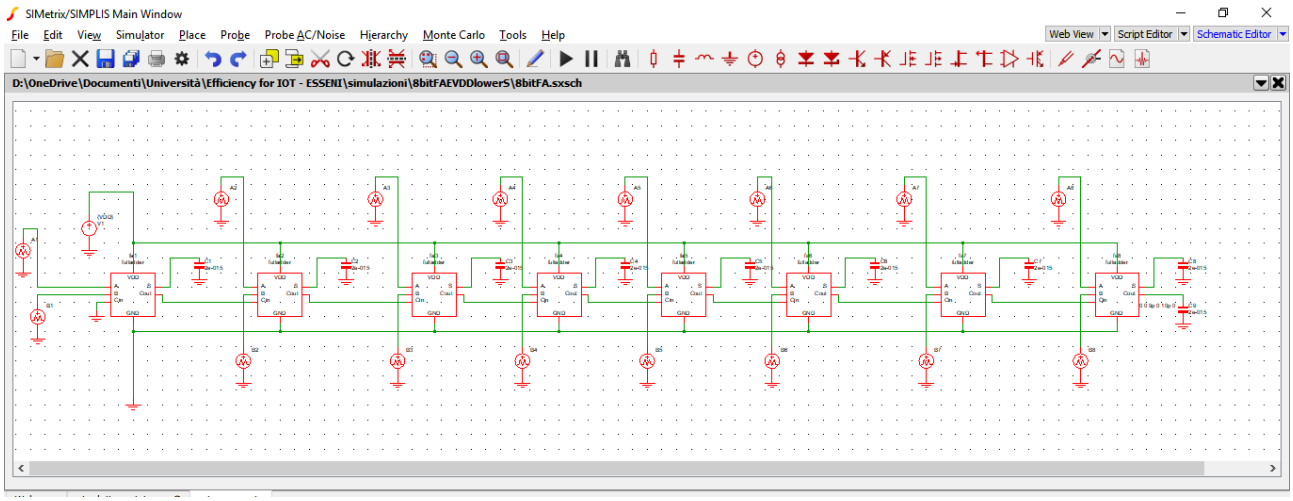


- un file *.sxsch che rappresenta la topologia del circuito da simulare
- un file *.sxscr contenente uno script SIMetrix che definisce modalità, tipologie di simulazione nonché i valori dei componenti utilizzati nei circuiti ed i parametri di modello.
- Una cartella data dove sono salvati i dati prodotti da varie simulazioni, che in certi casi possono essere onerose dal punto di vista tempistici.

1.2.1 I file schematici

È stato deciso di utilizzare gli schematici solamente per definire facilmente la topologia del circuito, senza che vengano imposte grandezze da misurare direttamente sullo schematico (non vengono utilizzati i probes). Questo perché le grandezze da misurare possono essere definite dinamicamente nello script file e questa metodologia permette di raggruppare i dati in modo da poterli elaborare sequenzialmente in maniera più efficiente ed automatica.

[1] https://www.simetrix.co.uk/



In questa fase, cioè quando si crea la topologia del progetto, non è necessario definire i valori dei dispositivi inseriti nello schematico, infatti verranno anch'essi definiti dinamicamente dallo script file.

1.2.2 Gli script file

SIMetrix permette l'utilizzo di particolare Script (scritti in un linguaggio simile al BASIC) per automatizzare diverse operazioni. In queste simulazioni i file di scripting sono utilizzati ampiamente per definire dinamicamente parametri simulativi, tipologie di simulazioni e tutto il necessario all'ambiente simulativo per compiere le operazioni necessarie. Sono stati utilizzato i cicli per esegue simulazioni ricorsive su parametri ottenuti da simulazioni precedenti. Per quanto riguarda la sintassi si rimanda alle pagine[2] dedicate. Ogni progetto ha un suo script dedicato che varia in base alle operazioni da svolgere. In ogni caso la struttura di base rimane sempre la stessa:

1. Definizione e applicazione dei parametri sul circuito da simulare;

```
75 Unselect
76 Select /prop Ref C1
77 Prop value {stdCap}
```

2. Generazione della netlist e simulazione

```
173
174         Netlist design.net
175         Run /noerr /file design.net /an {'.TRAN 0 '& STR(propagationtimeinterval) &' 0 '& STR(maxstep)}
176
```

3. Elaborazione dei dati di simulazione

```
291
292         let supplycurrent = V1#n
293         let powerSupply = supplycurrent * VDDs[idVDD]
294         let energyHigh = integ (powerSupply)
295
```

4. Visualizzazione dei risultati

[2] https://help.simetrix.co.uk/8.0/simetrix/simetrix_docs.htm#mergedProjects/script_manual/topics/sr_com_curve.htm

```
311
312 for energyindex = 0 to NEnergy-1
313     let energy = vector(NVDD)
314     for i=0 to NVDD-1
315         let energy[i] = totalenergies[energyindex * NVDD + i]
316     next NVDD
317     if energyindex == 0 then
318         Plot XY(energy, VDDs) /name {'Energy curve ' & str(energyindex)} /ylabel {'Energy / OP'} /yunit {'J'}
319     else
320         Curve XY(energy, VDDs) /name {'Energy curve ' & str(energyindex)}
321     endif
322 next energyindex
```

Ogni volta che si lancia una simulazione, SIMetrix crea un Gruppo dove salvare, per ogni step simulativo, i valori di ogni nodo della rete. In questo modo è possibile lanciare diverse simulazioni salvando di volta in volta l'identificativo del Gruppo creato per tale simulazione. L'elaborazione e la visualizzazione dei dati di una determinata simulazione possono quindi essere eseguite a posteriori, richiamando l'identificativo del Gruppo che si vuole elaborare.

L'impostazione così definita è stata scelta per facilitare l'utilizzo da parte di terzi degli ambienti simulativi che altrimenti potrebbero risultare piuttosto ostici nelle fasi iniziali del loro impiego. Inoltre, essendo il codice intrinsecamente sequenziale, facilita molto la lettura delle operazioni svolte e la replicabilità di simulazioni complesse come quelle attuate in questo lavoro.

1.2.3 Nota

In questa relazione, al di sotto di ogni titolo di capitolo, è presente una didascalia tra parentesi quadre [esempio]. Questo è il riferimento alla cartella contenente il progetto della simulazione. Successivamente sono elencati i parametri simulativi e di modello impiegati per ottenere nuovamente i risultati che verranno trattati nel corso del rispettivo capitolo. Saranno inoltre presenti dei riferimenti a formule contenute nelle dispense del corso di DEVICES AND CIRCUITS WITH HIGH ENERGY EFFICIENCY FOR IOT APPLICATIONS tenuto dal professor D. Esseni nell'anno accademico 2018 - 2019.

2 Capacità di gate

[Gatecapacity]

2.1 Parametri di simulazione

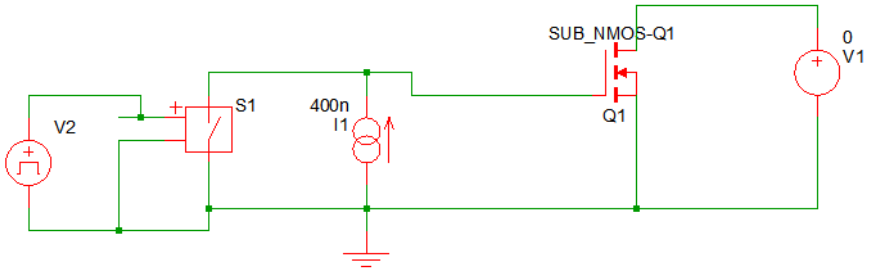
- maxTime = 400p
- NMAX = 50

2.2 Parametri del modello

T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,35	-0,35	1	0,7	1,23e5	0,02	0,04	60

2.3 Circuito e simulazione

Il circuito impiegato per la simulazione della capacità di gate di un transistore fa uso di un generatore ideale di corrente continua che al tempo t = 0 risulta cortocircuitato su un interruttore ideale. Successivamente l'interruttore viene aperto imponendo una corrente costante al terminale di gate dell'NMOS. Tale corrente tende a caricare gradualmente la capacità di gate, quindi misurando la tensione di gate è possibile ottenere (sapendo che I1 è costante) la capacità di gate al variare della tensione Vgs.

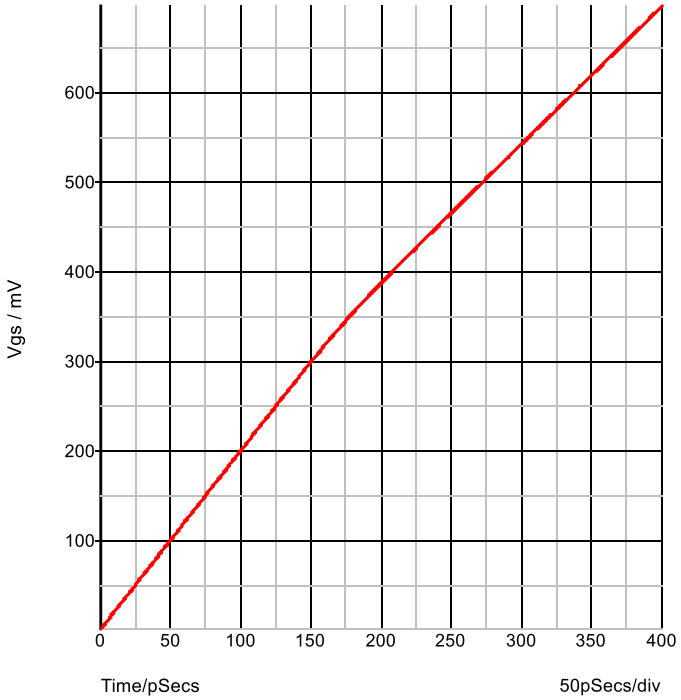


Infatti:

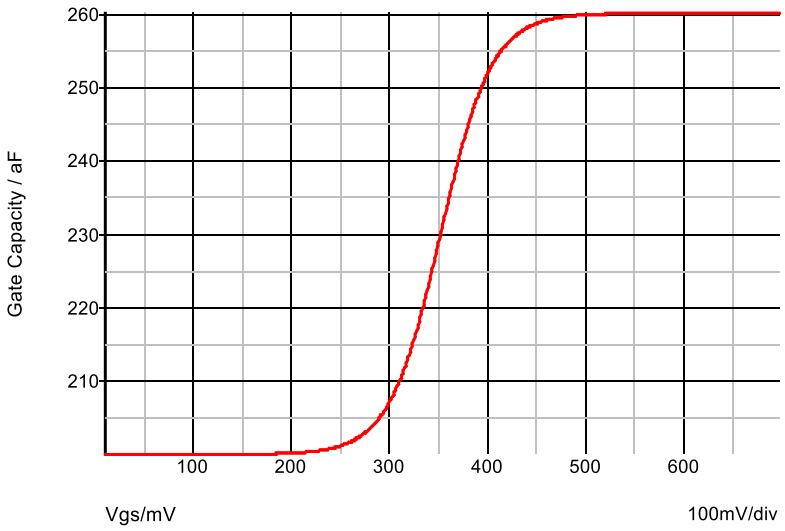
$$C = \frac{\partial Q}{\partial V_G} = \frac{\partial Q}{\partial t}\frac{\partial t}{\partial V_G} = i_G \left(\frac{\partial V_G}{\partial t}\right)^{-1}$$

Cioè differenziando la tensione di ingresso, facendone il reciproco e successivamente moltiplicando per la corrente di ingresso è possibile ottenere la capacità di ingresso al gate.

La tensione al gate simulata è:



Mentre la capacità simulata ha il comportamento in figura:



Come è possibile notare da questa ultima immagine, a tensioni basse (sottosoglia) la capacità di gate è dominata prevalentemente dai parassiti. Successivamente, quando il transistore va in inversione, la capacità si alza notevolmente per via dell'accumulazione di carica di canale.

3 Analisi delle caratteristiche di una porta NOT

[NOT]

3.1 Parametri di simulazione

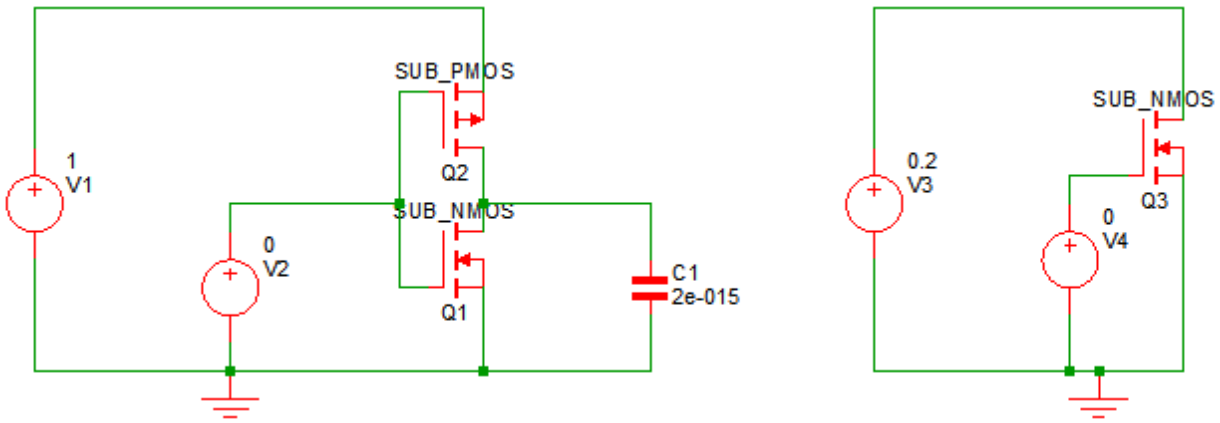
- VDD = 1
- stdCap = 2f
- minVGSstep = 10m
- maxVDD = 1
- minVDD = 0.1
- NVDD = 10
- maxKDS = 0.2
- NKDS = 5

3.2 Parametri del modello

T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,3	-0,3	1.3	0,7	1,23e5	0,02	0,2	60

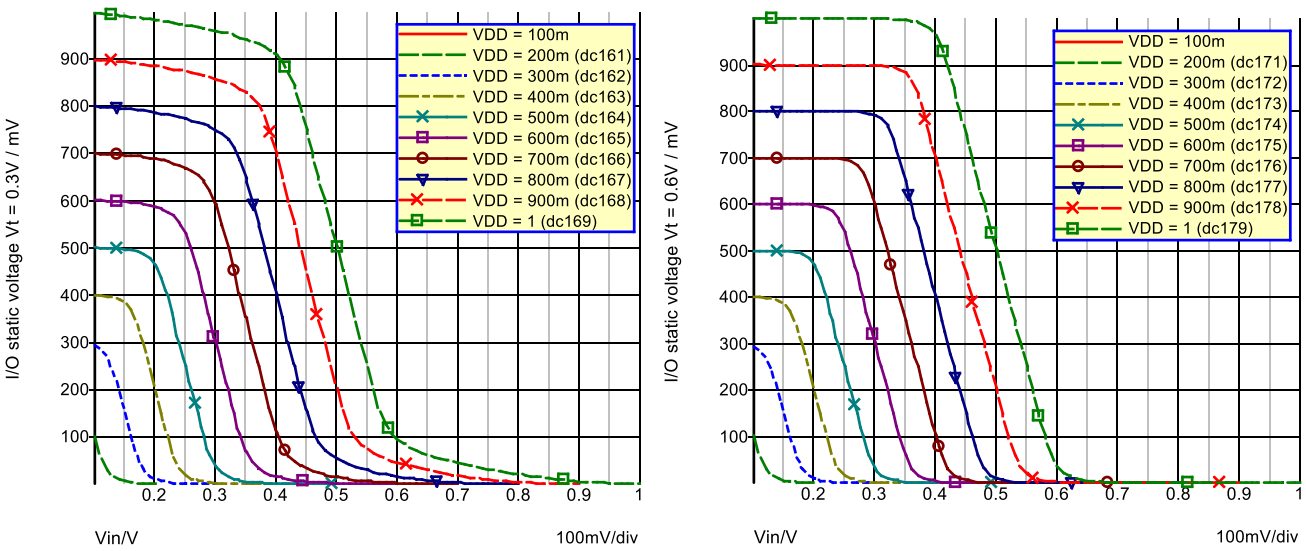
3.3 Circuito e simulazione

Per la simulazione della porta not e la comprensione di come variano le caratteristiche di tale porta logica in relazione alla variazione dei parametri del transistore è necessario utilizzare i due circuiti mostrati nella figura precedente:



Il primo circuito è stato utilizzato per l'analisi della transcaratteristica, mentre il secondo per la dipendenza del DIBL sul singolo transistore e della resistenza di contatto.

