

**Politecnico di Milano**

Scuola di Ingegneria dell'Informazione  
Corso di Laurea Magistrale in Ingegneria Elettronica



Front-End a Basso Consumo  
per l'Acquisizione e la Digitalizzazione  
di Segnali Neuronalì.

Relatore: Prof. Andrea L. LACAITA

Correlatore: Ing. Andrea BONFANTI, PhD

Tesi di Laurea di: Stefano BRENNÀ

matricola 748959

Anno Accademico: 2011-2012



# Ringraziamenti

L'opportunità di questo lavoro si è presentata in un momento difficile e la devo soprattutto al Prof. Lacaita. Lui ha permesso che io prendessi parte a questo progetto e per questo vorrei ringraziarlo. In ogni caso, senza la guida dell'Ing. Bonfanti sarei andato poco lontano da quella scrivania davanti alla quale chiesi questa tesi. Tutto ciò che so del progetto di circuiti integrati lo devo a lui, e da lui spero di imparare ancora molto. Ringrazio anche Roberto Modaffari per gli innumerevoli pranzi insieme e la condivisione di ogni momento di stress, delusione e gioia ma soprattutto di quei momenti di soddisfazione passeggeri, quando pensi che il tuo progetto funzioni e invece non è così, ma tu ancora non lo sai. Portare a termine questo lavoro ha richiesto "Carattere!", direbbe lui. Gli auguro grandi soddisfazioni. Ringrazio Federico Pepe per l'aiuto che mi ha dato nel progettare il layout. Saluto tutti i dispositivi ed in particolare Giovanni Paolucci, uno dei pochi autentici "mirandolisti". Se tutto va come previsto, saremo insieme ancora tre anni.

Più volte mi sono chiesto come sarebbe andata se non ci fossero state le persone che in questi anni hanno condiviso le mie esperienze e soprattutto se non ci fossi stata tu, Isabella. Sei in tutti i migliori ricordi di questi ultimi due anni ed anche se forse non lo sai, anche tu hai reso questo lavoro possibile, migliore e con un dolce significato in più. Particolari ringraziamenti li devo anche ad Andrea Bonetti per avermi offerto così gentilmente la sua esperienza, è stata molto preziosa. Grazie al mio amico di sempre Marco, a Giulio, Enrico e l'altro Marco. Grazie a Katia per avermi incoraggiato a iniziare Ingegneria, ora so che sarebbe stato meglio non ascoltarla.

Ringrazio, naturalmente, mio fratello e la mia famiglia; spero di aver lavorato in questi mesi anche per dar loro un motivo di felicità.

# Indice

<b>1</b>	<b>Elettronica e Neuroscienze</b>	<b>1</b>
1.1	Generalità . . . . .	1
1.2	Il Neurone . . . . .	3
1.3	L'Interfaccia tra Neuroni ed Elettronica . . . . .	6
1.3.1	Caratteristiche Elettriche dell'Elettrodo . . . . .	7
1.4	Sistemi Integrati per il Monitoraggio Neuronale . . . . .	9
1.5	Elettronica <i>Ultra-low Power</i> . . . . .	11
<b>2</b>	<b>Specifiche, Criteri e Scelte Progettuali di Sistema</b>	<b>13</b>
2.1	Requisiti di un Sistema di Neural Monitoring . . . . .	13
2.1.1	Generalità . . . . .	13
2.1.2	Amplificazione e filtraggio . . . . .	15
2.1.3	Digitalizzazione . . . . .	17
2.1.4	Alimentazione . . . . .	22
2.1.5	Trasmissione dei Dati . . . . .	22
2.2	I Sistemi Esistenti . . . . .	25
2.2.1	Il Sistema di Harrison . . . . .	25
2.2.2	Il Sistema dell'Università della California Santa Cruz. . . . .	26
2.2.3	Il Sistema WINeR . . . . .	27
2.2.4	Il Sistema Genov-Abdelhalim. . . . .	28
2.2.5	Il Sistema del Politecnico di Milano . . . . .	29
2.3	Motivazioni e Obiettivi per un Nuovo Progetto . . . . .	31
2.3.1	Il Front-End di Acquisizione e Digitalizzazione . . . . .	32
<b>3</b>	<b>Lo stadio di Amplificazione e Filtraggio</b>	<b>36</b>
3.1	Aspetti Fondamentali . . . . .	36
3.1.1	Circuiti CMOS Sottosoglia . . . . .	38
3.1.2	Pseudoresistori MOS . . . . .	40
3.2	Caratteristiche di Sistema . . . . .	45
3.2.1	Architettura Generale . . . . .	45
3.2.2	Analisi del Rumore . . . . .	47
3.3	Low Noise Amplifier . . . . .	51
3.3.1	Progetto e Dimensionamento del Primo Stadio . . . . .	51



3.3.2	Progetto e Dimensionamento del Secondo Stadio . . . . .	59
3.3.3	Il Controllo della Tensione di Modo Comune . . . . .	63
3.3.4	Risultati . . . . .	75
3.4	Programmable Gain Amplifier . . . . .	77
3.4.1	Progetto e Dimensionamento dell'OTA . . . . .	77
3.4.2	Il Controllo della Tensione di Modo Comune . . . . .	79
3.4.3	Risultati . . . . .	81
3.5	Layout . . . . .	84
<b>4</b>	<b>Valutazioni Preliminari al Progetto dell'ADC</b>	<b>89</b>
4.1	Aspetti Generali . . . . .	89
4.1.1	Specifiche e Panorama dei Convertitori ADC . . . . .	89
4.1.2	Il SAR a Ridistribuzione di Carica . . . . .	91
4.2	Topologie Alternative . . . . .	94
4.2.1	Generalità . . . . .	94
4.2.2	Consumo delle Topologie ad Array Binario . . . . .	95
4.2.3	Linearità delle Topologie ad Array Binario . . . . .	99
4.2.4	Bridge Capacitor Array . . . . .	108
4.2.5	Consumo di topologie SAR con Bridge Capacitor . . . . .	109
4.2.6	Linearità di topologie SAR con Bridge Capacitor . . . . .	111
4.2.7	Confronto tra Topologie . . . . .	119
4.3	L'Architettura Scelta . . . . .	124
<b>5</b>	<b>Progetto e Dimensionamento del Convertitore</b>	<b>129</b>
5.1	Dimensionamento dell'array e Risultati Pre-Layout . . . . .	129
5.1.1	Interruttori di ingresso . . . . .	137
5.2	Il Comparatore . . . . .	141
5.2.1	Funzionamento . . . . .	141
5.2.2	Sensibilità ai disturbi e al rumore . . . . .	144
5.3	La Logica SAR . . . . .	148
5.3.1	Caratteristiche Generali e Funzionamento . . . . .	148
5.3.2	Il Temporizzatore . . . . .	150
5.3.3	Flip-Flop e Latch Dinamici . . . . .	152
5.4	Risultati . . . . .	156
5.4.1	Layout . . . . .	156
5.4.2	Prestazioni del Convertitore . . . . .	158
<b>6</b>	<b>Il Test Chip</b>	<b>168</b>
6.1	Prestazioni del front-end . . . . .	169
6.2	Circuiti Ausiliari . . . . .	171
6.2.1	I Riferimenti di Corrente . . . . .	171
6.2.2	Gli INA . . . . .	173
6.2.3	I Traslatori di Livello . . . . .	175

<b>A Inaccuratezza dei Modelli di Rumore in Regime di Debole Inversione</b>	<b>179</b>
<b>B Effetto del Mismatch sull'ENOB nelle Varie Topologie di CR SAR</b>	<b>182</b>
<b>Bibliography</b>	<b>186</b>

# Elenco delle figure

1	Nuovo sistema integrato di monitoraggio neuronale frutto della collaborazione tra Politecnico di Milano e Università di Padova. . . . .	15
1.1	Schematizzazione di alcune applicazioni per sistemi di monitoraggio neuronale. . . . .	2
1.2	Immagine di un neurone ottenuta con microscopio elettronico a scansione. . . . .	3
1.3	Schematizzazione dei processi di innesco del potenziale d'azione:(A) Situazione a riposo; (B) Superamento della soglia di rigenerazione a cavallo della membrana alla base dell'assone, per sommazione dei potenziali in ingresso ; (C) Apertura dei canali del Sodio e ingresso dei relativi ioni; (D) Propagazione del potenziale d'azione alle porzioni di membrana adiacenti e all'assone; (E) Apertura dei canali del Potassio; (F) Ripristino della condizione iniziale. . . . .	5
1.4	Schematizzazione della tipica morfologia del potenziale d'azione. . . . .	6
1.5	Alcune tipologie di MEAs. . . . .	7
1.6	Microwire array: foto e schema tecnico. . . . .	7
1.7	Modello equivalente di un singolo elettrodo di un MEA. . . . .	8
1.8	Tracciato comprensivo di LFPs e APs. . . . .	9
1.9	Scaling di lunghezza minima di canale e tensione di alimentazione in tecnologia CMOS. . . . .	11
1.10	Recenti evoluzioni della tecnologia dei circuiti a semiconduttori . . . . .	12
2.1	Schema generale del front-end di acquisizione e digitalizzazione. . . . .	14
2.2	Le tre alternative per l'architettura di sistema. . . . .	15
2.3	Alternative per l'organizzazione dei blocchi di amplificazione e filtraggio. . . . .	16
2.4	Esempio di caratteristica di un ADC a 3 bit (in rosso) a confronto con quella ideale (in nero). . . . .	20
2.5	Evoluzione della FoM degli ADC nell'ultima decade [5]. . . . .	21
2.6	Confronto tra le prestazioni di alcuni recenti sistemi di comunicazione wireless . . . . .	23
2.7	Il sistema in [13]. . . . .	26
2.8	Sistema di neural monitoring di Chae et al. [18]. . . . .	27
2.9	Il Sistema WINeR [30]. . . . .	28
2.10	Il Sistema Polimi [39]. . . . .	29

2.11	Il Nuovo Sistema. . . . .	33
2.12	Architettura del Front-End di digitalizzazione del nuovo sistema e con le principali specifiche schematizzate. . . . .	34
3.1	Schema elementare della struttura di un dispositivo N-MOSFET. . . . .	38
3.2	Curva di efficienza (normalizzata rispetto al massimo teorico) del dispositivo MOSFET rispetto al coefficiente di inversione. . . . .	40
3.3	NMOS e PMOS configurati a pseudo-resistori [49]. . . . .	41
3.4	Resistenza di uno pseudoresistore MOS HG UMC130nm, analisi DC al variare della tensione Drain Source. . . . .	43
3.5	Resistenza di uno pseudoresistore MOS HG UMC130nm, analisi AC. . . . .	43
3.6	Controllo in corrente della pseudoresistenza. . . . .	43
3.7	NMOS e PMOS configurati a pseudo-resistori come nel sistema di Harrison (PMOS). . . . .	44
3.8	Catena di amplificazione del front-end. . . . .	45
3.9	I due Arrays di capacità del PGA. . . . .	46
3.10	Schematizzazione dell'amplificatore e ruolo della capacità parassita. . . . .	48
3.11	NEF di lavori recentemente pubblicati( Polimi [39], Sarpeshkar [17], Harrison 2003 (1) [14] e 2007 (2) [13], Wise [15], Perelman [16] . . . . .	50
3.12	Stadio di ingresso Tail-Current-Reuse. . . . .	52
3.13	Stadio di ingresso Tail-Current-Reuse con generatori di coda reali. . . . .	55
3.14	Stadio di ingresso dell'LNA. . . . .	56
3.15	Nucleo del circuito di reazione intrinseco del primo stadio dell'LNA. . . . .	56
3.16	Stadio di ingresso dell'LNA con rete di CMF esterna. . . . .	57
3.17	Schema equivalente di un amplificatore fully differential a due stadi. . . . .	59
3.18	Esempio di OTA con stadio di uscita Common Source reazionato in modo dedicato. . . . .	61
3.19	Secondo stadio dell'OTA utilizzato per l'LNA. . . . .	61
3.20	Struttura completa di primo e secondo stadio dell'OTA utilizzato per l'LNA. . . . .	62
3.21	Descrizione schematica di un amplificatore fully-differential con rete di CMFB. . . . .	65
3.22	Generico amplificatore fully differential reazionato in modo resistivo (da [44]). . . . .	68
3.23	Individuazione grafica dei punti di lavoro di un amplificatore fully differenziale semplice a due stadi (da [44]). . . . .	69
3.24	Risultati di simulazioni in Cadence Virtuoso che mostrano punti di lavoro multipli. . . . .	70
3.25	Soluzione di CMFB agente sui MOS di carico del primo stadio, proposta in [44] per garantire la polarizzazione dello stadio. . . . .	71
3.26	Grafico che mostra il singolo punto operativo del LNA. . . . .	72
3.27	Rete di reazione di modo comune del secondo stadio del LNA. . . . .	72
3.28	Rete di reazione di modo comune del secondo stadio del LNA. . . . .	73
3.29	Schematizzazione del ruolo della capacità $C_Y$ . . . . .	74