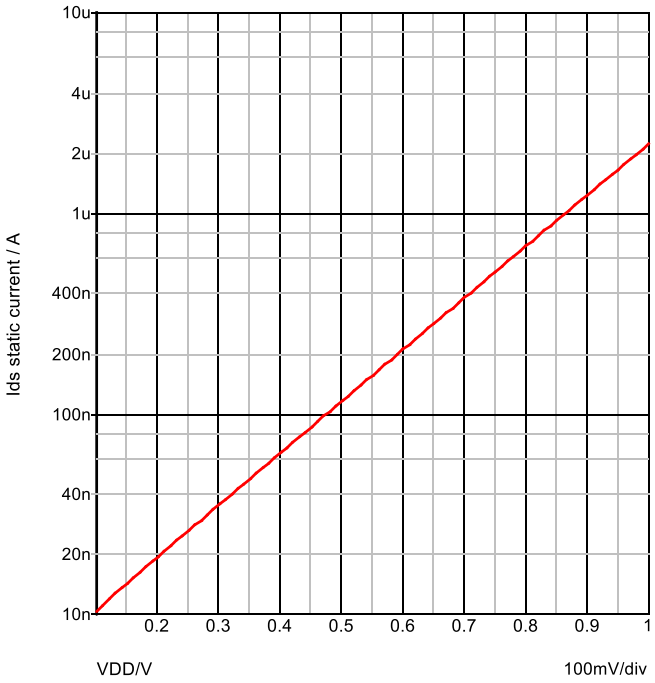
L'analisi più semplice da operare sulla porta not è quella della transcaratteristica ingresso-uscita. In particolare, si può notare dai seguenti grafici come tale transcaratteristica abbia profili più netti al crescere della V_{t0} .



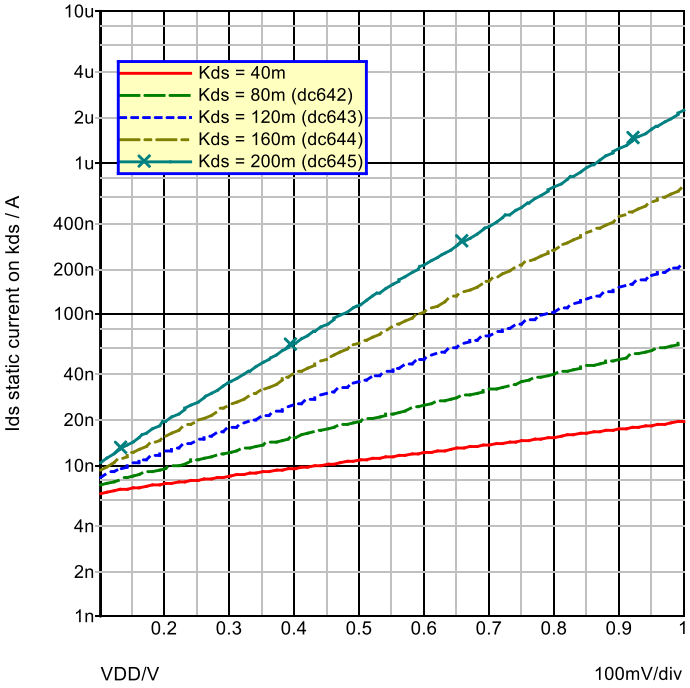
La corrente statica quando l'NMOS è spento (mentre il PMOS è acceso) è definita dalla Ioff del NMOS, cioè dall'equazione 2.11 con Vgs = 0:

$$I_{ds} = W\, I_0 \ln\left(1 + e^{\frac{V_{GS}-V_t}{n\phi_t}}\right)\left(1 - e^{-\frac{V_{DS}}{\phi_t}}\right)$$

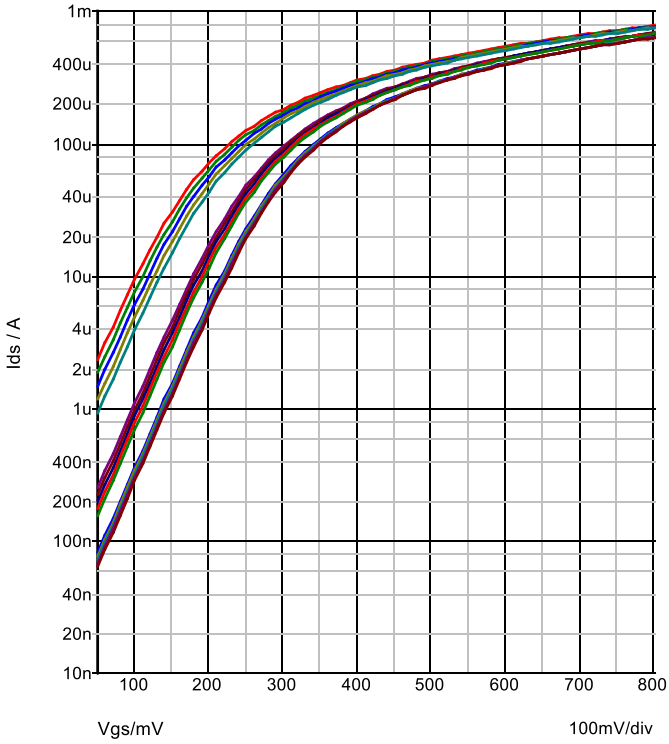
Con $I_0 = \left(\frac{t_{ch}}{2-t_{ch}}\right) C_{inv}\, n\, \phi_t$



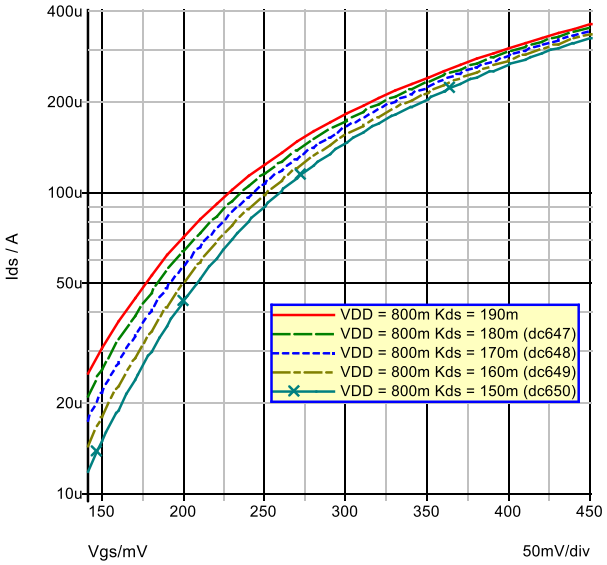
Si noti come il DIBL influisca molto sulla corrente OFF del transistore: diminuendo l'effetto di abbassamento della barriera prodotto dal drain (DIBL) la corrente Ids si abbassa sensibilmente in quanto la barriera di potenziale nel canale è molto più dipendente dalla tensione di gate (in questo caso Vgs = 0).



L'effetto del DIBL è presente anche sulla transcaratteristica del singolo transistore. Nella seguente simulazione si è deciso di considerare 3 tensioni di alimentazione applicate a Vds (0.8, 0.4 e 0.2) V e di far variare linearmente il parametro λ_{DS}, per ogni tensione di alimentazione,



Per maggiore chiarezza è stato considerato un range di tensione Vgs attorno alla Vt e visualizzate solamente le curve per Vds = 0.8V:

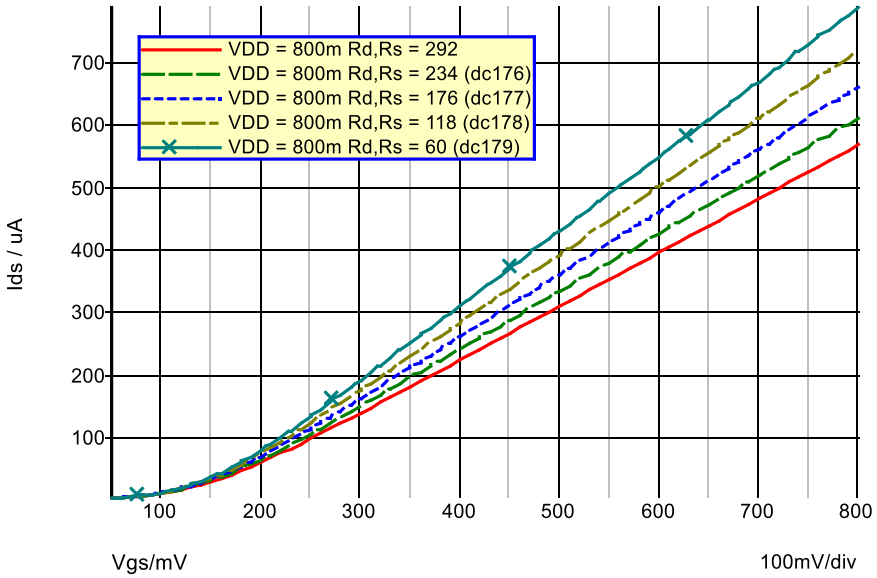


Come si vede da tale grafico, al crescere del DIBL la curva si sposta verso sinistra, cioè verso più alte correnti. Tale andamento è spiegato dalla equazione:

$$I_{DS} \cong W I_0 10^{\frac{V_{GS}-V_{t0}}{S}} 10^{\frac{\lambda_{DS} VDD}{S}}$$

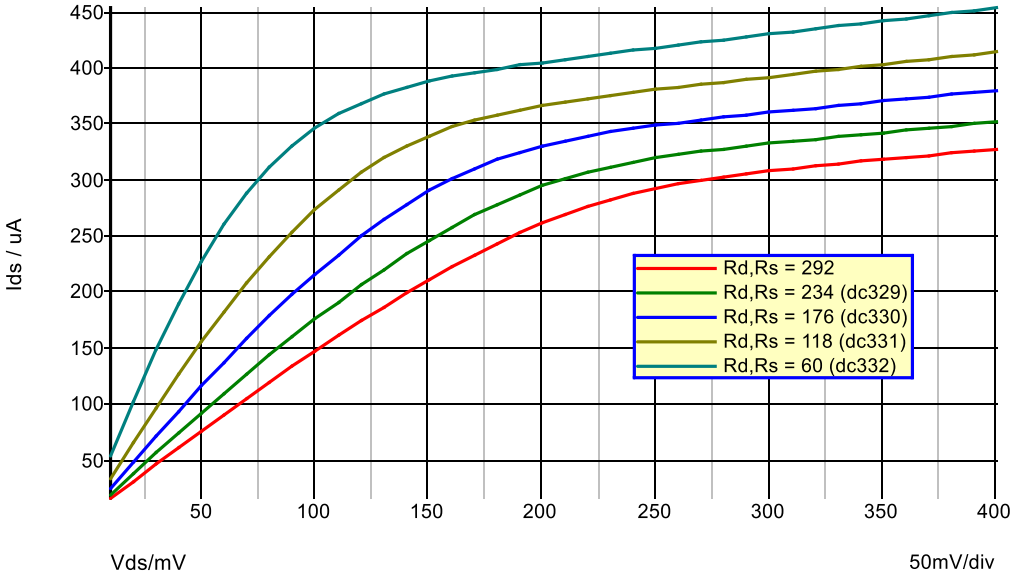
Che descrive l'andamento della corrente del transistore in sottosoglia. Si può notare la dipendente esponenzialmente dal DIBL.

È stata considerata oltremodo una variazione della resistenza di contatto al drain ed al source del transistore. Come è possibile notare dal grafico successivo, l'innalzamento della resistenza di contatto è deleterio per la corrente erogata dal transistore:



La dipendenza della corrente è strettamente legata alla "resistenza equivalente" del transistore. Infatti l'andamento lineare per tensioni di gate elevate fornisce una spiegazione esaustiva per i differenti coefficienti di linearità visibili nella simulazione.

Un ulteriore effetto della resistenza di contatto è visibile nella caratteristica di uscita del transistore:



In questo grafico si può vedere come mantenendo la Vgs costante la caratteristica, all'aumentare della resistenza di contatto, abbia una transizione meno brusca. Oltre a confermare un abbassamento della corrente complessiva come mostrato anche dal grafico precedente.

4 Metodo di riduzione dell'energia: stacking

[stacking]

4.1 Parametri di simulazione

- VDD = 1
- stdCap = 2f
- NVDD = 25
- maxVDD = 1
- minVDD = 0.1
- propagationtimeinterval = 350p
- maxstep = 0.5p

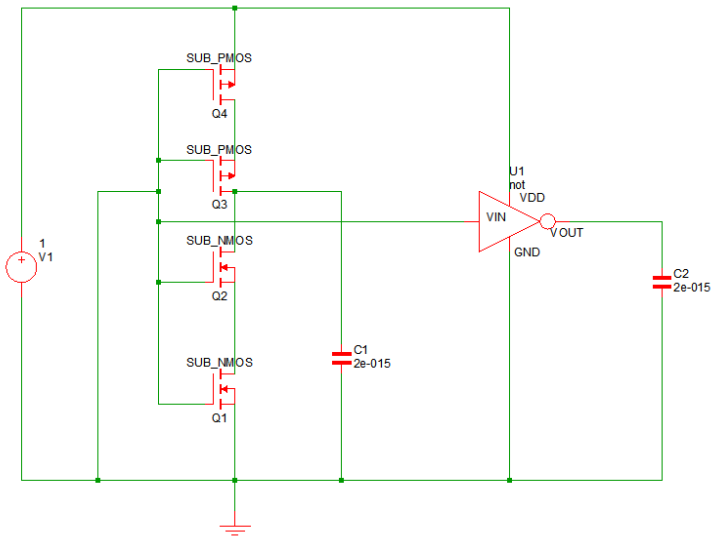
4.2 Parametri del modello

T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,3	-0,3	1.3	0,7	1,23e5	0,02	0,2	60

4.3 Circuito e simulazione

Per la simulazione dello stacking di transistori è stato utilizzato un circuito contenente 2 blocchi fondamentali:

- Il primo è una porta not con i transistor in configurazione stacking a coppie;
- Il secondo è una porta not standard, senza alcun tipo di accorgimento circuitale al fine di ridurre il leakage (possiede la medesima tecnologia).



La prima analisi che andremo a considerare sarà quella inerente alla tensione intra-stacking (Vx) cioè andremo ad analizzare come varia la tensione al source di Q2 (vedi figura) quando i due NMOS sono spenti.

Supponendo che i dimensionamenti dei transistori siano uguali tra loro, il sistema di equazioni 3.1 permette di calcolare la corrente Ids che passa attraverso lo stacking:

$$I_{UP} = W_{UP}\, I_0 10^{-\frac{(V_{T0}+\lambda_{SB}V_X)}{S}} 10^{-\frac{V_X}{S}} 10^{-\frac{\lambda_{DS}(VDD-V_X)}{S}} \left(1 - e^{-\frac{(VDD-V_X)}{\phi_T}}\right)$$

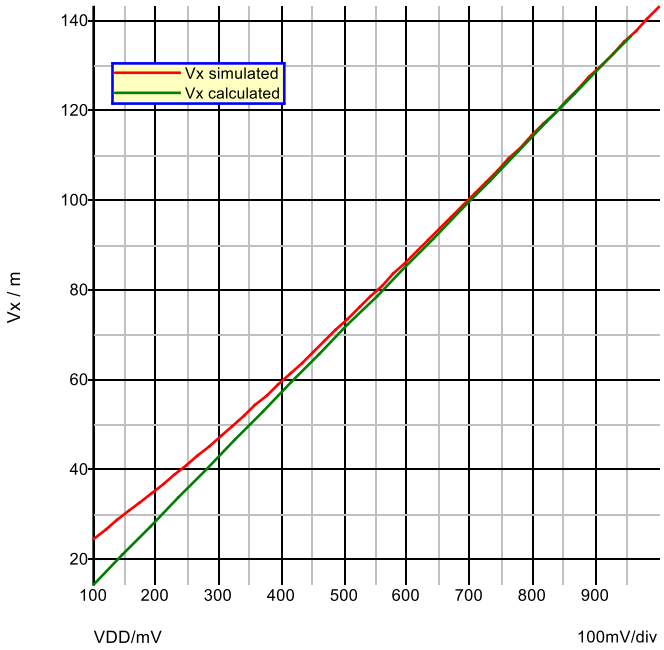
$$I_{DOWN} = W_{DOWN}\, I_0 10^{-\frac{(V_{T0})}{S}} 10^{-\frac{V_X}{S}} 10^{-\frac{\lambda_{DS}V_X}{S}} \left(1 - e^{-\frac{(-V_X)}{\phi_T}}\right)$$

Nel nostro caso $\lambda_{SB} = 0, W_{UP} = W_{DOWN}$

Mettendo insieme le due equazioni si ottiene che:

$$V_x = \frac{\lambda_{DS}VDD + Slog\left(\frac{W_{UP}}{W_{DOWN}}\right)}{1 + 2\lambda_{DS} + \lambda_{SB}}$$

Simulando si ottiene il seguente risultato:

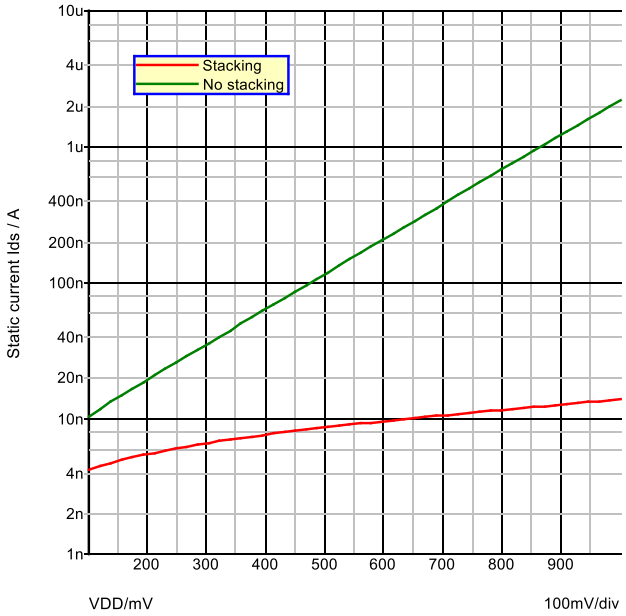


Come si vede c'è una piccola discrepanza per basse VDD tra Vx simulata ed il valore calcolato: infatti per tali tensioni non è più valida l'approssimazione 3.1b (versione semplificata per $V_x \gg 3 - 4\,\phi_t$. Sostituendo Vx nelle equazioni precedenti è possibile ottenere la corrente Ids dei transistori di stacking.

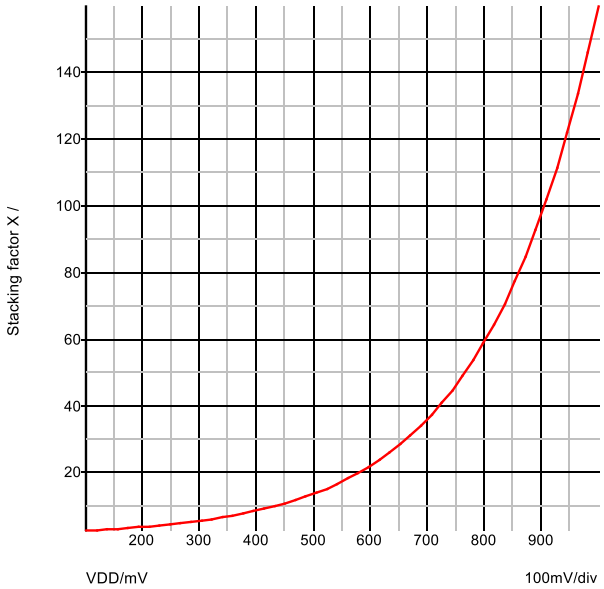
Si definisce quindi lo stacking factor come

$$X = \frac{I_{off}}{Istack} = 10^{\frac{\lambda_{DS}VDD}{S}} \left[\frac{1 + \lambda_{DS} + \lambda_{SB}}{1 + 2\lambda_{DS} + \lambda_{SB}}\right]$$

Per simulare tale fattore è possibile effettuare il rapporto tra la corrente Ids dei transistori in stacking e quella del transistore appartenente alla porta NOT semplice. Le due correnti al variare della VDD simulate sono:

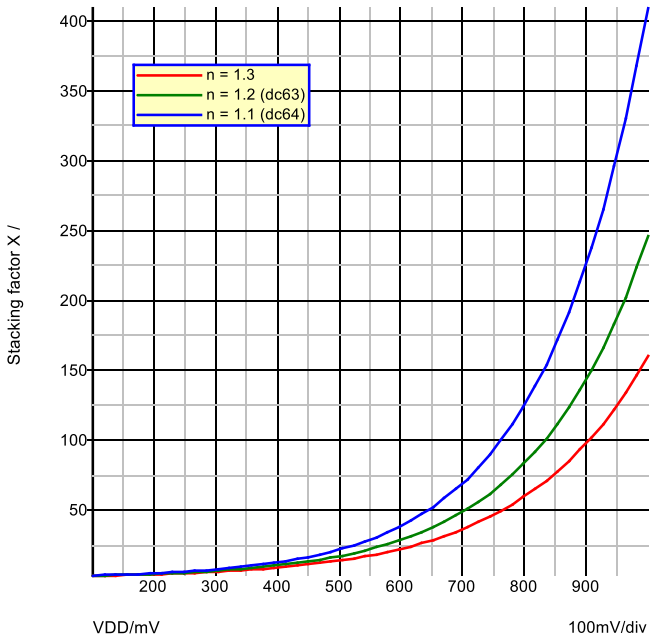


Facendo il rapporto tra le due si ottiene lo stacking factor X:



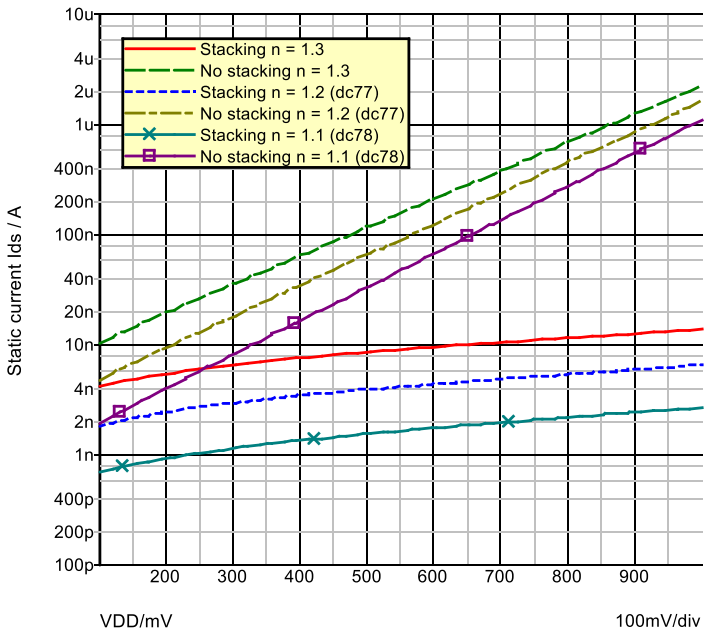
Variando S e λ_{DS}è possibile modificare le prestazioni dello stacking, in particolare facendo variare S:

- $n = 1.3 \rightarrow S = n\,\phi_t \ln(10) \cong 0.078 \frac{mV}{dec}$ (curve rosse)
- $n = 1.2 \rightarrow S = n\,\phi_t \ln(10) \cong 0.072 \frac{mV}{dec}$ (curve verdi)
- $n = 1.1 \rightarrow S = n\,\phi_t \ln(10) \cong 0.066 \frac{mV}{dec}$ (curve blu)



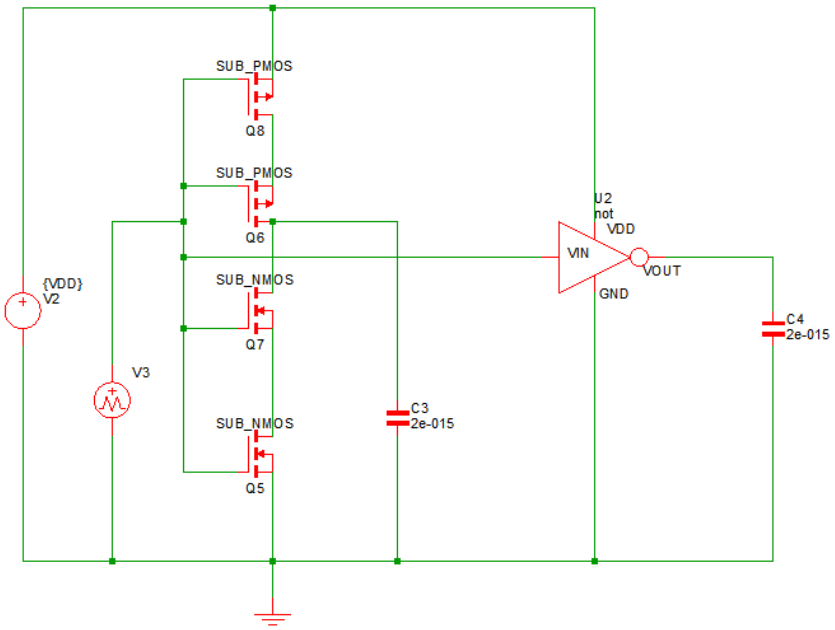
Si vede dal grafico precedente come migliorando le prestazioni del sottosoglia (S) è possibile aumentare il rapporto X.

Ma dal punto di vista delle correnti cosa implica? Diminuire S aumenta le prestazioni del sistema dal punto di vista energetico, quindi della corrente erogata. In assoluto le correnti diventano:

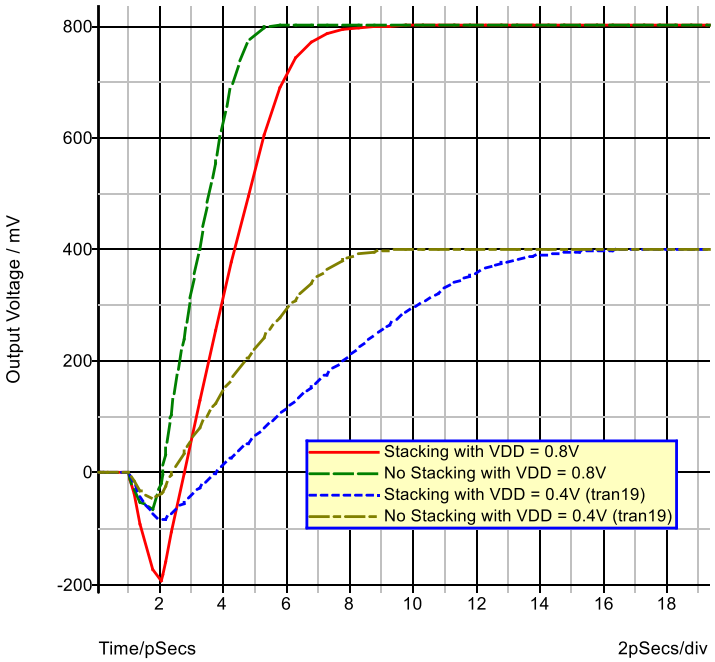


Da tale grafico si nota un accavallamento delle curve di corrente a tensioni di alimentazione basse. Questo vuol dire che lo stacking può non essere efficace per tali alimentazioni, se gioca a discapito dei parametri tecnologici. Per alte tensioni di alimentazione, invece, la discrepanza tra le correnti Ioff e di stacking è talmente alta che anche variando i parametri dei transistori non si nota una inefficacia di tale tecnica per la riduzione del consumo energetico.

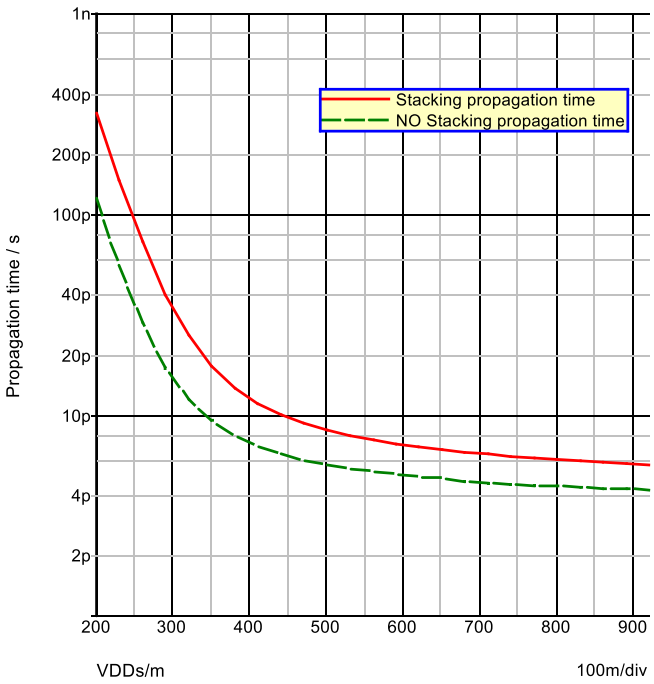
Lo stacking influisce anche sui tempi di propagazione: di seguito viene mostrata l'uscita dei due invertitori (stacking e standard) a fronte di uno stimolo in ingresso che porta a 0 la tensione di gate su entrambi i circuiti. Per simulare le caratteristiche dinamiche bisogna attuare una modifica allo schema di simulazione introducendo un generatore di tensione a gradino V3:



Si vede dal grafico come l'invertitore con lo stacking risponda più lentamente rispetto all'invertitore normale.



In generale è possibile svolgere tale analisi a differenti VDD e calcolare il tempo di propagazione a 0.9 * VDD per entrambi gli invertitori.



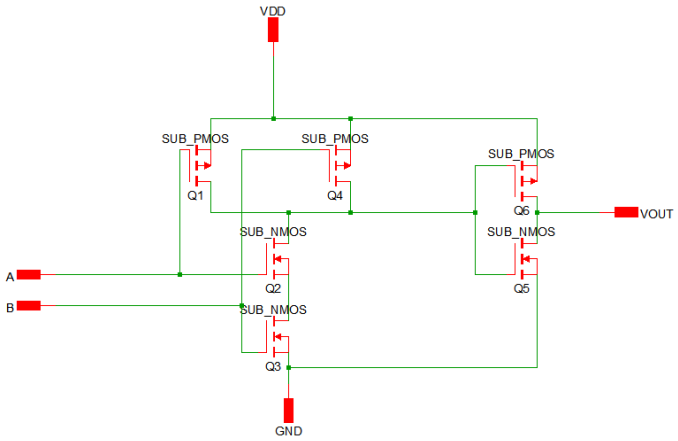
Come si può vedere dal precedente grafico, la soluzione con lo stacking è intrinsecamente più lento rispetto alla versione dell'invertitore tradizionale.