3.30	Modulo e fase del guadagno di anello della reazione di modo comune del secondo stadio del LNA. . . . .	75
3.31	Modulo e fase del trasferimento d'anello da modo differenziale a modo differenziale dell'LNA . . . . .	76
3.32	OTA utilizzato per il PGA. . . . .	78
3.33	Modulo e fase del trasferimento da modo differenziale a modo differenziale in anello aperto del PGA nella sua configurazione a minimo guadagno. . . . .	79
3.34	Grafico che mostra il singolo punto operativo del PGA. . . . .	80
3.35	Rete di reazione di modo comune del secondo del PGA. . . . .	81
3.36	Rete di reazione di modo comune del secondo stadio del PGA. . . . .	82
3.37	Modulo e fase del trasferimento di reazione di modo comune del PGA. . . . .	82
3.38	Layout del LNA. . . . .	85
3.39	Layout del PGA. . . . .	86
3.40	Trasferimenti programmabili della catena di amplificazione e filtraggio. . . . .	86
3.41	Transitorio dell'uscita del PGA in corrispondenza di un ingresso sinusoidale di $250\mu V_{pp}$ e dei vari guadagni. . . . .	87
3.42	Regolazione della tensione di modo comune di uscita del PGA. . . . .	88
4.1	Single slope ADC [56]. . . . .	90
4.2	Utilizzo degli ADC in funzione di risoluzione e velocità [56]. . . . .	91
4.3	SAR ADC [56]. . . . .	92
4.4	Binary Weighted Array Charge Redistribution SAR ADC. . . . .	92
4.5	Binary Weighted Array (Binario), in alto. Bridge Capacitor Array, in basso. Esempi per DAC a 10bit . . . . .	95
4.6	Split Capacitor Binary Weighted Array. . . . .	96
4.7	Consumo energetico per codifica normalizzato rispetto a $E_{unit}$ delle topologie ad array binario esaminate. . . . .	97
4.8	Schematizzazione della procedura di commutazione monotona per un ADC SAR CR fully differential binario a 3 bit [58]. . . . .	97
4.9	Riduzione della tensione di modo comune all'ingresso del comparatore associato alla procedura di switching monotona a 5 bit [58]. . . . .	99
4.10	Legame tra configurazioni capacitive e caratteristica dell'ADC CR SAR binario. . . . .	100
4.11	Distribuzione della deviazione standard della DNL di un convertitore ad array binario classico, per tutte le codifiche. . . . .	103
4.12	Distribuzione della deviazione standard della DNL di un convertitore ad array binario con split capacitor, per tutte le codifiche. . . . .	104
4.13	Simulazioni alternative effettuate per estrarre informazioni sulle prestazioni di linearità delle tipologie di array. . . . .	105
4.14	Istogrammi della distribuzione del massimo modulo di DNL in un binary weighted array con algoritmo di commutazione tradizionale. . . . .	105
4.15	Cambiamento delle distribuzioni delle variabili di interesse a seconda delle elaborazioni effettuate in simulazione. . . . .	106

4.16	Confronto tra massimo modulo della DNL di ADC CR SAR binario classico e con split capacitor. . . . .	106
4.17	Risultati delle simulazioni effettuate con i modelli di ADC binario monotono (2000 realizzazioni). . . . .	107
4.18	Parassita al top plate di un binary weighted array a 6 bit. . . . .	107
4.19	Topologia con Bridge Capacitor Array. . . . .	108
4.20	Consumo energetico per codifica delle topologie bridge analizzate. . . . .	110
4.21	Schematizzazione della situazione potenzialmente peggiore per la DNL: la transizione a cavallo della codifica MSB. . . . .	114
4.22	Deviazione standard della DNL in funzione della codifica per un Bridge Array a 10bit e capacità unitaria di 20fF. . . . .	115
4.23	Principali parassiti di interesse in un array con capacità di ponte. . . . .	116
4.24	Caratteristica DNL di un Bridge Capacitor Array con parassiti. . . . .	119
4.25	Istogrammi per un array con capacità di attenuazione e capacità minima 20fF. . . . .	122
4.26	Esempio di pattern di distribuzione delle deviazioni standard della DNL. . . . .	123
4.27	Schema delle simulazioni utilizzate per ottenere l'enob. . . . .	123
4.28	Topologia con Bridge Capacitor array scelta. . . . .	125
4.29	Confronto tra le energie spese durante il ciclo di commutazione del Monotono e delle altre topologie Bridge. . . . .	126
4.30	Schema di funzionamento della topologia scelta. . . . .	126
4.31	Risultati delle simulazioni effettuate con i modelli di ADC Bridge Monotono per la valutazione dell'ENOB in funzione della capacità unitaria e del mismatch tecnologico. . . . .	127
5.1	Legame tra parasita, deviazione standard della DNL e Capacità unitaria, per l'ADC SAR CR Bridge Monotonico scelto. . . . .	130
5.2	Andamento dell'ENOB in funzione del picco massimo di DNL dovuto al parassita sul top plate del LSB. . . . .	131
5.3	Istogrammi e parametri descrittivi dell'impatto del mismatch nell'array progettato. . . . .	131
5.4	Istogrammi e parametri descrittivi dell'impatto di parassiti (ai piatti superiori) e mismatch nell'array progettato. . . . .	132
5.5	Istogrammi relativi all'ENOB stimato per l'ADC progettato con i mismatch tecnologici e parassiti di 30fF previsti ad ogni top plate dei subarray. . . . .	132
5.6	INL e DNL simulate in <i>CadenceVirtuoso<sup>TM</sup></i> per l'ADC SAR realizzato. Array e comparatore sono stati simulati in modalità schematico tenendo conto di parassiti modellizzati con capacità ideali di 40fF ai top plate dei subarray. La logica è stata implementata in linguaggio Verilog-A. . . . .	135
5.7	Singola realizzazione di INL simulata in MATLAB per l'ADC SAR realizzato. . . . .	136
5.8	Architettura dello switch con circuito di bootstrap. . . . .	138
5.9	Forme d'onda di interesse per il funzionamento dello switch bootstrappato. . . . .	139
5.10	Comparatore progettato. . . . .	142

5.11	Caratteristiche del segnale di uscita al primo stadio del comparatore durante la commutazione. . . . .	143
5.12	Commutazione del comparatore. . . . .	144
5.13	Logica SAR classica (esempio a 9 bit). . . . .	148
5.14	Logica SAR monotonica realizzata. . . . .	149
5.15	Struttura del temporizzatore. . . . .	150
5.16	Schema di funzionamento del temporizzatore. . . . .	151
5.17	Timing diagram dei segnali logici del convertitore durante un ciclo di conversione. . . . .	152
5.18	Timing diagram dei segnali logici del convertitore durante un ciclo di conversione. . . . .	153
5.19	Circuiti logici dinamici full-custom. . . . .	154
5.20	Schema di funzionamento del latch della seconda fila. . . . .	155
5.21	Layout dell'array. . . . .	157
5.22	Layout del comparatore. . . . .	158
5.23	Layout dell'ADC. . . . .	158
5.24	Alcuni dettagli del Layout dell'ADC. . . . .	159
5.25	Layout del front-end. . . . .	160
5.26	DNL e INL simulate sull'estratto del convertitore in ambiente <i>Cadence<sup>TM</sup></i> . . . . .	161
5.27	SNDR post-layout al variare dell'ampiezza di una sinusoide di ingresso. . . . .	162
5.28	Deviazioni standard della DNL in funzione della codifica per l'ADC realizzato in base ai parametri di mismatch della tecnologia. . . . .	163
5.29	Istogrammi dell'ENOB associato alla caratteristica dell'ADC realizzato affetta da parassiti e mismatch. . . . .	163
5.30	FoM del convertitore al variare della frequenza di campionamento. . . . .	164
5.31	Trend di evoluzione delle FoM in funzione del nodo tecnologico. . . . .	166
5.32	Migliori FoM in funzione del nodo tecnologico. . . . .	166
6.1	Schema del test chip e dei circuiti di servizio. . . . .	168
6.2	Immagine del Miniasic. . . . .	170
6.3	Riferimento di corrente utilizzato. . . . .	172
6.4	Struttura degli INA utilizzati come buffer di lettura delle uscite di LNA e PGA. . . . .	173
6.5	Struttura d'interna degli OTA utilizzati per l'INA. . . . .	174
6.6	Traslatore di livello. . . . .	175
6.7	Traslatore di livello: commutazione delle uscite (analisi transitoria Montecarlo). . . . .	176
A.1	Schema di simulazione per la verifica del rumore dei dispositivi. . . . .	180
A.2	Densità spettrale di rumore di un NMOS AMS035 in debole inversione e sottosoglia, confrontata con la stime teoriche. . . . .	180
A.3	Densità spettrale di rumore di un NMOS HGLV UMC130nm in debole inversione e sottosoglia, confrontata con la stime teoriche. . . . .	181

B.1	Topologia Binaria Classica (Single Ended)	182
B.2	Topologia Bridge Tradizionale (Single Ended)	183
B.3	Topologia Bridge Tradizionale (Fully Differential)	183
B.4	Topologia Split Bridge (Single Ended)	184
B.5	Topologia Split Bridge (Fully Differential)	184
B.6	Topologia Binaria Monotona	185
B.7	Topologia Bridge Monotona	185