5 Analisi energetica di un Full Adder

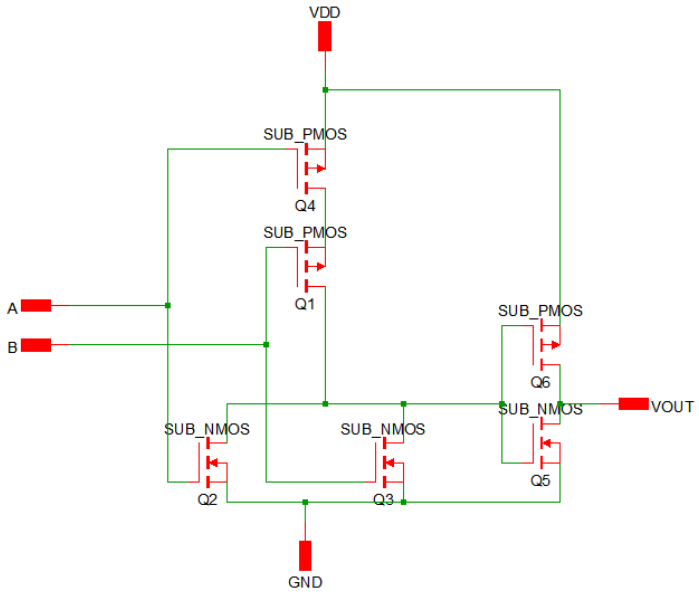
5.1 Circuito e simulazione

Successivamente alla caratterizzazione e simulazione del singolo transistor e della porta not, si è deciso di progettare e simulare altre porte logiche quali:

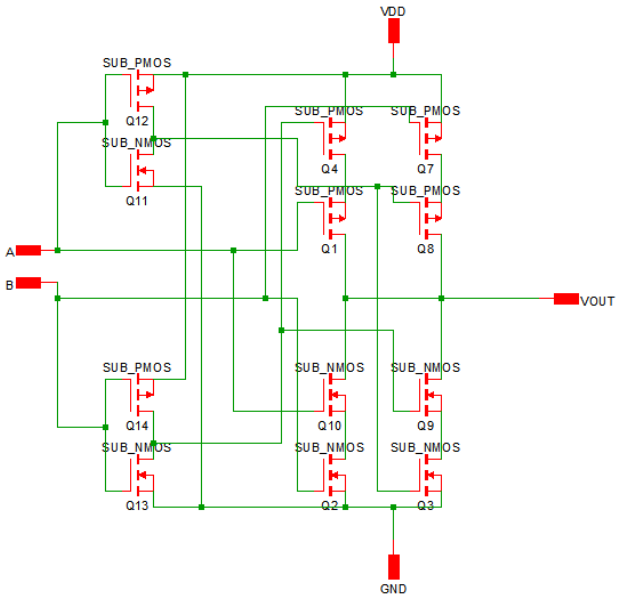
la porta AND



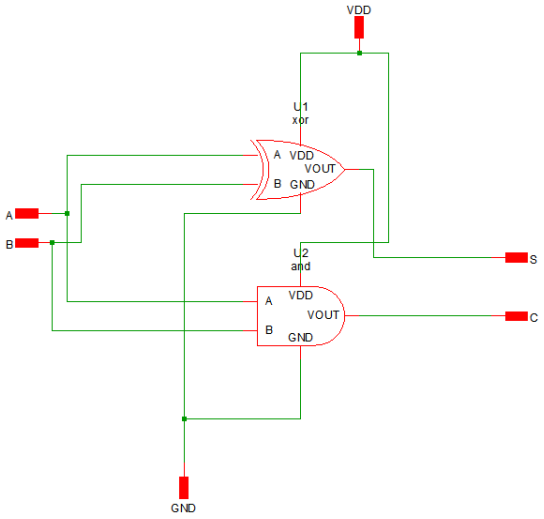
La porta OR:



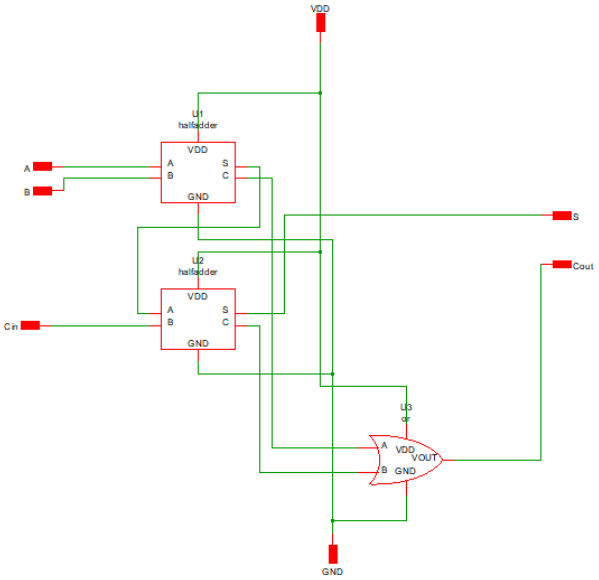
La porta XOR:



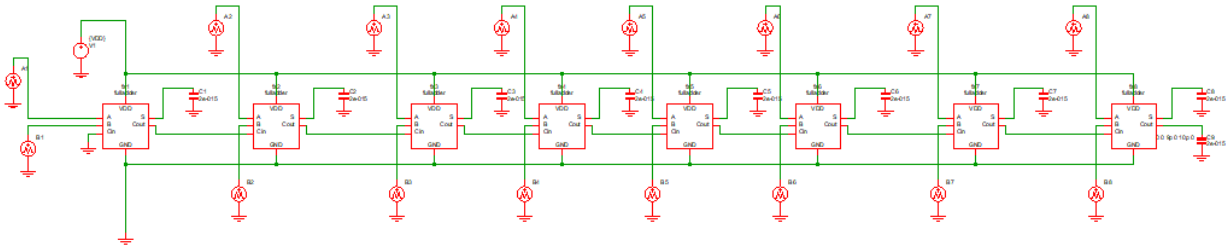
Tramite queste porte fondamentali è stato possibile sviluppare il modello di un HALF ADDER:



E quindi di un FULL ADDER, ad un singolo bit:



Successivamente si è deciso di simulare le precedenti strutture per confermarne il corretto funzionamento (simulazioni omesse per brevità), si è deciso di realizzare un FULL ADDER a 8 bit.



Tale circuito è sufficientemente complesso per analizzare alcuni caratteri generali dei circuiti digitali al variare di determinati parametri come ad esempio la tensione di alimentazione (VDD) e/o la pendenza di sottosoglia del transistore (variando il fattore di idealità n), che con circuiti più semplici sarebbero difficili da interpretare.

L'idea di base è simulare una serie di somme, ed analizzare l'energia consumata del dispositivo per operazione, al fine di ottenere un grafico al variare della tensione di alimentazione del circuito. Per fare ciò è necessario seguire un certo iter:

1. Volendo simulare 6 somme (cioè far svolgere al circuito 6 operazioni) è necessario definire la lunghezza temporale (il periodo di clock del sistema) per la quale i dati di input sono mantenuti stabili. Per fare ciò bisogna misurare il tempo di propagazione all'uscita dell'ultimo full-adder (sia per Carry-out e somma, poi trovare il massimo) al 90% del massimo valore del segnale ottenibile (VDD). Questo lo si fa facendo eseguire a sua volta una somma e misurando la propagazione della tensione di uscita.
2. Trovato il tempo di propagazione del circuito bisogna dimensionare gli ingressi (PWL source) associando il valore dello specifico ingresso ad un intervallo equivalente al tempo di propagazione raddoppiato misurato precedentemente (raddoppiato per permettere una migliore leggibilità dei risultati).
3. Eseguire la simulazione per le diverse tensioni di alimentazione.
4. Estrapolare la corrente assorbita dall'alimentazione moltiplicandola con la relativa tensione di alimentazione per ottenere la potenza (istantanea).

5. Integrare la potenza ottenendo l'energia consumata dal circuito per eseguire le 6 operazioni.
6. Dividere il valore massimo dell'energia per il numero di operazioni (in questo caso 6).

5.2 Simulazione della curva E - VDD

[8bitFAEVDD]

5.2.1 Parametri di simulazione

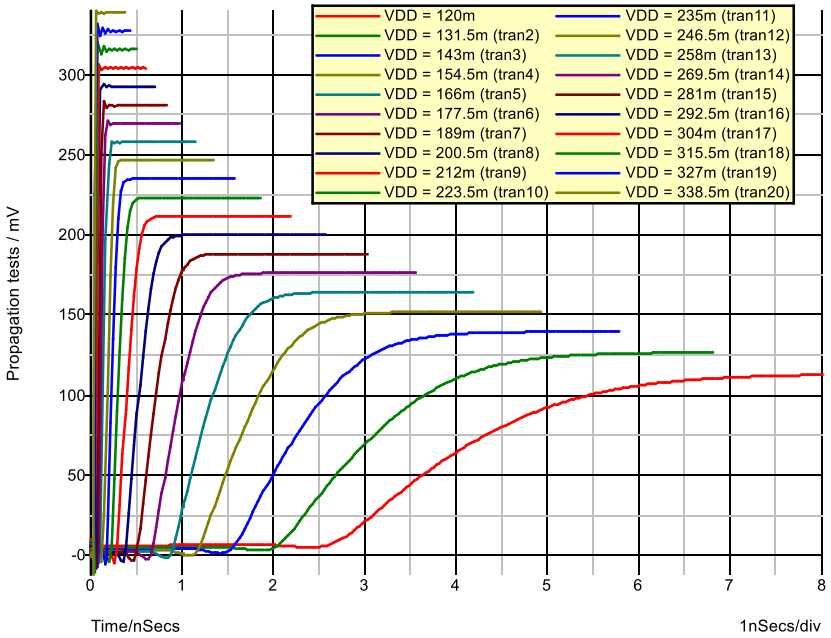
- VDD = 1
- stdCap = 2f
- minstep = 30p
- timeoffset = 5p
- MAXpropsim = 8n
- maxVDD = 0.35
- minVDD = 0.12
- NVDD = 20

5.2.2 Parametri del modello

T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,3	-0,3	1.3	0,7	1,23e5	0,02	0,2	60

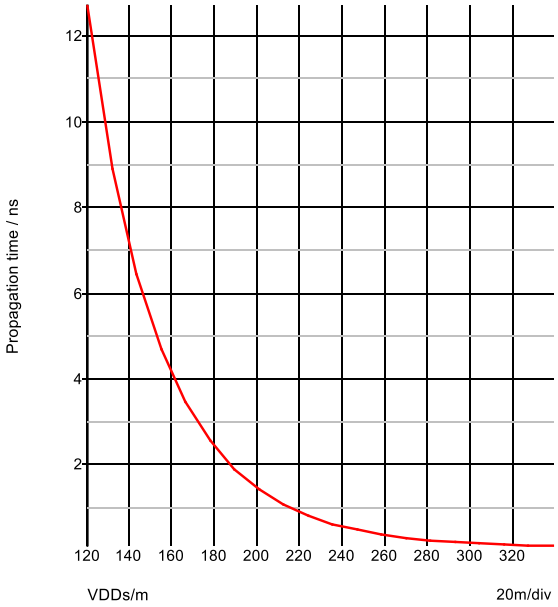
5.2.3 Circuito e simulazione

Il seguente grafico mostra il segnale di uscita del FULL ADDER al variare della tensione di alimentazione. Si vede molto bene come il tempo di propagazione diminuisca sensibilmente all'aumentare della tensione di alimentazione.

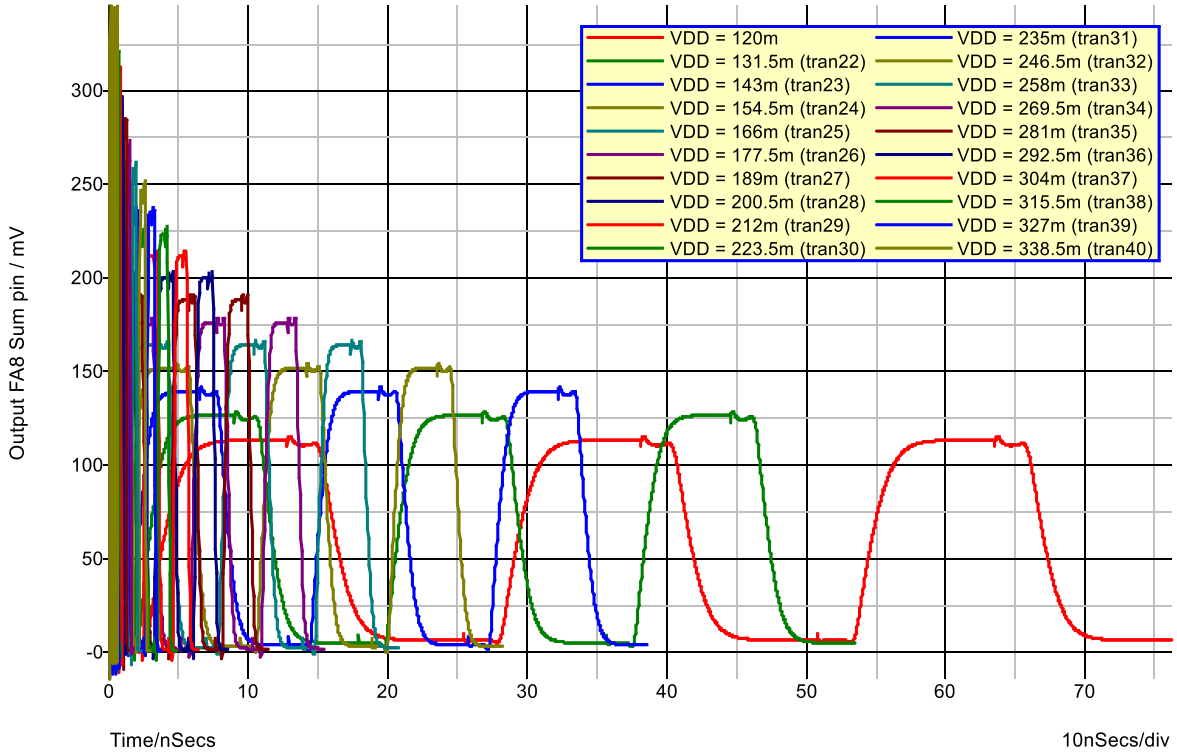


Si noti che, per ridurre il tempo totale di simulazione in questa fase, è stato creato un algoritmo di riduzione del tempo di simulazione adattivo rispetto alla tensione di alimentazione prevista. Per questo motivo le caratteristiche a VDD maggiore, una volta raggiunto il valore "alto", terminano e non proseguono fino al termine definito dalla caratteristica a tensione di alimentazione più bassa.