# Elenco delle tabelle

2.1	Confronto tra le prestazioni dei principali sistemi integrati di monitoraggio neuronale . . . . .	31
3.1	Relazioni tra bit di controllo, capacità d'ingresso, capacità di compensazione e guadagno del PGA. . . . .	47
3.2	Tabella riassuntiva delle prestazioni del Low Noise Amplifier. . . . .	77
3.3	Tabella riassuntiva delle prestazioni del PGA. . . . .	83
3.4	Tabella riassuntiva delle prestazioni dello stadio di amplificazione e filtraggio. . . . .	84
4.1	Tabella riassuntiva dei legami tra capacità unitaria, parametri di tecnologia e deviazione standard massima della DNL per ciascuna topologia esaminata. . . . .	120
4.2	Tabella riassuntiva delle capacità unitarie, totali e consumi medi per avere il triplo della deviazione standard massima della DNL inferiore a 0.5. . .	120
4.3	Tabella riassuntiva delle capacità unitarie, totali e consumo medio per ogni topologia e per avere un ENOB superiore a 9. . . . .	124
5.1	Tabella riassuntiva delle prestazioni attese per l'array ottenute da simulazioni in fase di schematico. . . . .	134
5.2	Tabella riassuntiva e di confronto con i principali lavori internazionali delle prestazioni del convertitore. . . . .	165
6.1	Tabella riassuntiva e di confronto con i principali lavori internazionali delle prestazioni del front-end. . . . .	171
6.2	Risultati delle simulazioni Montecarlo (500 realizzazioni) per le correnti di riferimento generate. . . . .	173
6.3	tabella riassuntiva delle prestazioni degli INA (Simulazioni MonteCarlo). . . . .	174

# Introduzione

Nel panorama della neurofisiologia, l'indagine dei meccanismi di comunicazione all'interno di popolazioni neuronali è un processo necessario per giungere alla comprensione del funzionamento del cervello. Una conoscenza più approfondita delle modalità con cui il sistema nervoso organizza le funzioni dell'organismo potrebbe aprire vasti orizzonti nel campo del trattamento di malattie attualmente non curabili. Inoltre, la ricerca in questa direzione potrà favorire lo sviluppo di protesi ad alta capacità di interazione con il sistema somatosensoriale umano ed in particolare con il cervello. Genericamente definite BMI (Brain Machine Interfaces), queste tipologie di protesi potrebbero permettere il ripristino di funzioni motorie, e in generale fisiologiche, perse a seguito di traumi e patologie invalidanti. Ciò è possibile sfruttando la natura elettrica dei segnali attraverso i quali avviene la trasmissione dell'informazione tra le popolazioni cellulari del tessuto nervoso. Questo tipo di segnali racchiude sia i potenziali d'azione, tipicamente emessi o trasmessi dai singoli neuroni, che altre categorie riconducibili all'attività coordinata di più gruppi neuronali e che ad oggi possono essere rilevati nella pratica clinica attraverso esami superficiali come l'EEG o con tecniche più invasive come l'elettrocorticografia. Lo sviluppo di sistemi elettronici dedicati all'acquisizione ed all'elaborazione di questi tipi di segnali è fondamentale per lo sviluppo delle BMI e nel corso degli ultimi dieci anni è stato oggetto di una crescente attività scientifica. Per poter essere utilizzato in-vivo nel contesto di esperimenti di neurofisiologia o all'interno di prototipi di interfacce neurali, un circuito integrato di questo tipo deve garantire la possibilità di acquisire il segnale da tanti canali quanti sono quelli offerti dai sensori attualmente utilizzati (32, 64, 100 o 128). Il sistema deve inoltre avere una banda adeguata (0.1Hz - 10kHz) e offrire la possibilità di trasmettere in modalità wireless dell'informazione acquisita a unità di calcolo remote. Al fine di evitare fenomeni di surriscaldamento e necrosi dei tessuti circostanti l'impianto e per estendere la durata degli esperimenti con sistemi alimentati a batteria, queste funzionalità devono essere garantite congiuntamente ad un consumo molto ridotto, inferiore a pochi mV. Contenere entro questi limiti il dispendio energetico e contemporaneamente garantire la trasmissione a distanza di metri con bit-rate superiori a diversi MSps, tipicamente richiesti per l'applicazione, rappresenta una sfida nel campo dell'elettronica a basso consumo. In questo contesto, nel 2010 il Politecnico di Milano ha realizzato un sistema integrato di neural monitoring a 64 canali in tecnologia 350nm, denominato *Windowing*, alimentato a 3V e dotato di trasmettitore wireless narrowband che soddisfa i requisiti dell'applicazione. Le più recenti pubblicazioni nel campo degli integrati per il monitoraggio dell'attività neuronale hanno evidenziato un significativo

miglioramento dell'efficienza di questo tipo di SoCs (System-On-Chip) dimostrando la possibilità di ridurre i consumi e contemporaneamente aumentare la quantità di informazione trasmessa. Lo scaling tecnologico ha naturalmente favorito questo processo. A fronte di questi sviluppi il chip realizzato al Politecnico di Milano è risultato appartenente ad una vecchia generazione le cui prestazioni non erano più in linea con quelle offerte da altri, più recenti, sistemi. Come altri lavori della sua generazione, il sistema Windowing non era in grado di acquisire i segnali alle più basse frequenze (0.1-10Hz) ed effettuava una compressione dei dati acquisiti per ridurre il bitrate e l'informazione trasmessa all'essenziale. Quest'ultima aveva lo scopo di contenere i consumi, già elevati a causa della scarsa efficienza della tecnica di trasmissione adottata. Questa attività di tesi inizia a Novembre del 2011 nel contesto del progetto PRIN (Programmi di ricerca scientifica di Rilevante Interesse Nazionale), cominciato nel 2011 in collaborazione con l'Università di Padova ed il cui scopo è la realizzazione di un nuovo chip di monitoraggio neuronale. Il sistema dovrà essere in grado di acquisire e trasmettere da 64 canali tutta l'informazione disponibile alla sorgente e di garantire consumi inferiori anche rispetto ai più recenti lavori presentati in campo internazionale. L'unità di trasmissione prevista è di tipo UWB (Ultra-Wide Band) e il suo progetto sarà curato dall'Università di Padova; al Politecnico di Milano sarà invece realizzata l'elettronica di interfaccia con i sensori e di amplificazione, filtraggio e digitalizzazione del segnale.

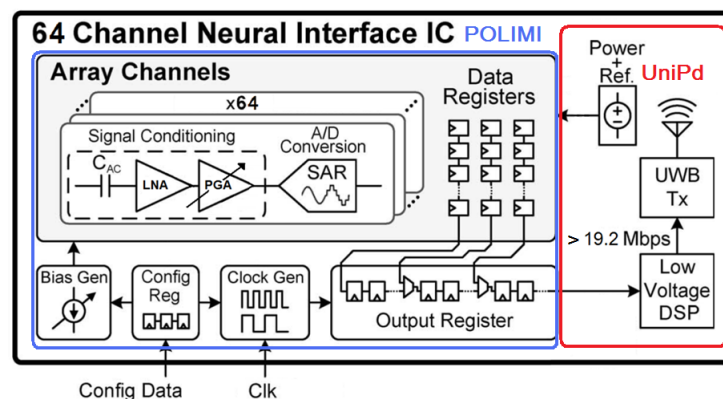


Figura 1: Nuovo sistema integrato di monitoraggio neuronale frutto della collaborazione tra Politecnico di Milano e Università di Padova.

In questo lavoro di tesi è illustrato il progetto completo di un circuito integrato costituito dalla cascata di un amplificatore a basso rumore, un amplificatore a guadagno programmabile e di un convertitore analogico-digitale a 10bit in tecnologia CMOS UMC 130nm, alimentati a 0.5V ed il cui consumo complessivo è inferiore al  $\mu W$ . L'insieme di questi circuiti costituisce il front-end di acquisizione e digitalizzazione completo per un singolo canale del sistema, la cui struttura, nella versione definitiva, sarà replicata per

64 volte.

La ricerca della massima efficienza ha richiesto lo sforzo di progettazione di ciascun blocco del sistema (amplificatori, ADC), con le prestazioni di efficienza migliori rispetto a quanto finora presentato in letteratura.

Il testo è suddiviso in sei capitoli, conclusione ed appendice. Nel **primo capitolo** sono presentati in modo generale i meccanismi che rendono possibile l'interazione tra le tecnologie elettroniche e i sistemi biologici nonché il ruolo dell'elettronica nelle neuroscienze e gli orizzonti che l'attività di ricerca in questa direzione potrebbe aprire in campo medico e scientifico.

Il **secondo capitolo** descrive le specifiche che un sistema di monitoraggio neuronale deve soddisfare per essere considerato uno strumento per esperimenti di neuroscienze. Nelle sezioni dello stesso capitolo si trovano diversi approfondimenti relativi ad alcuni sistemi di neural monitoring rappresentativi dello stato dell'arte, compreso il sistema Windowing realizzato al Politecnico di Milano. Alla fine dello stesso capitolo sono presentate e motivate le scelte progettuali generali per il progetto del sistema.

Il **terzo capitolo** illustra i dettagli del progetto dei due stadi di amplificazione e filtraggio del front-end realizzato. Sono approfondite le motivazioni dell'adozione di tutte le topologie circuitali utilizzate e sono presentate le relazioni analitiche che ne determinano le prestazioni e le relazioni di riferimento per i dimensionamenti. Sono presentati in ordine e con due sezioni dedicate, il primo amplificatore della catena, cioè il Low Noise Amplifier, e l'amplificatore a guadagno programmabile. Al termine di ciascuna sezione sono riassunte le prestazioni di ciascun blocco supportate dai risultati di simulazioni effettuate sia precedentemente che successivamente al progetto del layout. Di quest'ultimo ne sono mostrate le caratteristiche per ciascun blocco.

Il **quarto capitolo** è dedicato a valutazioni preliminari alla base del progetto del convertitore analogico-digitale. Sono analizzate le principali topologie di convertitori ad alta efficienza compatibili con l'applicazione. Per alcune di queste ne sono ricavate le caratteristiche di consumo e linearità finora non formalizzate in letteratura. E' inoltre proposto un approccio differente a quello tradizionale, per il dimensionamento dei componenti più influenti sulle prestazioni di linearità del convertitore. Le considerazioni effettuate sono state supportate da un lavoro di modellizzazione in *MATlab<sup>TM</sup>* delle diverse topologie di convertitore. Alla luce delle informazioni ricavate nel corso del capitolo, è infine illustrata e giustificata l'adozione di una nuova topologia di convertitore SAR a ridistribuzione di carica.

I dettagli del dimensionamento e del progetto dei singoli componenti del convertitore sono illustrati nel **quinto capitolo**. In particolare è da sottolineare la realizzazione di logica asincrona full-custom costituita da topologie di registri appositamente progettati per l'applicazione e di un comparatore completamente dinamico. Al termine del capitolo sono dimostrate le dimensioni, le prestazioni e l'efficienza del convertitore attraverso il

riassunto dei risultati di layout e di simulazione post-layout.

Il **sesto capitolo** riassume le caratteristiche e le prestazioni dell'intero front-end confrontandole con quanto presente in letteratura. Sono illustrate le caratteristiche del test chip realizzato ed attualmente in corso di fabbricazione presso le fonderie UMC ed è dedicato un breve approfondimento a ciascuno dei circuiti ausiliari appositamente progettati per il funzionamento ed il testing del sistema.

Per non appesantire oltremodo il corpo del lavoro ma contemporaneamente non privarlo dei necessari strumenti di analisi, alcuni risultati di simulazione a supporto delle scelte progettuali sono state relegate, con opportuni riferimenti, in appendice.