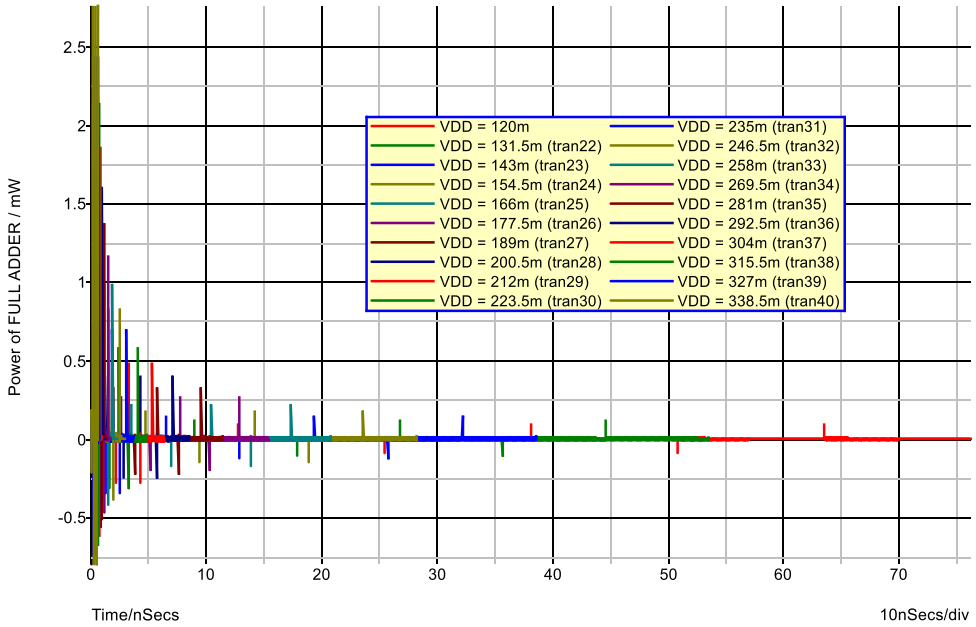
Il valore del tempo di propagazione (elaborazione delle curve precedenti) è riportato nel seguente grafico al variare della tensione di alimentazione:

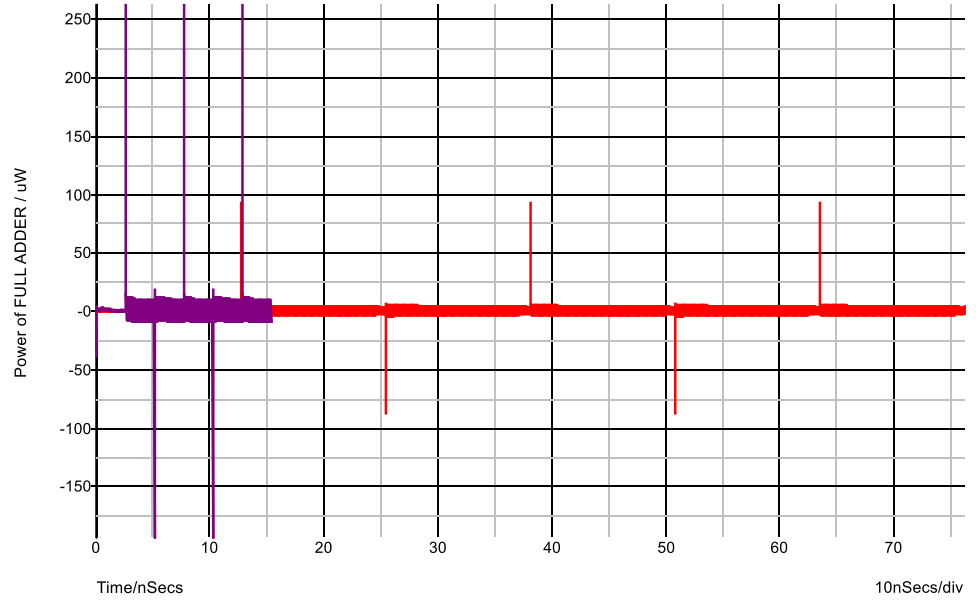


Successivamente sono state simulate le 6 somme, il seguente grafico mostra la forma d'onda dell'uscita del FULL ADDER (in questo caso il pin Sum):



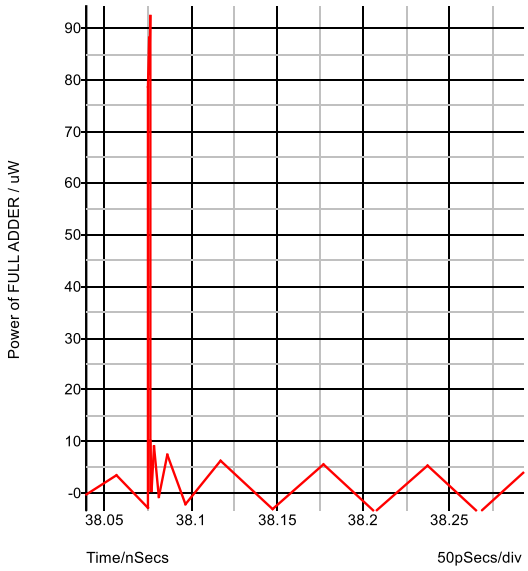
La curva rossa a VDD minore mostra molto bene il transitorio a cui è soggetto il circuito, una volta che l'input è cambiato. Questo grafico enfatizza il concetto di propagazione visto nei grafici precedenti e mostra l'enorme differenza che c'è tra la forma d'onda a 120 mV (curva rossa) e quella a 338 mV (curva oro) nonostante ci sia una differenza di poco più di 200mV tra una tensione di alimentazione e l'altra. La differenza in termini prestazionali è abissale ma, come si vedrà in seguito, bisogna tenere conto del trade-off tra consumi e prestazioni.





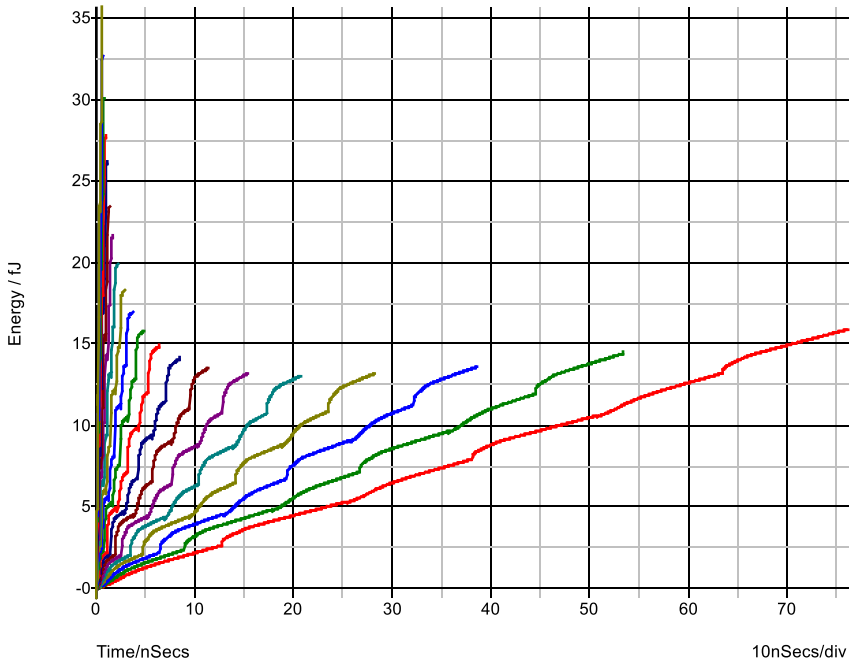
I due precedenti grafici mostrano l'andamento globale (il primo) e quello per sole 2 alimentazioni VDD = 120mV e VDD = 177mV (il secondo) della potenza assorbita durante lo svolgimento delle 6 somme.

Si possono apprezzare gli spikes di corrente, e quindi di potenza, nell'istante di commutazione degli ingressi. Tali picchi sono dovuti alla commutazione dei transistor all'interno del dispositivo e determinano l'assorbimento di energia dinamica. Successivamente ad un picco, il profilo di potenza si abbassa molto velocemente, ma non si azzera mai. Questo vuol dire che esiste sempre un leakage statico che determina un assorbimento continuo di energia, seppur piccolo.

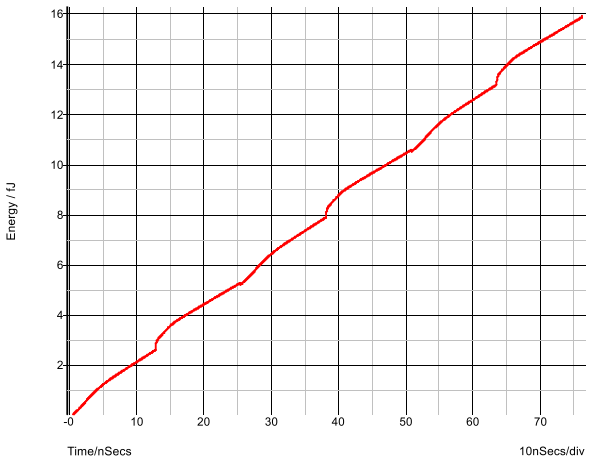


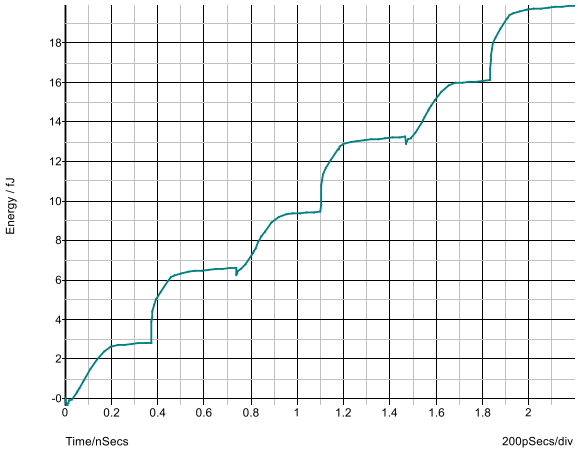
Da questo ingrandimento della caratteristica sembra che il circuito abbia un comportamento oscillatorio, ma non è così. Tale effetto è frutto dell'approssimazione temporale del simulatore. Diminuendo il valore del massimo timestep è possibile ottenere una approssimazione molto migliore. Tuttavia non è necessario ottenere una precisione altissima per tale dato in quanto successivamente, integrando su tutto il tempo di elaborazione per trovare l'energia assorbita, il valore che ci interessa di tale caratteristica è la media, quindi il comportamento informativo oscillatorio viene perso siccome il periodo di oscillazione è molto minore all'intervallo di integrazione.

L'integrale della potenza è dunque l'energia assorbita:

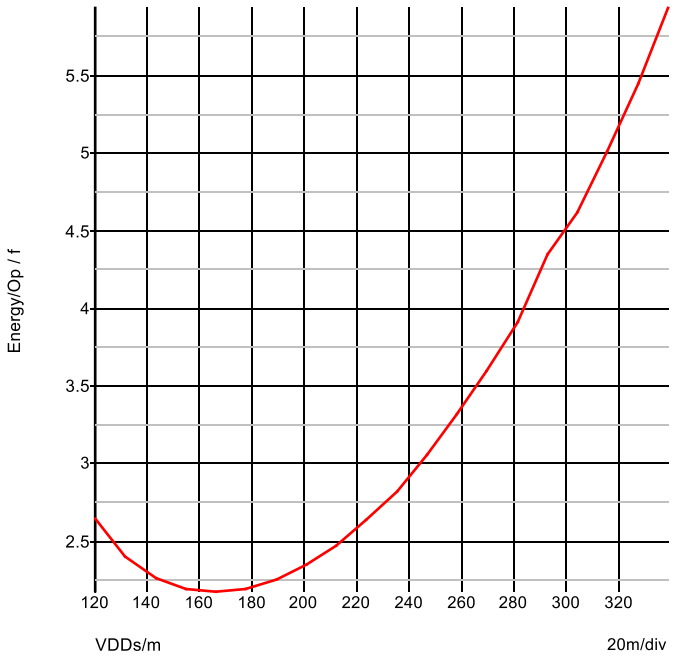


In tale grafico si possono distinguere due tipologie di curve: le curve a bassa tensione di alimentazione in cui il termine di leakage è dominante, e quelle a tensione di alimentazione alta in cui l'energia dinamica è il termine più importante.





A sinistra vediamo la curva a VDD=120mV in cui il gradino dovuto all'assorbimento dinamico è quasi inesistente a discapito di un consumo statico che fa da padrone. Nel grafico di destra (VDD = 258mV) possiamo osservare come la pendenza dovuta al leakage sia ininfluente nel computo totale energetico in confronto all'assorbimento dinamico. Si può anche notare che nel grafico di destra è possibile distinguere due tipologie di gradini, uno più accentuato dell'altro, che corrispondono di fatto alla somma di 0xFF + 0xFF e 0x00 + 0x00. Può essere quindi fatta una analogia tra assorbimento di corrente e tipologia di somma da eseguire. Già nel grafico principale dell'energia è possibile notare che esiste una sella nell'energia massima assunta dal dispositivo in relazione alla tensione di alimentazione, e quindi una energia ottima. Visualizzando l'energia massima in relazione alla VDD si ottiene infatti:



Come si può facilmente notare esiste un minimo per l'energia/operazione che corrisponde ad una tensione di alimentazione di circa 170 mV. A sinistra di tale punto l'energia cresce siccome i tempi di elaborazione si dilatano e il leakage diventa gradualmente sempre più significativo. A destra del minimo il leakage risulta essere meno importante dell'assorbimento dinamico.

5.3 Simulazione di curve E – VDD variando S

[8bitFAEVDDlowerS]

5.3.1 Parametri di simulazione

- VDD = 1
- stdCap = 2f
- maxstep = 30p
- timeoffset = 5p
- MAXpropsim = 22n
- maxVDD = 0.28
- minVDD = 0.12
- NVDD = 20
- NEnergy = 3
- n_param = [1.3, 1.2, 1.1]

5.3.2 Parametri del modello

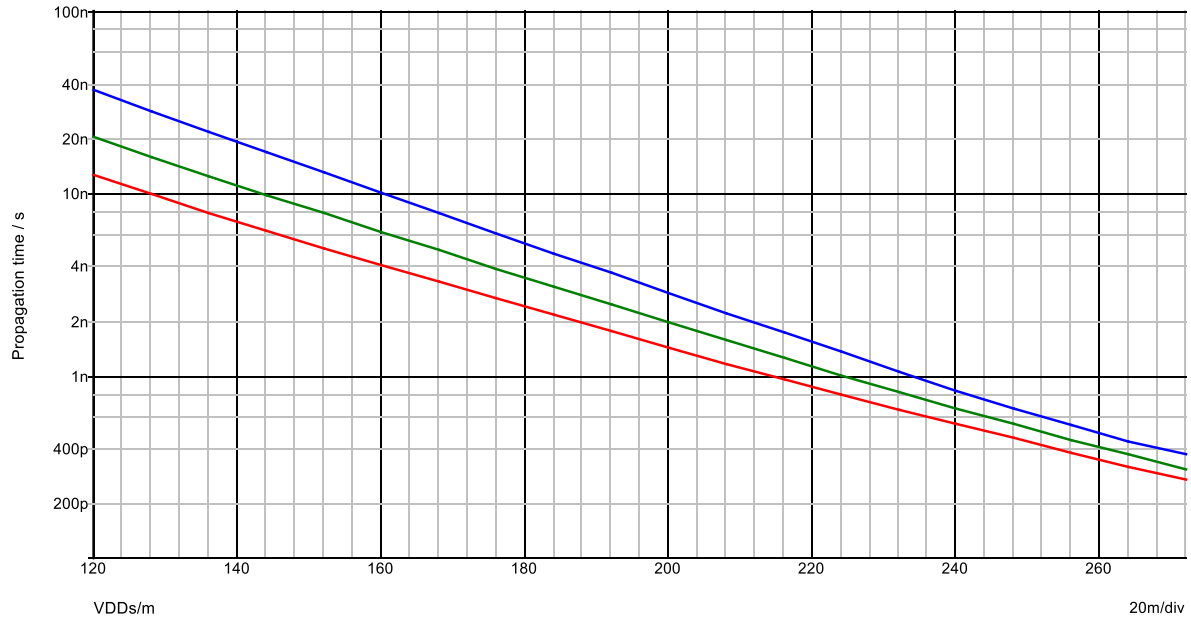
T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,3	-0,3	1.3	0,7	1,23e5	0,02	0,2	60

5.3.3 Circuito e simulazione

Successivamente si è voluto ottenere tale grafico al variare di alcuni parametri. In particolare, si è deciso di variare solamente la Subthreshold Swing (S) variando il fattore di idealità n in tre step:

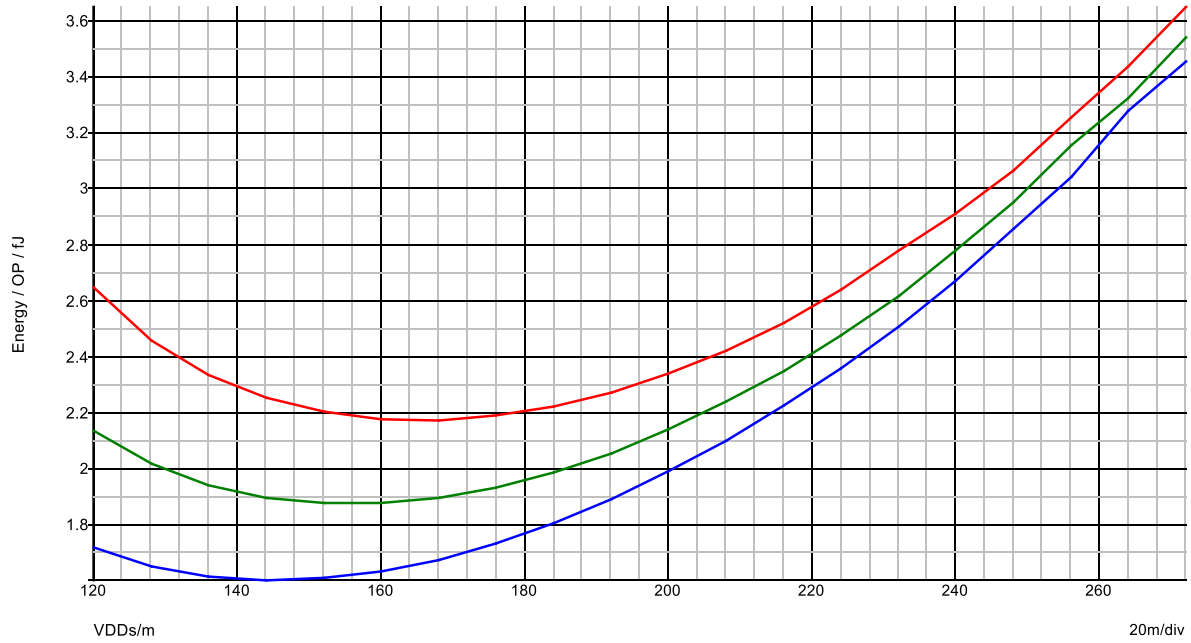
- $n = 1.3 \rightarrow S = n\,\phi_t \ln(10) \cong 0.078 \frac{mV}{dec}$ (curve rosse)
- $n = 1.2 \rightarrow S = n\,\phi_t \ln(10) \cong 0.072 \frac{mV}{dec}$ (curve verdi)
- $n = 1.1 \rightarrow S = n\,\phi_t \ln(10) \cong 0.066 \frac{mV}{dec}$ (curve blu)

L'impatto di tale valore sui parametri dinamici del circuito è importante, facendo un confronto tra i tempi di propagazione si nota molto bene la distinzione tra le caratteristiche (tempo di propagazione in scala log):

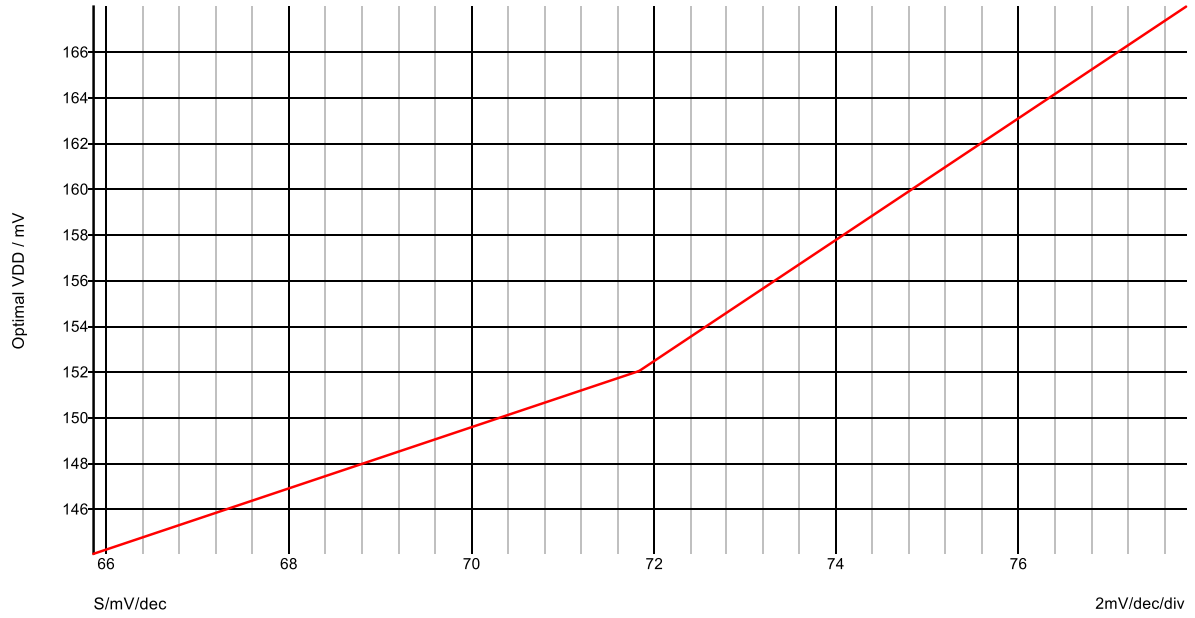


Si noti come aumentando S, il tempo di propagazione diminuisca. La curva rossa (S = 66 mV/dec) è circa 4 volte inferiore, per VDD = 120mV, rispetto alla curva blu (S = 78mV/dec). Inoltre, si vuole far notare che le caratteristiche non sono parallele e tendono ad avvicinarsi al crescere della tensione di alimentazione.

Analizzando quindi il grafico dell'energia per operazione si nota che diminuendo la subthreshold swing S verso il minimo teorico di 60 mV/dec si ottiene un minore dispendio energetico globale. Tale dato è in accordo con l'analisi dei tempi di propagazione riportata sopra: diminuendo S l'andamento dei tempi di propagazione sale e quindi l'energia necessaria a svolgere una operazione (alla medesima VDD) diminuisce.



Un ulteriore dato è possibile estrapolare da tale grafico: al variare di S, il minimo dell'energia per operazione (cioè l'energia ottimale per operazione) cambia valore in ascissa. Cioè varia la tensione di alimentazione al quale ho il minimo dispendio di energia. L'andamento del minimo è mostrato nel seguente grafico in relazione all'avanzamento di S:



Per descrivere l'andamento in termini qualitativi è possibile avvalersi delle soluzioni viste nel capitolo 2:

La corrente off del transistore (Vds = VDD e Vgs = 0V), con VDD < Vt quindi sottosoglia, è secondo la 2.20:

$$I_{OFF} \cong W I_0 10^{-\frac{V_{t0}}{S}} 10^{\frac{\lambda_{DS} VDD}{S}}$$

Mentre la corrente ON, sempre sottosoglia, secondo la 2.23:

$$I_{ON} \cong W I_0 10^{-\frac{VDD - V_{t0}}{S}} 10^{\frac{\lambda_{DS} VDD}{S}}$$

Per cui il rapporto tra le due correnti è:

$$\frac{I_{ON}}{I_{OFF}} \cong 10^{\frac{VDD}{S}}$$

Aumentando S il rapporto tra la corrente ON e quella OFF diminuisce, enfatizzando quindi la componente di leakage che, come si vede dalle simulazioni, produce uno spostamento verso valori maggiori della tensione per energia ottimale.

5.4 Simulazione di curve E-VDD con Ioff costante

[8bitFAEVDDsameIoff]

5.4.1 Parametri di simulazione

- VDD = 1
- stdCap = 2f
- maxstep = 50p
- timeoffset = 5p

- MAXpropsim = 7n
- maxVDD = 0.28
- minVDD = 0.12
- NVDD = 20
- Vt0_param = [0.3, 0.25, 1.2]

5.4.2 Parametri del modello

T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,3	-0,3	1.3	0,7	1,23e5	0,02	0,2	60

5.4.3 Circuito e simulazione

Successivamente si è deciso di far variare il parametro S, come nel caso precedente, mantenendo costante la Ioff. Per fare ciò, una volta fissato S è necessario far variare un ulteriore parametro che compone la Ioff per mantenerla costante.

Come abbiamo già detto la corrente off (Vgs = 0) è rappresentata dall'equazione 2.20:

$$I_{OFF} \cong W I_0 10^{-\frac{V_{t0}}{S}} 10^{\frac{\lambda_{DS} VDD}{S}}$$

Con

$$I_0 = \left(\frac{t_{ch}}{2 - t_{ch}}\right) C_{inv}\, n\, \phi_t$$

Volendo mantenere Ioff costante variando Vt0, bisogna invertire la 2.20:

$$V_{t0} = \lambda_{DS}\, VDD + S \log\left(\frac{W I_0}{I_{OFF}}\right)$$

Quindi per ogni VDD deve essere ricalcolato il parametro di Vt0 al fine di bilanciare l'effetto dovuto alla variazione di S. I risultati non sono stati riportati in quanto congruenti a quelli ottenuti alla sola riduzione di S.

5.5 Simulazione curve E – VDD variando la frequenza

[8bitFAEVDDmorefrequency]

5.5.1 Parametri di simulazione

- VDD = 1
- stdCap = 2f
- maxstep = 40p
- timeoffset = 5p
- MAXpropsim = 12n
- maxVDD = 0.28
- minVDD = 0.12
- NVDD = 20
- NEnergy = 3

- time_const = [3, 2, 1.2]

5.5.2 Parametri del modello

T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,3	-0,3	1.3	0,7	1,23e5	0,02	0,2	60

5.5.3 Circuito e simulazione

Riguardando l'equazione 4.20 si nota come l'energia dinamica sia dipendente da un fattore α_{sw}:

$$E_T = \alpha_{sw}\, C_{eff} VDD^2 + VDD \frac{I_{off}}{X_{stack}}\, LD_{stage} t_d + \frac{E_{reg} LD_{tot}}{LD_{stage}}$$

Tale fattore è denominato activity factor e prende in considerazione lo sfruttamento della logica da parte del sistema. Nelle precedenti analisi abbiamo simulato il massimo tempo di propagazione del full adder a 8 bit per ogni VDD: analizzando la forma d'onda in uscita, è stato possibile determinare (al 90% dell'escursione massima) il T_{prop}. Successivamente è stato necessario definire il ciclo di clock con il quale modificare gli ingressi del nostro circuito. Per le precedenti analisi abbiamo supposto:

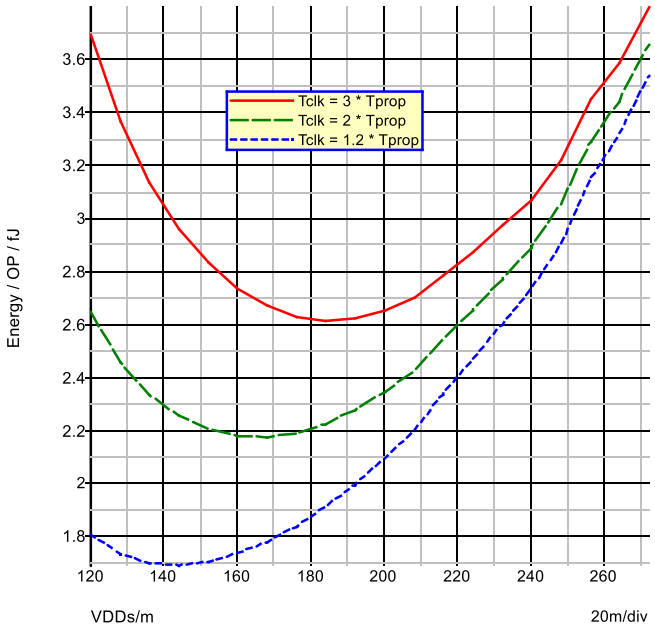
$$T_{clk} = 2 \times T_{prop}$$

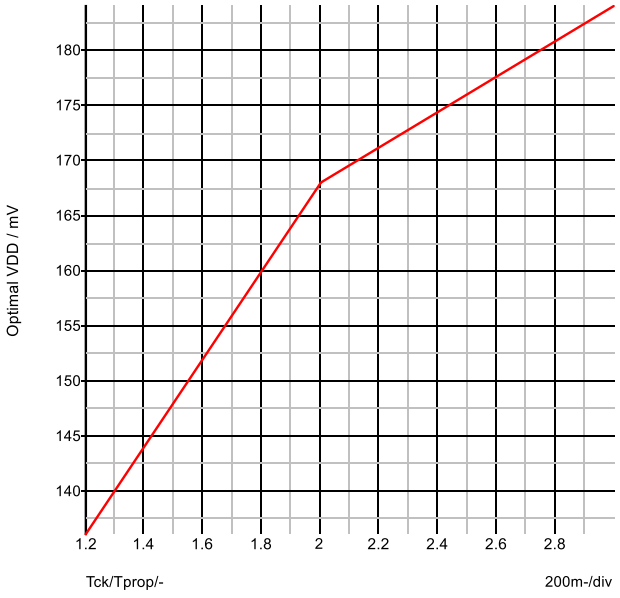
Ora se viene modificato il fattore di proporzionalità (2 in questo caso) è possibile modificare la frequenza del nostro circuito, quindi il fattore di sfruttamento dello stesso.

Successivamente è stato deciso di analizzare il circuito ponendo tale fattore di proporzionalità (da adesso in avanti chiamato α) pari a 3, 2 e 1,2. Il limite minimo è 1 per il quale il circuito lavora alla frequenza massima, oltre la quale l'uscita non riesce a raggiungere il 90% dell'escursione totale. Si trova quindi che:

$$T_{clk} = \alpha \times T_{prop} \rightarrow \alpha = \frac{T_{clk}}{T_{prop}}$$

Le simulazioni prendono seguono l'andamento previsto, in quanto all'aumentare di α il leakage aumenta, spostando verso tensioni più alte la VDD ottima:


Energy / OP / fJ
3.6
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
Tclk = 3 * Tprop
Tclk = 2 * Tprop
Tclk = 1.2 * Tprop
120
140
160
180
200
220
240
260
VDDs/m
20m/div


Optimal VDD / mV
180
175
170
165
160
155
150
145
140
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
Tck/Tprop/-
200m-/div

6 Tecnica di riduzione energetica: Stacking naturali

[naturalstacking]

6.1 Parametri di simulazione

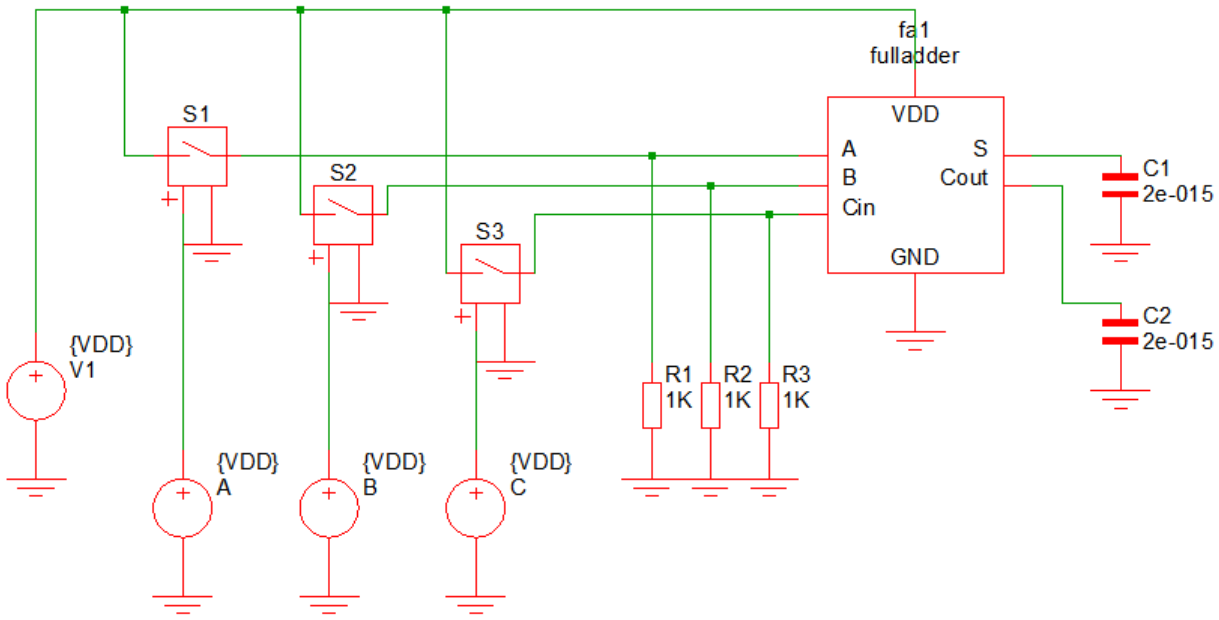
- stdCap = 2f
- conf = ['000','001','010','011','100','101','110','111']
- maxVDD = 1.2
- minVDD = 0.1
- lVDDs = [0.2, 0.8, 1.2]
- DIBLs = [0.2, 0.15, 0.1]

6.2 Parametri del modello

T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,3	-0,3	1.3	0,7	1,23e5	0,02	0,2	60

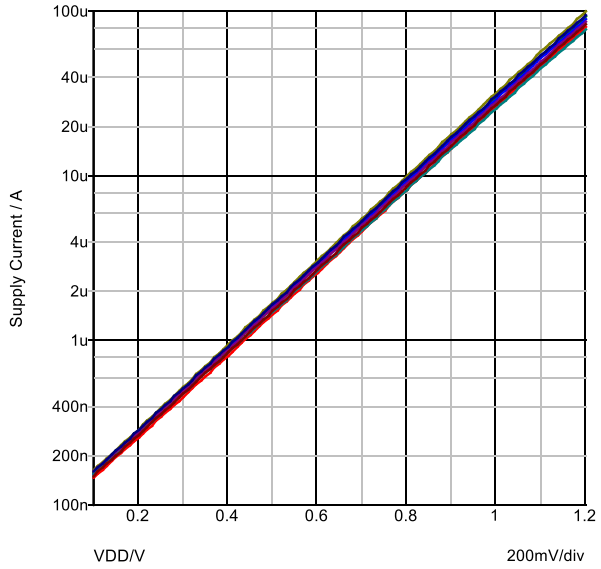
6.3 Circuito e simulazione

Una semplice tecnica per ridurre l'assorbimento di corrente da parte di un circuito è il così detto natural stacking. Si basa sulla variazione degli ingressi di un circuito se questo è in stato IDLE, cioè non interessa l'uscita del circuito stesso, per raggiungere una configurazione con il più basso stato energetico. Questa tecnica infatti è utile per sfruttare gli stacking naturali insiti nella topologia dei circuiti, senza dover aggiungere stadi di stacking artificiosamente. È particolarmente efficace per circuiti di medie dimensioni: in queste simulazioni prenderemo in considerazione un Full Adder.