# Capitolo 1

## Elettronica e Neuroscienze

### 1.1 Generalità

L'Epilessia, l'emivrania, la sclerosi multipla, le malattie di Parkinson e di Alzheimer sono disfunzioni neurologiche che da sole costituiscono il 35% di tutte le patologie Europee [1] per un costo annuo complessivo stimato intorno a 80 miliardi di euro. Nel mondo, 450 milioni di individui sono affetti da queste patologie e, a seguito delle relative complicanze, circa 6.8 milioni di questi pazienti muoiono ogni anno [2]. Il carattere senile di molti dei disturbi considerati (pensiamo al Parkinson e all'Alzheimer) lascia intuire come questi numeri siano destinati a crescere all'interno di società che, come quelle occidentali, sono in costante invecchiamento. In questo contesto ricerca e industria elettronica possono giocare un ruolo determinante contribuendo a sviluppare nuovi sistemi diagnostici, di misura, terapeutici e riabilitativi in grado di interfacciarsi e interagire con il sistema nervoso, dalle sue strutture più complesse fino all'unità funzionale elementare: il neurone.

Cervello, tronco encefalico e midollo spinale sono strutture deputate all'organizzazione della maggior parte dei fenomeni fisiologici. Si tratta di attività volontarie come le funzioni motorie o la comunicazione verbale ed attività involontarie come la respirazione. Ogni emozione, pensiero, colore o dolore percepito è il risultato dell'attività coordinata di un elevato numero di cellule chiamate neuroni. Questi si differenziano in differenti sottotipi e si localizzano, solitari o raggruppati in popolazioni, in regioni anatomiche precise a seconda del compito per il quale si sono evoluti. Ciascuno di loro è parte di una estesa rete di interconnessioni che costituisce il sistema nervoso. Il loro insieme può essere pensato come un processore costituito da dispositivi intrinsecamente lenti (come sono i neuroni se confrontati con i dispositivi a semiconduttore) organizzati in parallelo per effettuare un grande numero di operazioni al secondo.

Come le funzioni dell'organismo, anche le disfunzioni possono essere comprese attraverso l'indagine della comunicazione tra neuroni. Le neuroscienze sono le discipline che si occupano di rivelare i segreti ancora nascosti del funzionamento del sistema nervoso; a tale scopo sfruttano ciò che la tecnologia mette loro a disposizione.

Negli ultimi vent'anni i progressi nell'acquisizione e nell'elaborazione di segnali elettroe-

necefalografici (EEG) hanno reso possibile tecniche come il monitoraggio superficiale dell'attività corticale [3] , terapie come la stimolazione cerebrale profonda (DBS) nella cura del morbo di Parkinson [4] e neuroprotesi come gli impianti cocleari che hanno gradualmente integrato la pratica clinica migliorando in modo significativo la qualità della vita dei pazienti.

A fronte della enorme complessità dei fenomeni di interazione tra i neuroni tuttavia, la ricerca può andare molto lontano senza che le siano forniti strumenti adeguati allo studio ravvicinato (perciò più selettivo) , prolungato ed in situazioni di normale attività del loro comportamento. Simili strumenti, risultato di una crescente attività di sviluppo, sono i sistemi integrati di monitoraggio neuronale, grazie ai quali sono già possibili esperimenti di neuroscienze in cui può essere rilevata l'attività parallela di numerosi singoli neuroni e di gruppi di neuroni. Anche la neurochirurgia può beneficiare di questi sistemi: si pensi alla possibilità di registrare il tracciato dell'attività di vari neuroni localizzati in una specifica area. In tal modo sarebbe possibile capire se questa va asportata perchè interessata da un processo tumorale oppure può essere conservata al fine di non compromettere la funzione cui è dedicata, con il massimo vantaggio possibile per il paziente. Infine, una vera e propria frontiera di incontro tra neuroscienze e ingegneria è costituita dallo sviluppo delle BMI (Brain-Machine Interfaces). Altrimenti dette *Interfacce Neurali*, oggi le BMI sono sistemi di comunicazione diretta tra computer e cervello umano basate tipicamente su segnali EEG e per questo dotate di scarsa risoluzione spaziale. In futuro, grazie anche ai sistemi integrati di monitoraggio neuronale potrebbero rivoluzionare la medicina e, nel senso più generale del termine, amplificare le capacità degli esseri umani.

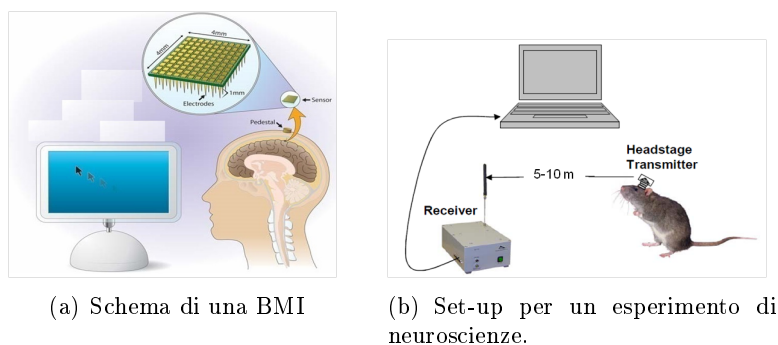


Figura 1.1: Schematizzazione di alcune applicazioni per sistemi di monitoraggio neuronale.

Mentre per molti esperimenti nel campo delle neuroscienze una connessione wireless dei sistemi di monitoraggio è solo consigliabile per migliorare la qualità del segnale ed eventualmente svincolarsi dall'ingombro e dai limiti di durata legati all'impiego delle batterie, nel caso di molte applicazioni terapeutiche impiantabili e nel campo delle BMI risulta praticamente indispensabile. Nonostante la recente esplosione dell'attività di sviluppo di sistemi di questo tipo sia in campo scientifico che in campo industriale, esiste ancora un ampio margine per l'innovazione e il miglioramento [5].

L'interazione del mondo elettronico con il sistema nervoso centrale è possibile grazie alla natura elettrica della attività cerebrale, ed è a partire da una breve descrizione dei fenomeni da cui essa origina che il lettore è invitato a procedere nel graduale avvicinamento al nucleo di questo lavoro.

## 1.2 Il Neurone

Le specifiche dei sistemi elettronici di interfaccia con il sistema nervoso traggono origine nella anatomia e fisiologia delle sue unità funzionali fondamentali ossia i neuroni. Il Neurone, mostrato in Figura 1.2, è una cellula comprendente tre differenti parti: un corpo cellulare chiamato *soma*, una serie di propaggini chiamate *dendriti*, adibite alla ricezione di informazione dai neuroni circostanti e un lungo ramo chiamato assone che si protende all'esterno della cellula, attraverso il quale il neurone è in grado di trasmettere informazione in uscita. Il soma ha un diametro tipico di  $20\mu\text{m}$  mentre dendriti e assoni, più sottili, possono estendersi da 1 millimetro fino ad un metro.

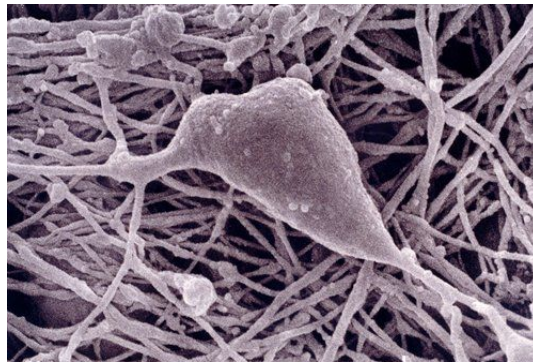


Figura 1.2: Immagine di un neurone ottenuta con microscopio elettronico a scansione.

Come ogni cellula del corpo umano, anche il neurone è uno sistema organizzato di unità funzionali genericamente indicate con il termine *organuli*, immersi in un volume di fluido detto *citoplasma* complessivamente delimitato dalla membrana cellulare. La membrana cellulare è costituita da un doppio strato di fosfolipidi ed è dotata di differenti tipi di canali proteici che in modo attivo (voltaggio o molecola dipendente) o passivo regolano il passaggio delle molecole. In condizioni di *riposo* ai capi della membrana si instaura un equilibrio elettrochimico tra specie ioniche: all'interno della cellula una concentrazione di ioni  $K^+$ , ai quali la membrana è debolmente permeabile, accompagnati da altrettanti organuli e biomolecole cariche negativamente determinano una carica citoplasmatica complessivamente negativa. All'esterno invece si ha una maggiore concentrazione di ioni  $Na^+$  e  $Ca^{2+}$ , a cui la membrana è praticamente impermeabile, e pochi ioni  $Cl^-$ . In queste condizioni la forza associata al campo elettrico (che agisce sulle particelle cariche) bilancia l'effetto dei gradienti di concentrazione delle diverse



specie garantendo una situazione di equilibrio. Il potenziale che si instaura tra interno ed esterno della membrana a seguito dello squilibrio di carica è di circa  $-90\text{mV}$  ed il suo mantenimento è possibile grazie all'azione di particolari canali proteici di membrana. L'esempio più importante di questi canali è la *pompa Na-K*, che scambia costantemente contro gradiente di concentrazione (tra l'interno e l'esterno della cellula) tre ioni Sodio con due di Potassio.

A qualsiasi spostamento esogeno o endogeno di carica che modifica i profili di concentrazione delle specie ioniche a cavallo della membrana è associata una transitoria alterazione del potenziale di membrana e viceversa. Se durante questo transitorio il potenziale transmembrana supera una determinata soglia (circa  $-60\text{mV}$ ) alcuni canali ionici di membrana, inizialmente chiusi, si attivano secondo un ordine preciso: in principio si aprono i canali del sodio che lasciano entrare nella cellula per diffusione numerosi ioni  $\text{Na}^+$ . In questa fase il potenziale transmembrana  $\delta V_{int-est}$  aumenta fino a raggiungere il picco massimo, detto *overshoot*, pochi millivolt sopra lo zero. Questo valore di potenziale innesca l'apertura dei canali del Potassio, con la conseguente fuoriuscita di ioni  $\text{K}^+$  cui si associa il fenomeno di ripolarizzazione della membrana, ovvero la diminuzione del potenziale ai suoi capi; nel frattempo i canali del sodio si chiudono. A seguito della fuoriuscita massiva di Potassio la membrana si trova in una momentanea condizione di iper-polarizzazione e refrattarietà a nuove depolarizzazioni dalla quale recupera in breve tempo fino alla condizione di partenza grazie all'azione costante delle pompe molecolari di membrana.

Il meccanismo appena descritto e illustrato in figura 1.3 realizza un'amplificazione del segnale originario, che altrimenti si attenuerebbe fino al ripristino delle condizioni di partenza. In questo modo una semplice variazione di tensione è tradotta in una autentica *onda di depolarizzazione* identificata con il termine di *potenziale d'azione* (PA, oppure AP, Action Potential) in grado di propagarsi e rigenerarsi lungo le regioni adiacenti della membrana per poi estendersi a tutta la cellula. La capacità di depolarizzazione è comune alla maggior parte delle cellule umane; la soglia di innesco del PA, la sua morfologia e i valori variano a seconda del tipo di cellula ma mostrano caratteri generali comuni. Nel Neurone il PA misurato tra l'interno e l'esterno di un punto della membrana ha un andamento temporale che può essere descritto dalla curva di figura 1.4.

I neuroni si interfacciano tra loro attraverso le *sinapsi*: connessioni mediate da molecole organiche chiamate neurotrasmettitori che permettono il trasferimento di carica ionica, cioè il passaggio di informazione nella forma di PA. Esse sono posizionate alle estremità dei dendriti e degli assoni, in corrispondenza dei punti di contatto con la membrana del neurone adiacente. Attraverso le vie di comunicazione dendritiche il neurone è in grado di rilevare potenziali cellulari di cellule a lui connesse. Quando la sommazione dei potenziali post-sinaptici (provenienti da neuroni direttamente connessi) rilevata alla base dell'assone supera la soglia di circa  $-60\text{mV}$ , il PA si scatena e si propaga attraverso la parte finale della membrana del soma e infine, per contiguità, lungo l'assone. La particolare struttura di quest'ultimo permette la trasmissione rigenerativa del segnale fino a destinazioni qualsivoglia lontane dell'organismo entro poche decine di millisecondi.

Per essere apprezzata e utilizzata a scopi medici l'informazione contenuta nelle variazioni di potenziale deve essere misurata con opportuni strumenti. Tali strumenti

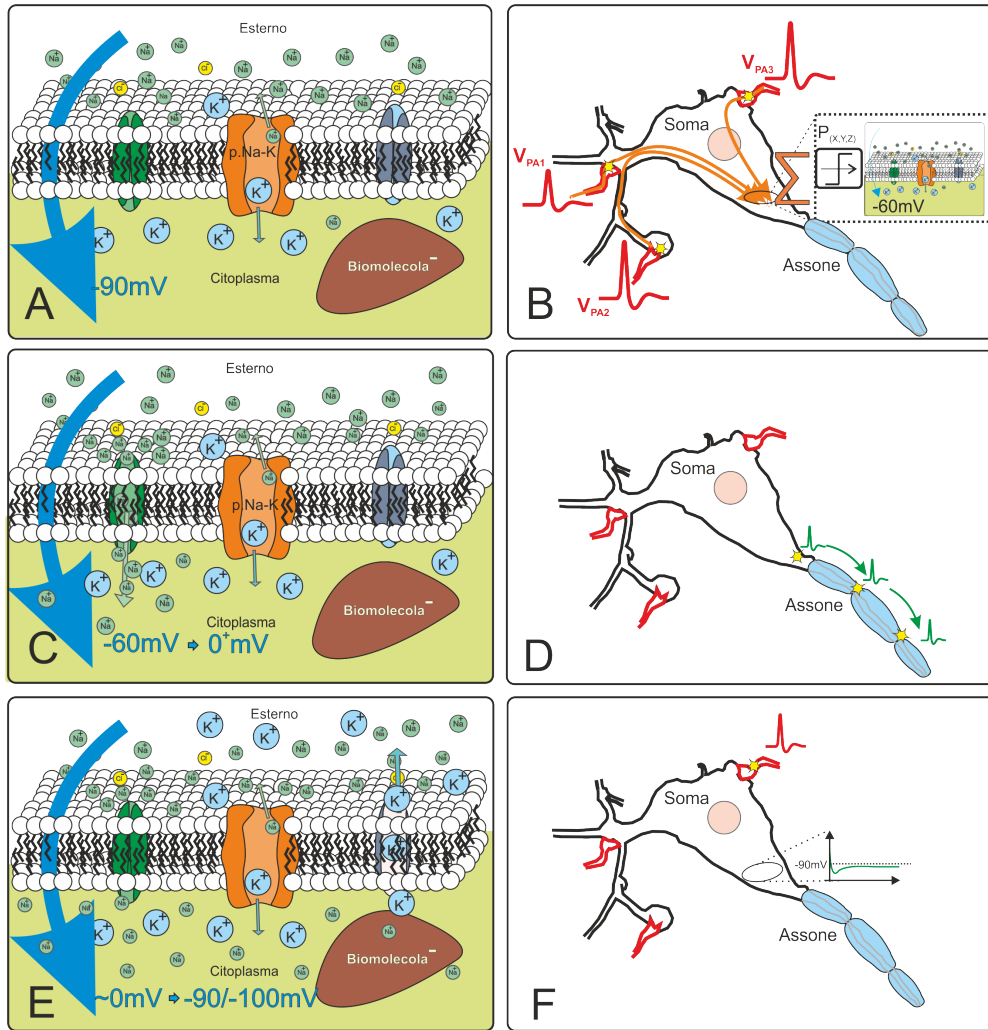


Figura 1.3: Schematizzazione dei processi di innesco del potenziale d'azione:(A) Situazione a riposo; (B) Superamento della soglia di rigenerazione a cavallo della membrana alla base dell'assone, per sommazione dei potenziali in ingresso ; (C) Apertura dei canali del Sodio e ingresso dei relativi ioni; (D) Propagazione del potenziale d'azione alle porzioni di membrana adiacenti e all'assone; (E) Apertura dei canali del Potassio; (F) Ripristino della condizione iniziale.

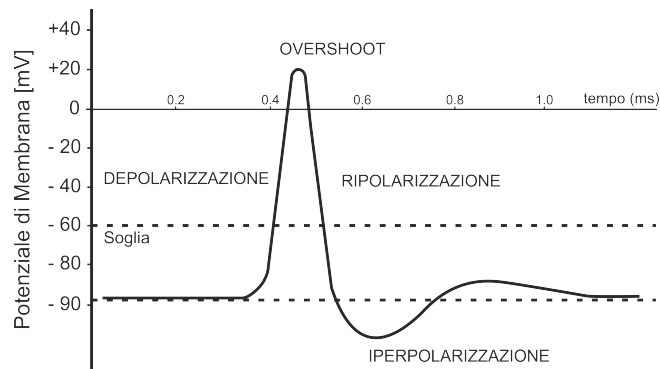


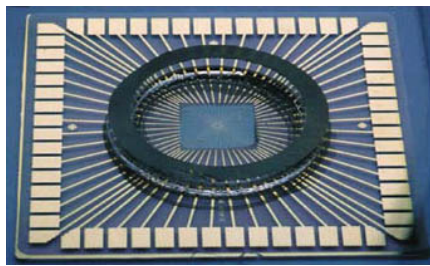
Figura 1.4: Schematizzazione della tipica morfologia del potenziale d'azione.

dovrebbero garantire una adeguata varietà di risoluzioni spaziali e temporali, così da poter osservare sia l'attività della singola cellula, sia l'attività collettiva di intere popolazioni neuronali. Mentre il secondo tipo di attività può essere già registrata in modo non invasivo con tecniche convenzionali come l'EEG, la prima richiede sensori in grado di leggere i potenziali extra-cellulari, cioè i segni dei potenziali d'azioni che sono rilevabili a poca distanza dalla membrana.

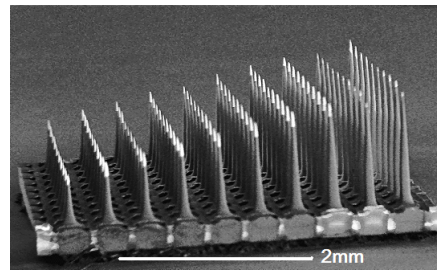
### 1.3 L'Interfaccia tra Neuroni ed Elettronica

Un importante crocevia nello sviluppo di sistemi elettronici per il monitoraggio dell'attività neuronale è stato raggiunto con lo sviluppo dei primi array di microelettrodi (MEAs, multi-electrode arrays) per la lettura di potenziali extra-cellulari, in grado di garantire selettività di acquisizione rispetto alla singola cellula e la lettura simultanea da molteplici sorgenti. L'intuizione di costruire simili elettrodi sfruttando processi litografici applicabili su larga scala tipici dell'industria dei semiconduttori risale agli anni '60 ed è attribuibile al Prof. Moll della Stanford University [6]. Da allora l'idea si è evoluta portando alla realizzazione di due famiglie di MEAs: una comprendente elettrodi per la registrazione in vitro (Figura 1.5(a)) (cui tipicamente ci si riferisce quando si parla di MEA) e l'altra per la registrazione in-vivo. I MEA dedicati al monitoraggio in-vivo di segnali nervosi sono di maggiore interesse clinico e si distinguono a loro volta costituiti in tre grandi architetture: *Michigan Array* [8], il più recente *Utah Array* [7] (Figura 1.5(b)) e i *microwire arrays*.

Michigan e Utah array sono entrambi realizzati sfruttando i processi di produzione tipici dell'industria dei semiconduttori ed hanno quindi il vantaggio di poter integrare circuiti di elaborazione del segnale acquisito sul substrato di supporto comune agli stessi elettrodi. Il primo rappresenta la naturale evoluzione dell'intuizione di Moll ed è realizzato in tecnologia planare su un substrato di Silicio: i punti di registrazione degli elettrodi, delimitati da diffusioni di Boro, sono ricoperti da uno strato metallico realizzabile in Platino, Oro oppure Iridio. L'array complessivo garantisce 64 canali di registrazione. Lo



(a) MEA per la registrazione in vitro



(b) Esempio di Utah array

Figura 1.5: Alcune tipologie di MEAs.

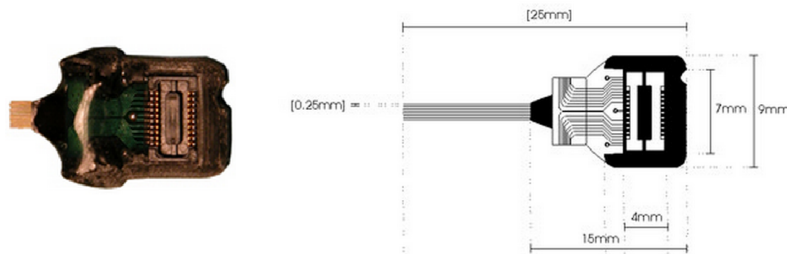


Figura 1.6: Microwire array: foto e schema tecnico.

Utah array, realizzato anch'esso su un substrato in Silicio, è una matrice tridimensionale a base quadrata ( $4\text{mm} \times 4\text{mm}$ ) di microelettrodi ad ago. Gli elettrodi sono distanziati di  $400\mu\text{m}$  gli uni dagli altri e la lunghezza di ogni singolo ago è programmabile fino a  $1.5\text{mm}$  prima della produzione. La punta di ciascun elettrodo è metallizzata in Platino o Iridio, mentre il resto della struttura è isolata con una copertura biocompatibile in Poliammide: questo permette di registrare non solo sorgenti multiple poste sullo stesso piano, come permetterebbe l'architettura Michigan, ma anche in specifiche configurazioni tridimensionali. I microwire arrays, che sfruttano processi produttivi diversi ed hanno una architettura alternativa rispetto ad entrambe le classi di elettrodi finora descritte, sono mostrati in fig. 1.6. Costituiti da veri e propri micro-fili paralleli di acciaio o alternativamente Tungsteno, presentano una parte terminale scoperta e dedicata alla registrazione mentre la restante è ricoperta da uno strato biocompatibile e isolante di Poliammide. Il numero di canali offerto dai microwire arrays del panorama commerciale varia tra 8 e 64.

### 1.3.1 Caratteristiche Elettriche dell'Elettrodo

Nei tessuti la conduzione è dovuta ad un flusso di ioni, mentre la corrente elettrica è causata da un flusso di elettroni; quando l'elettrodo entra in contatto con l'ambiente extracellulare, gli ioni caricati positivamente presenti in soluzione sono attratti verso

la sua superficie, (in genere dotata di carica negativa) e tendono a rivestirla formando il cosiddetto *strato di helmholtz interno*. Un ulteriore strato di ioni idrati posto in successione al precedente nella direzione di allontanamento dalla superficie metallica dell'elettrodo costituisce lo *strato di Helmholtz esterno*. All'interfaccia tra mondo ionico e mondo elettronico, ossia a cavallo del doppio strato tra tessuto biologico ed elettrodo, avvengono reazioni di ossidoriduzione cui si associa un trasporto di carica tra i due mondi: in tal modo gli ioni presenti nel tessuto biologico possono interagire con l'elettrodo cedendo o acquistando elettroni utili alla conduzione elettrica. All'equilibrio il numero delle ossidazioni eguaglia il numero delle riduzioni e non vi è conduzione tra i due mondi; al contrario, in condizioni dinamiche tale equilibrio si altera e si instaura una corrente di segnale. Al doppio strato di carica di Helmholtz corrisponde una capacità  $C_d$  ed una differenza di potenziale  $E_{hc}$  denominata *potenziale di semielemento*. Entrambe queste grandezze sono parametrizzate nel modello equivalente dell'elettrodo per il singolo canale [11], mostrato in figura 1.7.