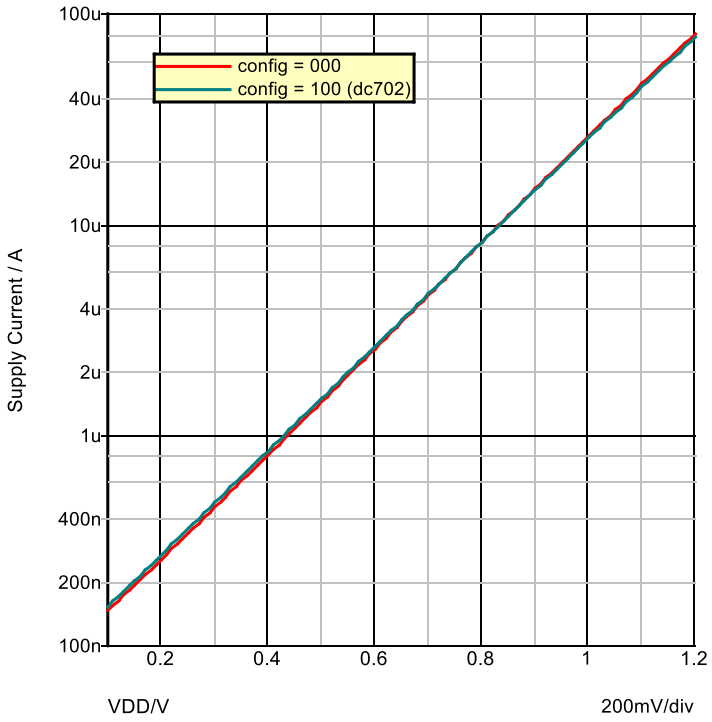


In particolare variando staticamente A, B, C e misurando quindi la corrente assorbita dal full adder sarà possibile determinare la configurazione a minore energia. La topologia del circuito di pilotaggio degli ingressi è complessa per via di un bug del simulatore Simetrix: è stato riscontrato che si verificano dei problemi durante

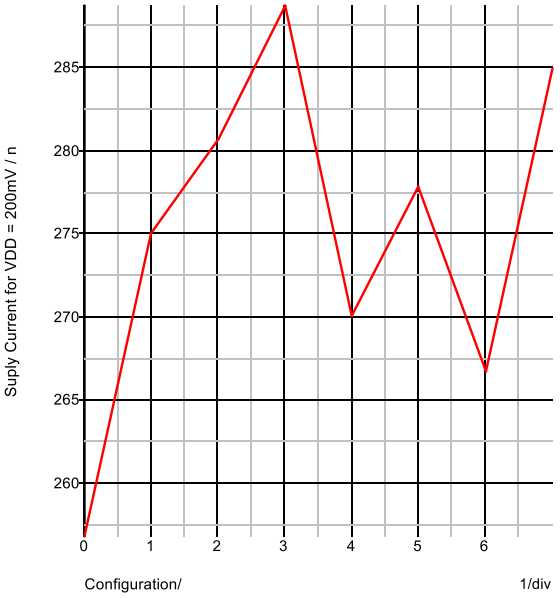
la simulazione DC (non vengono forniti dei risultati congrui alle aspettative) dovuti alla variazione di un parametro definito come GLOBALE, che in questo caso è la tensione di alimentazione VDD. Per ovviare a tale problematica si è deciso di pilotare gli ingressi con degli switch praticamente ideali, ed inserire dei pull down che portino la tensione agli ingressi a 0 quando questi sono spenti. La corrente circolante sulle resistenza non rientra nel computo totale della corrente assorbita dal Full-Adder.

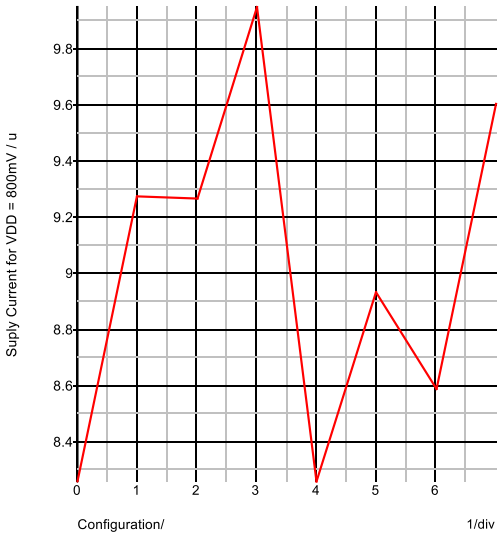


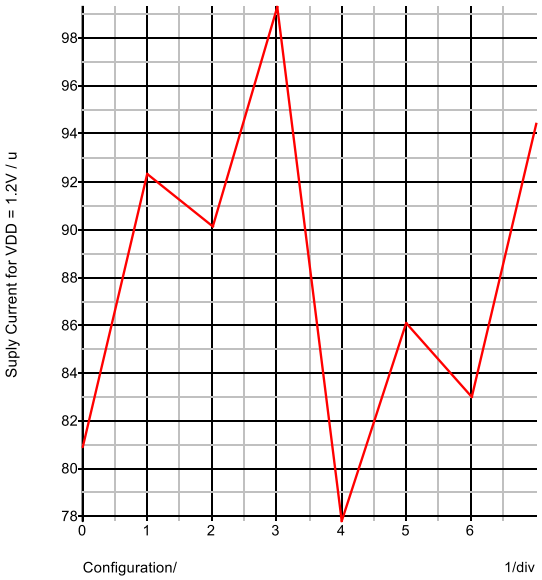
Se analizziamo l'assorbimento di corrente in base alla tensione di alimentazione applicata per tutte le configurazioni otteniamo un grafico difficile da leggere. Selezionando invece solo la configurazione 000 e quella 100 (A = 1 , B = 0, Cin = 0) è possibile notare un comportamento particolare:



Per tensioni basse configurazione 000 è quella più efficiente, mentre per tensioni più alte, da 0.8V in su, la corrente assorbita per lo stato 100 è minore rispetto a quella assorbita dallo stato 000. Tali grafici sono stati realizzati con $\lambda_{DS} = 0.2$

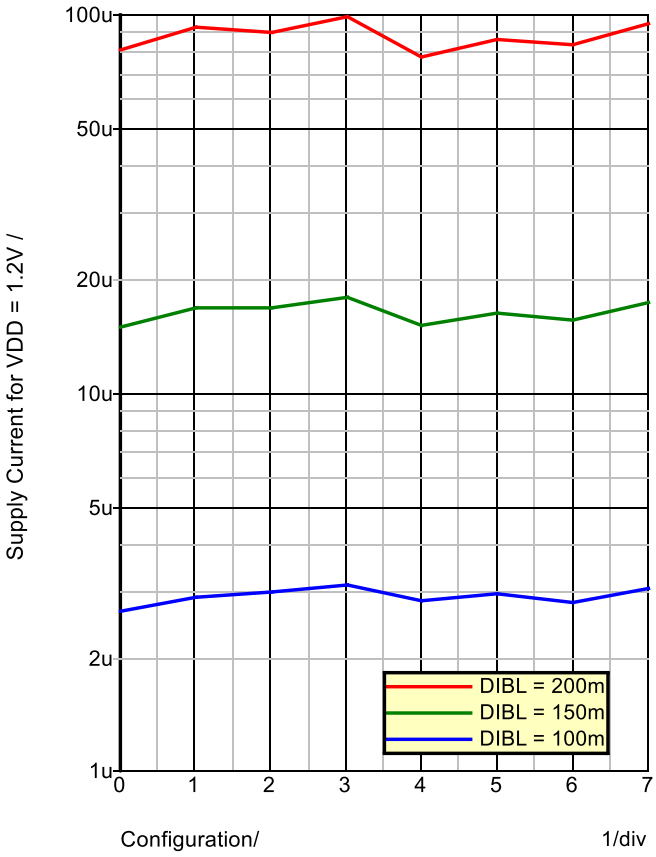






Come si vede dai grafici precedenti, per la tensione di alimentazione di 1.2 V, c'è un minimo per la configurazione 4 (=100, A=1, B=Cin=0), che per tensioni più basse risulta essere nella configurazione 0 (=000).

Tale modifica del minimo può essere imputata all'effetto del DIBL esercitato dalla tensione Vds sul profilo di potenziale interno al canale. Infatti se andiamo a ridurre tale effetto otteniamo una riduzione complessiva della corrente ed un ulteriore abbassamento della corrente per lo stato 000 che già a $\lambda_{DS} = 0.1$ diventa il minimo locale.


100u
50u
20u
10u
5u
2u
1u
Supply Current for VDD = 1.2V /
0
1
2
3
4
5
6
7
Configuration/
1/div
DIBL = 200m
DIBL = 150m
DIBL = 100m

7 Tecnica di riduzione energetica: Power gating

7.1 Simulazione statica

[powergatingstatic]

7.1.1 Parametri di simulazione

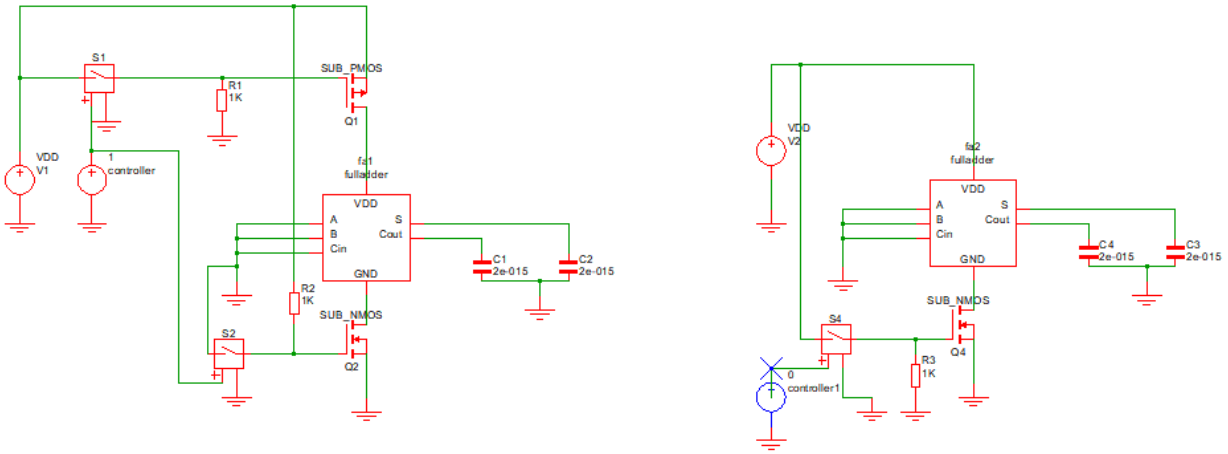
- VDD = 1
- stdCap = 2f
- maxVDD = 1.2
- minVDD = 0.1

7.1.2 Parametri del modello

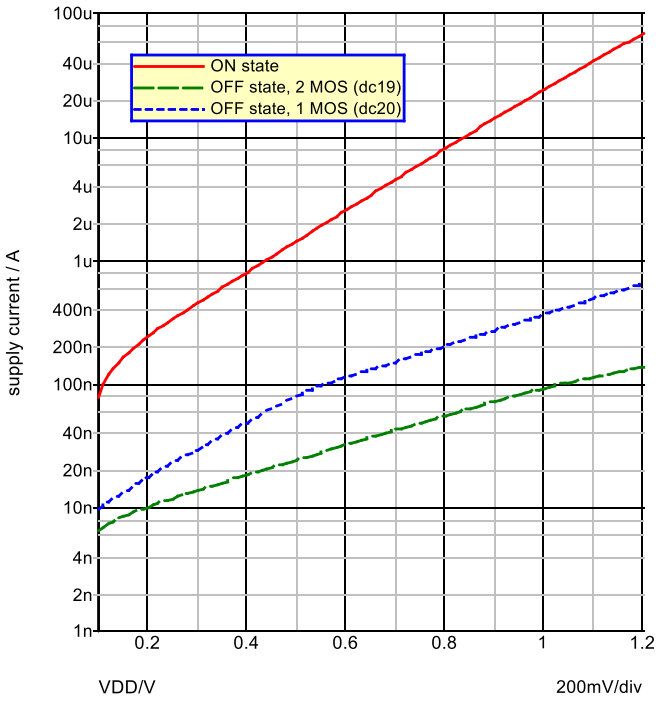
T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,3	-0,3	1.3	0,7	1,23e5	0,02	0,2	60

7.1.3 Circuito e simulazione

Il power gating permette di abilitare/disabilitare una determinata porzione di circuito per porla in SLEEP Mode. Il power gating è generalmente formato da uno o due mosfet che bloccano il flusso di corrente quando il segnale di SLEEP è attivo. La simulazione si avvale di un doppio circuito che ha come elemento di test un Full Adder. Tale circuito viene utilizzato per simulare due topologie di power gating: a singolo transistore (destra) e a doppio transistore (sinistra). Il dimensionamento dei transistor è un parametro di progetto fondamentale, che però in questo ambito è stato trascurato per semplicità rendendo i power gating transistors uguali ai transistor che compongono il Full Adder.



Il grafico successivo mostra l'assorbimento di corrente statica nello stato ON (cioè durante il normale funzionamento del dispositivo), nello stato OFF nella configurazione a 2 transistori e nello stato OFF nella configurazione a 1 transistore.



Come si può notare l'assorbimento di corrente cala drasticamente sfruttando il power gating, inoltre l'utilizzo della soluzione a doppio transistore è sicuramente più performante dal punto di vista energetico, ma quella a singolo NMOS può rappresentare un buon trade-off tra la soluzione con e senza power gating in quanto l'NMOS e/o il PMOS di power gating sono dimensionalmente più onerosi dei transistor che compongono la topologia del circuito per via dell'elevato flusso di corrente e delle prestazioni dinamiche richieste.