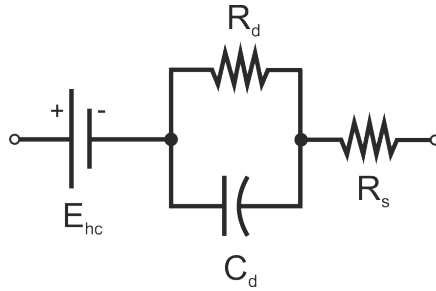


Figura 1.7: Modello equivalente di un singolo elettrodo di un MEA.

Le resistenze  $R_d$  e  $R_s$ , del circuito equivalente modellizzano rispettivamente le perdite di elettroni che fluiscono dall'elettrodo al fluido extracellulare e la resistenza serie attribuibile al tessuto biologico. Entrambe le resistenze dipendono dalla frequenza e rendono conto della parte reale dell'impedenza complessiva, tuttavia il loro contributo nella banda di interesse è trascurabile. Alla frequenza di 1kHz, in piena banda utile, il singolo elettrodo offre infatti un'impedenza dominata dal contributo reattivo e può variare da diverse centinaia di  $k\Omega$  fino ad un  $M\Omega$ , valori che corrispondono a capacità comprese tra 100pF e 350pF; Alla parte reale dell'impedenza si associa un rumore termico corrispondente a quello della sua resistenza equivalente. Sommando questo contributo termico ai disturbi e alle fluttuazioni di fondo si ottiene un rumore sovrapposto alla sorgente che calcolato su una banda di 10kHz è pari a circa  $5-10\mu V_{rms}$ , [29]. Questi valori, come vedremo nel capitolo 2, determinano le specifiche dell'elettronica a valle: essa deve potersi interfacciare con un'impedenza così grande senza significativa perdita di segnale e contemporaneamente il rumore che introduce deve essere tale da non degradare il SNR.

## 1.4 Sistemi Integrati per il Monitoraggio Neuronale

Indipendentemente dalla tipologia di elettrodo, l'obiettivo principale dello sviluppo di interfacce neurali in-vivo è l'integrazione di sistemi multicanale di acquisizione, digitalizzazione e trasmissione wireless dei segnali sullo stesso substrato degli elettrodi al fine di disporre dispositivi compatti con la minore lunghezza possibile di collegamenti così da ridurre rumore aggiuntivo, interferenze ma anche l'ingombro. Se da un lato il singolo elettrodo dev'essere sufficientemente piccolo da poter avvicinare selettivamente pochi neuroni o addirittura il singolo soma per leggerne l'attività, la presenza di più canali è indispensabile per monitorare intere aree funzionali e popolazioni neuronali. Uno dei grandi obiettivi nello sviluppo di una *neural probe* è la cattura dell'intero contenuto informativo disponibile alla sorgente. Esso è costituito da due tipi di segnali: gli spikes dei singoli neuroni e i Local Field Potentials (LFPs), cioè l'attività simultanea di popolazioni neuronali limitrofe al punto di registrazione ma troppo distanti dall'elettrodo perchè esso possa rilevare depolarizzazioni distinte [12]. Mentre la prima categoria ha un contenuto in frequenza compreso tra i 100Hz e i 5kHz, la seconda ha una banda che si estende nelle basse frequenze fino a 100mHz. Gli LFPs rappresentano la manifestazione intra-corticale del segnale EEG, che nella pratica clinica è acquisito a livello cutaneo; si tratta di segnali ampi che possono contenere parte delle istruzioni motorie e perciò risultano interessanti per la realizzazione di neuroprotesi. A valle dell'elettrodo i segnali presentano tipicamente un'ampiezza variabile tra poche decine di microvolts fino a 2mV, in figura 1.8 è mostrato un esempio di tracciato comprensivo di LFPs e APs.

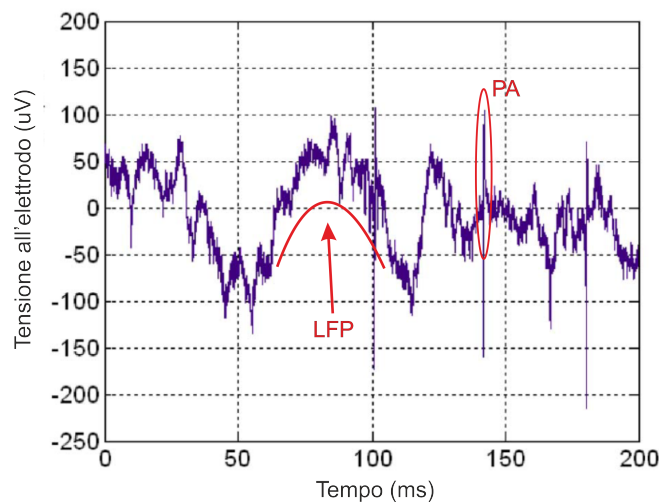


Figura 1.8: Tracciato comprensivo di LFPs e APs.

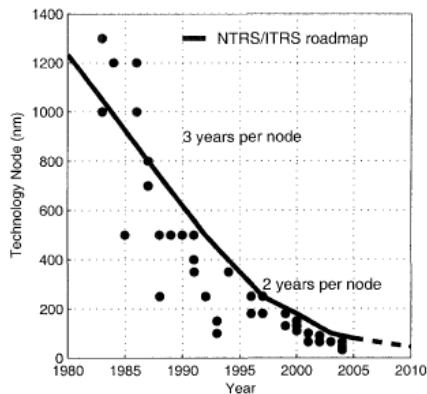
La possibilità di trasmettere senza fili i dati acquisiti permette di ridurre interferenze che spesso si concatenano ai cavi aprendo al contempo la possibilità di impianto del dispositivo. Un ulteriore requisito a tal fine è la sufficiente autonomia del dispositivo, ma anche in questo la tecnologia wireless rappresenta una soluzione interessante per i

sistemi di alimentazione. Per questi motivi le prestazioni dell'elettronica a valle dell'elettrodo e soprattutto il consumo di potenza rivestono una importanza fondamentale. I requisiti di consumo soffrono anche di un particolare legame con le specifiche di compattezza che sistema integrato impiantabile di monitoraggio neuronale deve soddisfare. Un chip impiantabile, il cui volume ragionevole assumendo di avere almeno 64 canali è di  $4\text{mm} \times 4\text{mm} \times 1.5\text{mm}$ , non dovrebbe mai consumare più di  $12.5\text{-}13\text{mW}$  (cifra già piuttosto bassa) per non superare la densità di potenza dissipata di  $800\mu\text{W}/\text{mm}^2$ , oltre la quale esiste il rischio di danno termico ai tessuti biologici [9] [10]. Il rispetto di questo vincolo imporrebbe un consumo di potenza per canale inferiore a  $195\mu\text{W}$ . Tutto questo deve essere realizzato pur garantendo l'acquisizione dell'intero contenuto informativo del segnale e la sua trasmissione in formato digitale, che come sarà spiegato nel prossimo capitolo, necessita di data-rates nell'ordine di diversi Mbit/s.

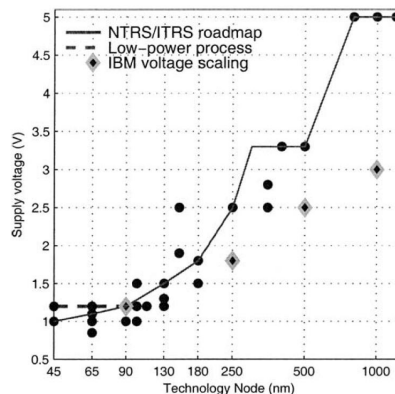
Il progetto di stadi amplificanti integrati a basso rumore e con banda estesa a frequenze così basse costituisce tuttora una sfida. Il primo circuito integrato completo comprendente amplificazione digitalizzazione e trasmissione è stato realizzato in tecnologia CMOS da Harrison et al. [13] nel 2007. Da allora, la produzione scientifica di sistemi integrati di elaborazione del segnale neurale è molto aumentata ed è tutt'oggi in crescita. Il sistema di Harrison, alimentato a  $3.3\text{V}$  per induzione elettromagnetica e realizzato in tecnologia  $0.5\mu\text{m}$ , garantiva 100 canali e prevedeva la sola elaborazione delle trasmissioni degli istanti di arrivo spikes neuronali senza essere in grado di acquisire i LFPs. In questo modo è stato possibile ottenere un bit-rate basso di circa 330kbps e non penalizzare eccessivamente il consumo del trasmettitore FSK a 433MHz utilizzato e ottenerne uno complessivo di  $13.5\text{mW}$ , pari a  $135\mu\text{W}/\text{ch}$ . La registrazione dell'intera traccia rispetto al singolo PA permetterebbe di identificare più accuratamente la soglia di rilevazione degli spikes e sarebbe di maggiore interesse clinico in quanto la completa morfologia dello spike (che si perderebbe in un sistema a soglia) contiene informazioni importanti come ad esempio il tipo di movimento comandato da neuroni della corteccia motoria [20]. Il prezzo di una trasmissione completa è un maggior bitrate e quindi un maggior consumo. In questo contesto numerosi lavori successivi hanno proposto varianti, ad esempio in quello di Ghovanloo et al.[30] sostituisce la digitalizzazione con una conversione amplitude-to-time riducendo drasticamente il bitrate ( $1.5\text{Mbit}/\text{s}$ ) della trasmissione, che è effettuata attraverso un trasmettitore wireless FSK ma che richiede una complessa elaborazione analogica del segnale a valle. Wise et al. [15] presenta un sistema che permette di collezionare il tracciato completo di due tra i 64 canali totali. Per gli altri canali è registrato unicamente l'istante di occorrenza degli spikes e la trasmissione complessiva dei dati avviene a 2Mbps. In [18] Chae et al. è descritto un sistema in  $0.35\mu\text{m}$  di 128 canali capace di rilevare e trasmettere il segnale dei singoli spikes o alternativamente l'intero tracciato utilizzando una trasmissione UWB (Ultra Wide Band) ad un bitrate superiore a 90Mbps, garantendo un range di oltre 1m e contenendo il consumo di potenza per canale a  $47\mu\text{W}$ . HermesE, il sistema messo a punto dal gruppo di T. Meng dell'università di Stanford, acquisisce e trasmette l'intero stream di dati acquisiti dal MEA, ma la codifica e la trasmissione sono implementate esternamente. Alla descrizione di questi sistemi e all'analisi di alcuni loro punti di forza e problemi sarà dedicata una parte del secondo capitolo.

## 1.5 Elettronica *Ultra-low Power*

Nonostante il consumo di potenza sia migliorato ad ogni passo dello scaling tecnologico, anche la densità di dispositivi integrati nei SoCs ( Systems-on-Chip ) è considerevolmente aumentata e l'energia spesa per ogni funzione di elaborazione del segnale è rimasta un elemento critico sia in campo digitale che in campo analogico.



(a) Evoluzione del nodo tecnologico.