7.2 Simulazione dinamica

[powergatingtransition]

7.2.1 Parametri di simulazione

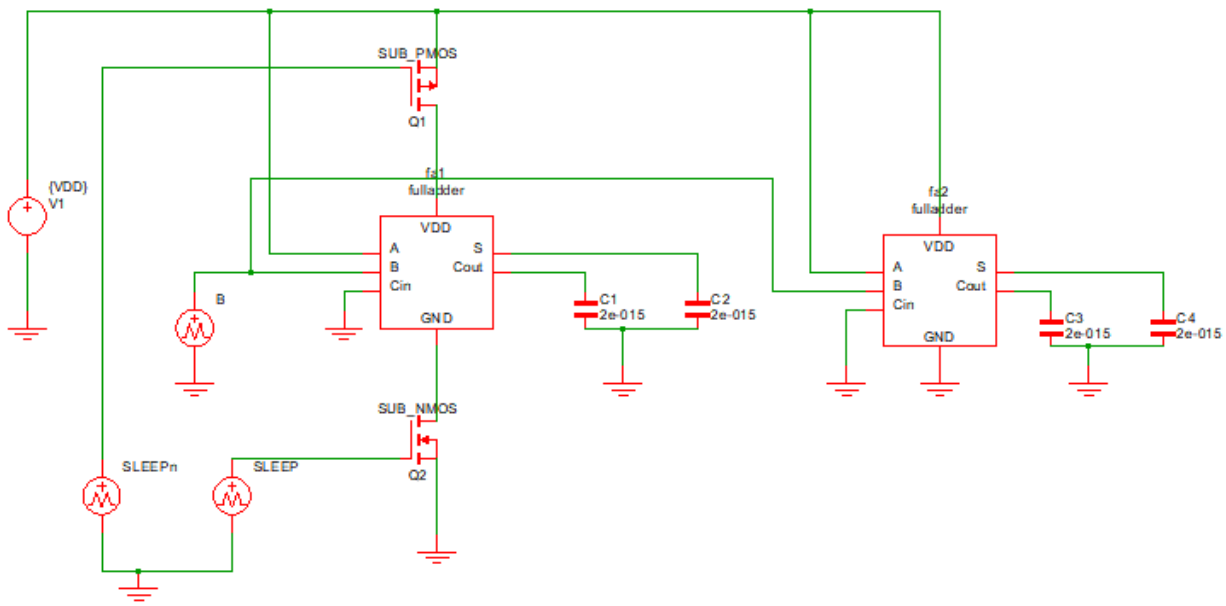
- VDD = 1
- stdCap = 2f
- VDDs = [1, 0.5, 0.3, 0.25]
- timesim = 450p

7.2.2 Parametri del Modello

T [K]	W [m]	L [m]	Vt0n [V]	Vt0p [V]	n	tch	Term_V [m/s]	Cinv [F/m^2]	Kds	rpara [Ω]
300	1e-6	30e-9	0,3	-0,3	1.3	0,7	1,23e5	0,02	0,2	60

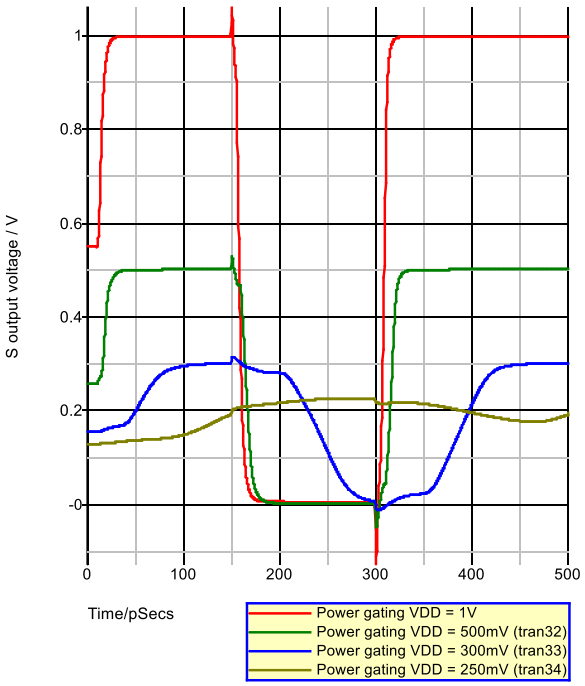
7.2.3 Circuito e simulazione

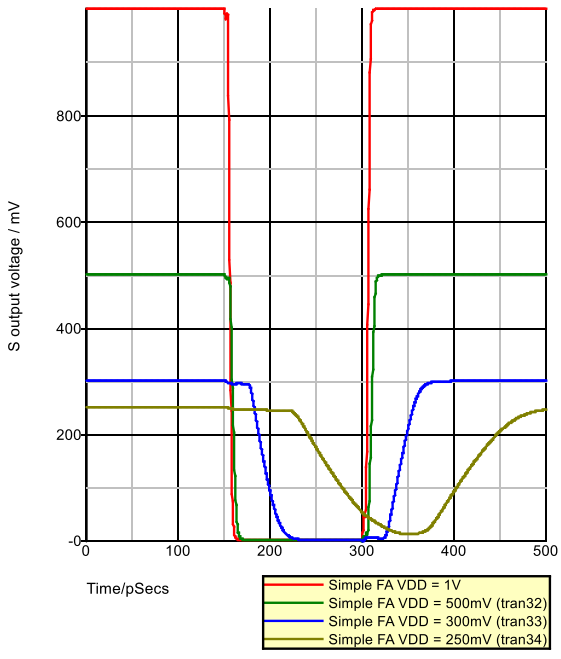
Per valutare le prestazioni dinamiche del Full Adder pilotato da Power Gating Transistors è stato necessario realizzare un circuito con la seguente topologia:



Come si nota sono presenti due Full Adder indipendenti tra loro e pilotati dagli stessi segnali di ingresso. Il primo (sinistra) è dotato di Power Gating Transistor che inizialmente disabilitano il circuito, mentre il secondo Full Adder (destra) è alimentato direttamente.

Il seguente grafico mostra la risposta del Full Adder con Power Gating a diverse tensioni di alimentazione (sinistra) e la risposta del Full Adder con alimentazione diretta (destra).



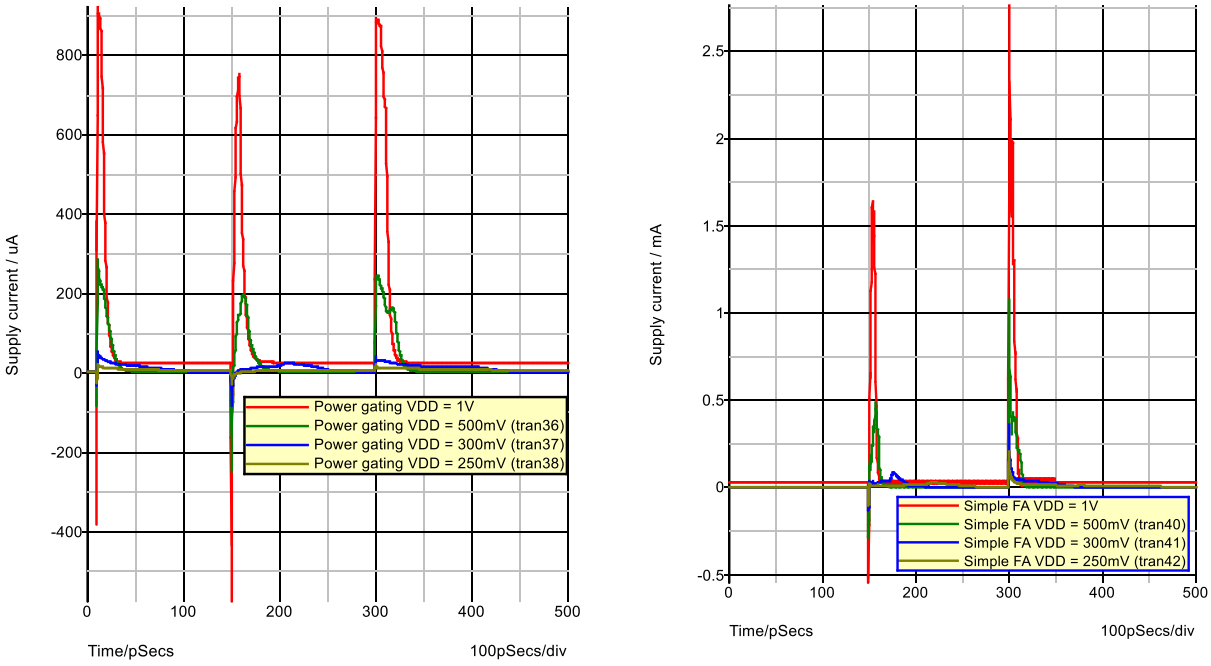


Come si vede il Power Gating, oltre a peggiorare le prestazioni dinamiche del circuito, peggiora anche la forma d'onda d'uscita che si presenta "rumorosa" per tensioni di alimentazione basse. Da questi risultati è possibile affermare che la tecnica del Power Gating (con transistori di Gating "piccoli" come quelli utilizzati in questa

simulazione) è sconsigliabile per tensioni al di sotto della tensione di soglia, quindi applicazioni very low voltage. È tuttavia da considerare il dimensionamento dei transistori di Gating: infatti maggiore è l'area impiegata migliori saranno le prestazioni dinamiche del sistema. Cioè ingrandendo i transistori di Gating questi diventeranno sempre più trasparenti al sistema pilotato e non produrranno più un effetto deleterio come quello mostrato in figura.

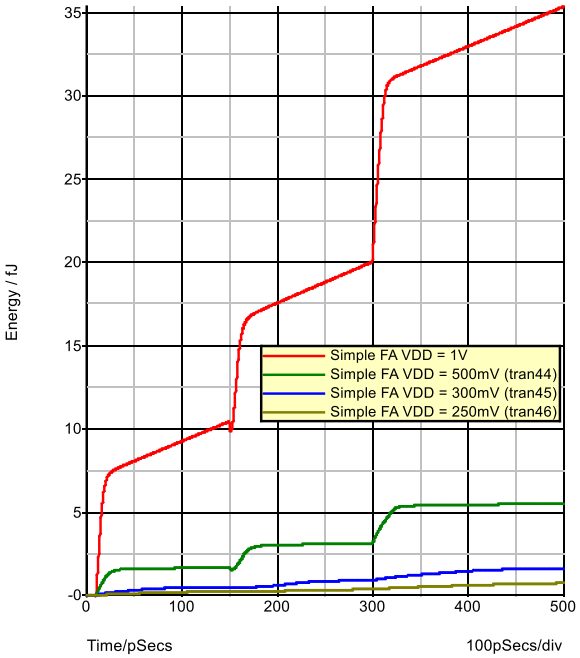
Notare come nel grafico di sinistra (Power Gating) il segnale, da uno stato intermedio, passi ad "alto". Questo comportamento deriva dall'abilitazione dei due transistori di Gating che, a 10 ps dall'inizio della simulazione, simulano l'accensione del dispositivo. Il resto della forma d'onda è dovuto alla variazione del segnale di ingresso al pin B.

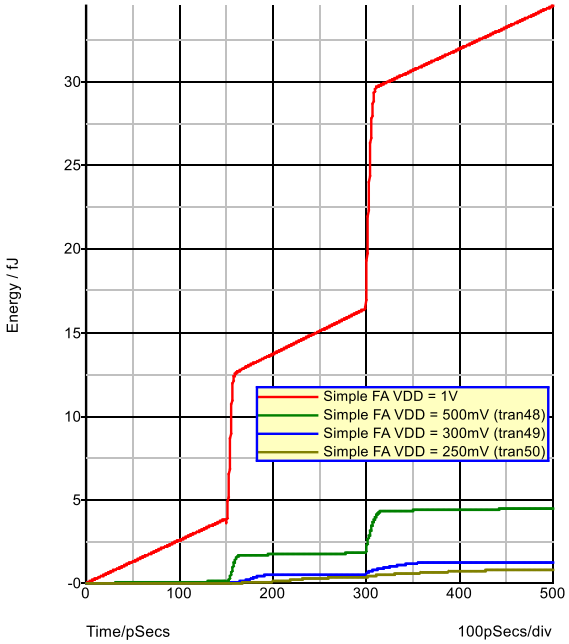
Allo stesso modo possiamo confrontare la risposta dinamica per la corrente assorbita:



La presenza dei tre picchi nel grafico del Power Gating (differentemente da quello ad alimentazione diretta) è dovuta al fatto che il primo picco di corrente è un effetto dell'abilitazione dei transistori di Gating, mentre gli ultimi due sono effetti dovuti alla commutazione degli ingressi, come lo sono gli unici due del grafico di destra. I picchi di corrente per effetto del Power Gating sono ridotti rispetto a quelli del circuito con alimentazione tradizionale, questo spiega il trade-off tra tempo di propagazione e consumo energetico: maggiore è l'assorbimento di corrente e migliori saranno le performance dinamiche.

Analogamente come è stato fatto per altre simulazioni è possibile ottenere, integrando la potenza assorbita, l'energia assorbita dal circuito di Power Gating (sinistra) ed alimentazione tradizionale (destra).





Si nota immediatamente come nel computo totale dell'energia i due circuiti non differiscano di molto, ma se confrontati attentamente è possibile notare una differenza sostanziale: l'abilitazione del circuito, nel caso di sinistra, produce un consumo di energia paragonabile a quello di una operazione tradizionale. Benché quindi, nel caso del Power Gating, per ogni operazione venga assorbita una quantità di energia minore rispetto al caso tradizionale, l'abilitazione del circuito stesso produce un assorbimento che potrebbe andare a vanificare l'efficacia della tecnica stessa. Si consiglia quindi di utilizzare periodi di inattività ed attività dei circuiti "lunghi" con poche transizioni dallo stato ON allo stato SLEEP, altrimenti l'energia dinamica necessaria a rendere attivo il circuito potrebbe superare la quota energetica risparmiata dal Gating.

8 Sommario

1 Introduzione 3
1.1 Il modello 3
1.2 Progetto di simulazione 4
1.2.1 I file schematici 4
1.2.2 Gli script file 5
1.2.3 Nota 6
2 Capacità di gate 7
2.1 Parametri di simulazione 7
2.2 Parametri del modello 7
2.3 Circuito e simulazione 7
3 Analisi delle caratteristiche di una porta NOT 9
3.1 Parametri di simulazione 9
3.2 Parametri del modello 9
3.3 Circuito e simulazione 9
4 Metodo di riduzione dell'energia: stacking 14
4.1 Parametri di simulazione 14
4.2 Parametri del modello 14
4.3 Circuito e simulazione 14
5 Analisi energetica di un Full Adder 20
5.1 Circuito e simulazione 20
5.2 Simulazione della curva E - VDD 23
5.2.1 Parametri di simulazione 23
5.2.2 Parametri del modello 23
5.2.3 Circuito e simulazione 23
5.3 Simulazione di curve E – VDD variando S 29
5.3.1 Parametri di simulazione 29
5.3.2 Parametri del modello 29
5.3.3 Circuito e simulazione 29
5.4 Simulazione di curve E-VDD con Ioff costante 31
5.4.1 Parametri di simulazione 31
5.4.2 Parametri del modello 32
5.4.3 Circuito e simulazione 32
5.5 Simulazione curve E – VDD variando la frequenza 32

5.5.1 Parametri di simulazione 32
5.5.2 Parametri del modello 33
5.5.3 Circuito e simulazione 33
6 Tecnica di riduzione energetica: Stacking naturali 35
6.1 Parametri di simulazione 35
6.2 Parametri del modello 35
6.3 Circuito e simulazione 35
7 Tecnica di riduzione energetica: Power gating 39
7.1 Simulazione statica 39
7.1.1 Parametri di simulazione 39
7.1.2 Parametri del modello 39
7.1.3 Circuito e simulazione 39
7.2 Simulazione dinamica 40
7.2.1 Parametri di simulazione 40
7.2.2 Parametri del Modello 40
7.2.3 Circuito e simulazione 40