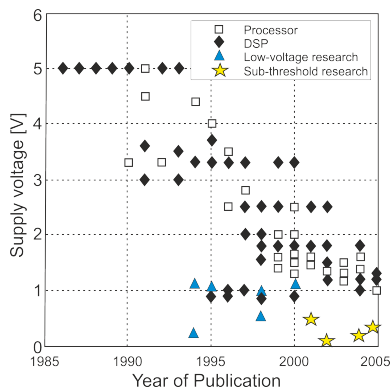


(b) Scaling della tensione di alimentazione con la tecnologia.

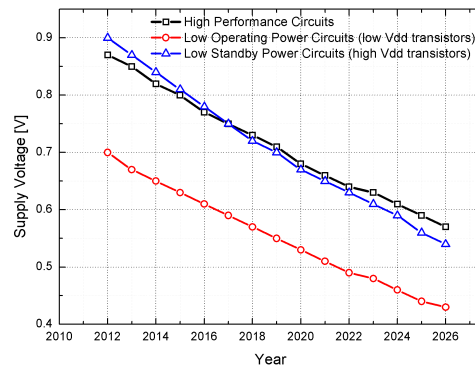
Figura 1.9: Scaling di lunghezza minima di canale e tensione di alimentazione in tecnologia CMOS.

L'estrema riduzione delle dimensioni minime dei dispositivi e al contempo delle tensioni di alimentazione per circuiti analogici e digitali, evidenziata nelle figure 1.10(a) e 1.10(b), favorisce lo sviluppo di circuiti digitali e analogici caratterizzati da consumi di potenza estremamente ridotti. Le pubblicazioni dell'ultimo decennio nel campo della ricerca elettronica a basso consumo (Fig. 1.9(a)) e le stime di riduzione della tensione di alimentazione previste per i circuiti integrati di futura commercializzazione (Fig. 1.9(b)) rappresentano ulteriori significativi indizi dell'accelerazione tecnologica su questo fronte. Il limite fondamentale dello scaling delle tensioni di alimentazione è stato evidenziato nel 1972 dal Prof. Meindl e dal Dr. Swanson [23]. Essi dimostrarono la possibilità di operazioni digitali a tensioni di alimentazione estremamente basse, prevedendone il limite a circa 200mV. Pochi anni dopo, a Losanna, il Prof. Eric Vittoz iniziò per primo l'esplorazione nel campo dei circuiti analogici a minima tensione di alimentazione, ottenendo risultati progettuali compatibili con comuni applicazioni dell'elettronica di consumo, quali orologi da polso e calcolatrici portatili. Vittoz fu il pioniere della modellizzazione e del progetto di circuiti CMOS analogici con transistori in *debole inversione* (weak inversion) e in regime di *sottosoglia* (sub-threshold), contribuendo a formalizzare il celebre modello EKV (Enz, Krummenacher, Vittoz) [24] [25]. Circuiti di questo tipo, operanti con tensioni di alimentazione tra 1V e 0.5V sono ora largamente presenti nel mondo della ricerca.





(a) ISSCC : Tensione di alimentazione dei principali lavori nella ricerca low-power [21].



(b) Proiezione futura delle tensioni di alimentazione per circuiti della catena di produzione[22].

Figura 1.10: Recenti evoluzioni della tecnologia dei circuiti a semiconduttori

Il progetto analogico low-power ha di fatto rappresentato una delle maggiori sfide degli ultimi dieci anni. Il crescente numero di articoli riguardanti SoCs a basso consumo e tensione di alimentazione oggi pubblicati su riviste e conferenze internazionali dimostrano il grande interesse che comunità scientifica e industria elettronica pongono sull'argomento.

Le applicazioni *ultra low power* sono trasversali, se non universali: partendo dai processori digitali, passando da amplificatori a basso rumore fino ad arrivare a circuiti mixed-mode come ADC, DAC e RF transceivers. In questo contesto le applicazioni bioelettroniche forniscono un campo vasto e ricco di sfide applicative. I sistemi bioelettronici spesso richiedono l'impiantabilità a lungo termine, alimentazioni compatte o wireless e ridotte dimensioni del circuito, che tuttavia deve comprendere tutto quello che serve al trattamento del segnale, dal sensore al trasmettitore. Nonostante alcuni risultati incoraggianti [26], le prime dimostrazioni di sistemi come quelli descritti indicano chiaramente che alcune sfide fondamentali devono ancora essere superate per rendere possibile la loro trasformazione in prodotti commerciali affidabili e dotati di prestazioni all'altezza delle richieste: miniaturizzazione, adeguata banda, risoluzione, range di trasmissione dei dati, sistema di alimentazione e biocompatibilità.

## Capitolo 2

# Specifiche, Criteri e Scelte Progettuali di Sistema

### 2.1 Requisiti di un Sistema di Neural Monitoring

#### 2.1.1 Generalità

Un sistema integrato per il monitoraggio dell'attività neuronale deve comprendere un front-end di acquisizione e digitalizzazione, un trasmettitore, una unità di alimentazione ed una serie di circuiti di servizio (riferimenti di corrente, tensione, registri, clock...). Mentre i requisiti del sistema di alimentazione saranno discussi in un paragrafo dedicato, è da subito possibile definire i quattro fondamentali criteri che tratteggiano le linee guida da cui discendono le specifiche di progetto per front-end e trasmettitore:

1. Molteplicità di canali: per sostenere una strategia di acquisizione in parallelo di più sorgenti corticali. Allo stesso scopo, l'impegno di un MEA è obbligatorio ed è consigliabile un numero di elettrodi superiore a 32;
2. Filtraggio passabanda: lo spettro del segnale e le caratteristiche elettriche del tessuto biologico impongono i limiti della banda di filtraggio del sistema (100mHz - 5kHz) ;
3. Basso rumore: il rumore sovrapposto al segnale a valle dell'elettrodo impone le specifiche di rumore del front-end di amplificazione (inferiore a  $5\mu V_{rms}$  sulla banda utile);
4. Basso consumo: le proprietà di conduzione termica dei tessuti biologici circostanti il chip determinano il power budget di un sistema completo (amplificatori, ADC, logica, trasmettitore) la cui alimentazione è fornita in modalità wireless;

Come già accennato nel capitolo introduttivo, i potenziali extracellulari sono impulsi bifasici di durata tipicamente compresa tra  $750\mu s$  e  $1ms$  e con un contenuto in frequenza esteso da pochi Hertz fino a 5-10kHz mentre LFPs contengono informazioni situate a

frequenze inferiori all'Hertz. L'elettrodo è in grado di leggere univocamente i singoli spikes solo se posizionato nel raggio di  $50\mu\text{m}$  dalla cellula sorgente, mentre a distanze superiori può comunque rivelare potenziali dovuti all'attività di fondo di gruppi neuronali [28]. La forma e l'ampiezza del segnale dipendono dalla cellula sorgente, dalla tecnologia dell'elettrodo e dalla sua distanza dalla cellula e variano tra poche decine di microvolt e 1-2mV. Ampiezze così piccole suggeriscono di porre immediatamente a valle dell'elettrodo, prima della digitalizzazione, uno o più stadi di amplificazione, secondo lo schema di fig. 2.1.

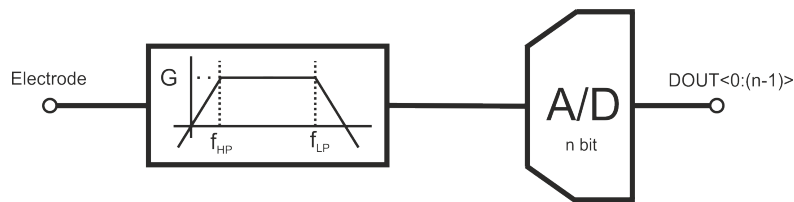


Figura 2.1: Schema generale del front-end di acquisizione e digitalizzazione.

Quanto descritto finora riguarda fenomeni circoscritti allo spazio circostante il singolo elettrodo di registrazione. In un MEA tuttavia la lettura dei segnali avviene in parallelo mentre la trasmissione riguarda il bit-stream comune a tutti i canali, e ad una frequenza molto superiore alla banda del segnale biologico. Per questo motivo l'organizzazione dei blocchi funzionali del sistema è un aspetto di primaria importanza perchè determina l'ordine delle elaborazioni che i segnali subiscono per essere digitalizzati e convogliati in un'unica sequenza di bit e può incidere molto sull'efficienza del sistema. In questo senso esistono tre architetture possibili, illustrate in figura 2.2 ed elencate di seguito:

1. Un ADC per canale (Fig 2.2a): non richiede particolare circuiteria di servizio ma solo l'ottimizzazione dei consumi del singolo convertitore, rilassa le specifiche del singolo ADC rispetto alla frequenza di campionamento ma ne inasprisce i requisiti di consumo;
2. Un ADC singolo per l'intero sistema (Fig 2.2b): richiede un buffer per ogni canale ed un MUX analogico, rilassa le specifiche di consumo dell'ADC, inasprendo quelle relative alla frequenza di conversione;
3. Un ADC ogni  $k$  canali (Fig 2.2c): è la soluzione intermedia rispetto alle due precedenti ed anche quella che richiede la maggiore complessità circuitale;

Ora che gli aspetti più generali sono stati definiti, è possibile procedere nell'analisi delle specifiche dei singoli blocchi. Si tenga presente che l'obiettivo di questa sezione è illustrare le possibilità progettuali e i requisiti necessari alla realizzazione di un sistema di utilità pratica. Ogni mancato riferimento alle scelte adottate in seguito nel progetto del sistema è voluto, queste saranno infatti presentate in modo generale nell'ultima sezione del capitolo e più in dettaglio nei capitoli successivi.

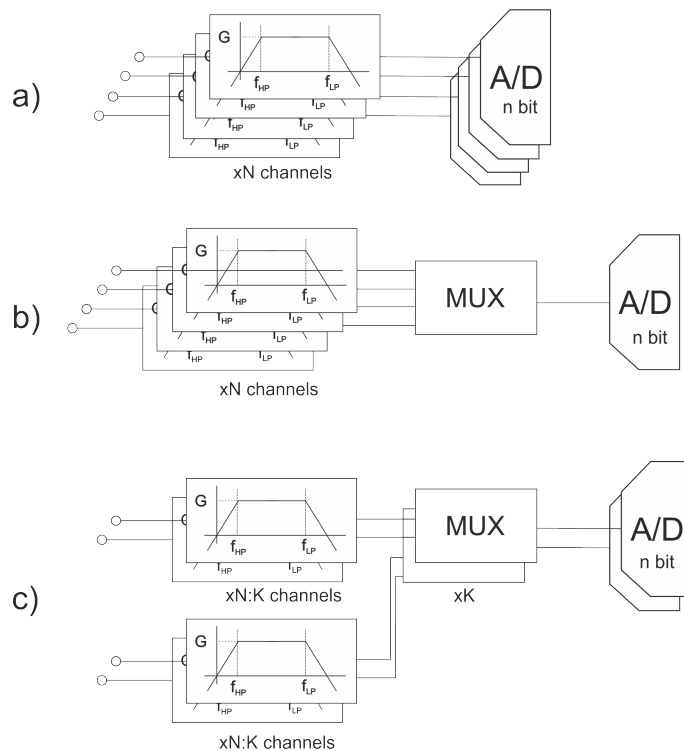


Figura 2.2: Le tre alternative per l'architettura di sistema.

### 2.1.2 Amplificazione e filtraggio

Le fluttuazioni fisiologiche di ioni nella matrice extracellulare rilevate dall'elettrodo possono causare lente variazioni del potenziale di ingresso anche superiori a 200mV. Tali segnali, oltre a non contenere informazioni di particolare utilità, sono assimilabili a vere e proprie variazioni di polarizzazione (Offset) che potrebbero portare fuori dal punto di funzionamento un generico circuito posto a valle. Per evitare tale inconveniente ma contemporaneamente acquisire tutta l'informazione disponibile all'elettrodo si impone un filtraggio passa banda tra 100mHz e 5-10kHz.

In generale sono tre le strategie che si possono alternativamente adottare per il progetto di uno stadio di amplificazione e filtraggio e sono illustrate in figura 2.3.

La soluzione 2.3(a) mostra un filtro passa alto con un frequenza di taglio corrispondente al limite inferiore della banda del segnale, seguito da un singolo stadio che realizza il guadagno complessivo e il filtraggio passa-basso. Una soluzione di questo tipo è da scartare perché non permette sufficiente riduzione del rumore riferito in ingresso [5]. La soluzione 2.3(b) sfrutta un primo pre-amplificatore che realizza parte del guadagno, seguito da uno stadio di filtraggio passa alto e da un secondo stadio che completa il guadagno complessivo. Anche questa soluzione, in cui il primo amplificatore è direttamente connesso all'elettrodo, è da scartare in quanto sottoporrebbe quest'ultimo alle grandi variazioni di offset di cui si è discusso in precedenza, con il rischio di saturare la dinamica.

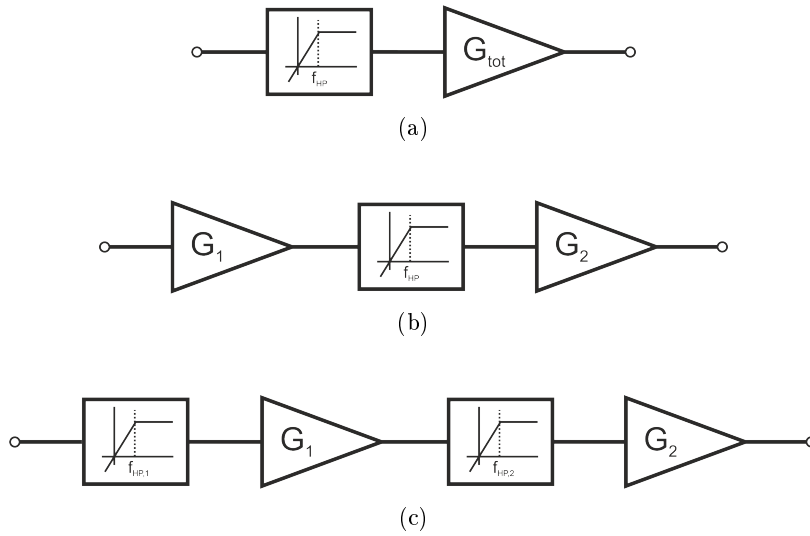


Figura 2.3: Alternative per l'organizzazione dei blocchi di amplificazione e filtraggio.

La terza configurazione (Fig. 2.3(c)) è la più indicata ed è costituita da un primo stadio di moderato guadagno accoppiato in AC ( $f_{HP,1}$  seguito in successione da un filtro passa alto con frequenza di taglio  $f_{HP,2}$  e dal secondo amplificatore che completa il guadagno e realizza il filtraggio passa basso. Separare la frequenza di taglio passa alto inferiore  $f_{HP,1}$ , dalla più alta  $f_{HP,2}$ , permette di minimizzare la figura di rumore della catena. La formulazione esatta del rumore riferito all'ingresso ed i passaggi con i quali essa si ricava dipendono dall'architettura dei singoli stadi, per questo motivo maggiori dettagli a riguardo saranno presentati nel capitolo relativo al progetto del front-end di amplificazione. E' tuttavia facile comprendere come il contributo dominante provenga dal primo amplificatore. Si tenga inoltre presente che l'impedenza di ingresso del front-end dev'essere molto superiore a quella offerta dall'elettrodo nella banda di interesse, pari ad  $1M\Omega$  ad  $1kHz$  circa. Il guadagno dovrebbe essere variabile per evitare la saturazione ma contemporaneamente sfruttare l'intera dinamica di uscita dell'amplificatore. Avendo in ingresso ampiezze variabili tra poche decine di  $\mu V$  e  $2mV$ , uno stadio single ended alimentato ad  $1V$  con una verosimile dinamica di uscita di  $0.6V$ , dovrebbe disporre di un guadagno selezionabile tra  $50dB$  e  $90dB$ . Una buona soluzione prevede la distribuzione di quest'ultimo attribuendone una porzione fissa e maggioritaria al primo amplificatore (per non degradare le prestazioni di rumore) e la restante, variabile o programmabile, ad un secondo amplificatore (Variable Gain Amplifier, VGA o Programmable Gain Amplifier, PGA). Quanto al consumo, in un amplificatore esso dipende prevalentemente dalle specifiche di rumore da soddisfare. Esse impongono infatti un limite inferiore alla corrente di polarizzazione dello stadio di ingresso che è tanto più elevato quanto più basso è il massimo rumore che si accetta di introdurre. I parametri di riferimento per valutare la qualità dell'elettronica di amplificazione sono due: si chiamano *Noise Efficiency Factor* (NEF) [41] e *Power Efficiency Factor* (PEF) [40] e sono calcolati come segue:

$$NEF = V_{n,in,rms} \sqrt{\frac{2I_{tot}}{\pi U_T 4kTBW}} \quad (2.1)$$

$$PEF = NEF^2 V_{DD} = V_{n,in,rms}^2 \frac{2P_{tot}}{\pi U_T 4kTBW} \quad (2.2)$$

Queste due figure di merito tengono conto del rumore riferito in ingresso e della banda sulla quale è calcolato. Mentre la prima fornisce indicazioni di efficienza soltanto rispetto alla corrente totale di polarizzazione richiesta per il funzionamento del circuito, la seconda considera anche la tensione di alimentazione offrendo un ragionevole indice di valutazione del trade-off tra rumorosità e potenza consumata. Sebbene la PEF sia più esaustiva, essa è stata proposta solo recentemente e pertanto la NEF è la principale figura di merito con la quale si sono finora confrontati i lavori in letteratura. L'unico stadio di interesse per la valutazione dell'efficienza è il primo amplificatore, classificato come Low-Noise-Amplifier (LNA), perchè essendo posto a monte della catena di acquisizione e elaborazione del segnale, è il blocco le cui specifiche di rumorosità sono più stringenti.

### 2.1.3 Digitalizzazione

Successivamente all'amplificazione il segnale dev'essere digitalizzato, eventualmente codificato e compresso ed infine trasmesso in modalità wireless ad un ricevitore remoto posto a qualche metro di distanza. L'operazione di digitalizzazione è effettuata dall'ADC (Convertitore Analogico-Digitale) in seguito ad una fase di campionamento dell'ingresso analogico. Il convertitore è in grado di discernere livelli di tensione del segnale di ingresso non inferiori ad una grandezza chiamata bit meno significativo (Least Significant Bit, LSB) e che dipende dalla massima dinamica di tensione (Full Scale Range, FSR) con cui lavora, secondo la relazione:

$$LSB = \frac{FSR}{2^n} \quad (2.3)$$

Il FSR è pari alla tensione di alimentazione  $V_{dd}$  se l'architettura del convertitore è single-ended, oppure a  $2V_{dd}$  se è fully differential. Il valore analogico d'ingresso è quindi approssimato al più vicino multiplo dell'LSB, cui corrisponde una precisa codifica digitale. Questa inevitabile operazione di quantizzazione degrada il rapporto segnale-rumore (SNR) introducendo un rumore (detto rumore di quantizzazione) la cui densità spettrale di potenza riferita all'ingresso del convertitore è pari a [32]:

$$V_{QN}^2 = \frac{LSB^2}{12} \quad (2.4)$$

Perchè la digitalizzazione non peggiori significativamente il SNR sono necessari almeno 8 bit di conversione. Infatti, ammettendo di rilevare in ingresso il minimo segnale disponibile e di sottoporlo ad una verosimile amplificazione  $G_{tot}$  di circa 60dB prima dell'ADC, dovremmo garantire che anche il rumore di quantizzazione riportato all'ingresso dello stadio amplificante sia inferiore allo stesso valore fissato per il rumore dell'elettronica di amplificazione:

$$\frac{LSB^2}{12} \frac{1}{G_{tot}^2} < (5\mu V_{rms})^2 \quad (2.5)$$

relazione dalla quale consegue una fondamentale ma non particolarmente stringente condizione sull'LSB :

$$LSB < 17.3mV \quad (2.6)$$

Sostituendo l'espressione del FSR si può dedurre il minimo numero di bit per la digitalizzazione:

$$n \geq \log_2\left(\frac{FSR}{G_{tot}\sqrt{12}(5\mu V_{rms})}\right) \quad (2.7)$$

Considerando un generico front-end single ended con  $V_{dd}$  pari a 3V, il numero di bit dovrebbe essere superiore a 7.43, pertanto 8 bit sembrerebbe a pieno titolo considerabile il minimo numero di bit necessario a non degradare il rapporto segnale-rumore. In realtà fermarsi a questo punto dell'analisi per fondare il dimensionamento del convertitore potrebbe risultare rischioso. Nel progetto di un front-end analogico e di un ADC infatti entrano in gioco fenomeni di distorsione e non idealità che contribuiscono a degradare ulteriormente la qualità del segnale, pertanto è un criterio più conservativo riferirsi all'ENOB (numero di bit effettivo) dell'ADC piuttosto che al suo numero di bit nominale. L'ENOB è funzione sia del rumore di quantizzazione che della distorsione, come emerge dalla relazione secondo la quale è definito:

$$ENOB = \frac{SNDR - 1.76}{6.02} \quad (2.8)$$

dove SNDR cioè il *Signal-to-Noise Distortion Ratio* è dato da:

$$SNDR = 10\log\left(\frac{V_{in,pp}^2/2}{\sum_{i=2}^{+\infty} \frac{A_{i,harm}^2}{2} + P_{Noise}}\right) \quad (2.9)$$

ed è il rapporto tra la potenza del segnale in ingresso e la somma della potenza delle armoniche e del rumore misurati in uscita attraverso un algoritmo di FFT, mentre il rumore di quantizzazione è responsabile di un fondo di rumore, detto *Noise Floor* (NF) calcolabile in dB come segue:

$$NF = -(6.02n + 1.76 + 10\log(\frac{N_{FFT}}{2})) \quad (2.10)$$

Quest'ultima grandezza è pari all'opposto del rapporto segnale-rumore di un ADC la cui unica non idealità è dovuta al rumore di quantizzazione (ADC ideale) e quando in ingresso è fornita una sinusoide a piena dinamica [32]. La scelta del numero di campioni per il calcolo della FFT deve essere opportunamente studiata, in generale esso deve essere sufficientemente elevato per garantire un basso noise floor ed al contempo multiplo del numero di campioni contenuti in un periodo della sinusoide digitale in uscita per evitare l'equivocazione del segnale. Misurando il SNDR in uscita nell'ipotesi fittizia in cui esso sia il risultato della conversione del segnale da parte di un ADC ideale si ha:

$$SNDR = P_{segnale,out} - P_{noise,eq} \quad (2.11)$$

$$SNDR = -NF_{eq} \quad (2.12)$$

Sostituendo l'espressione del noise floor si ricava infine:

$$SNDR = 6.02 \cdot ENOB + 1.76 \quad (2.13)$$

espressione che, ponendo  $n=ENOB$  (in accordo con l'ipotesi iniziale) corrisponde alla 2.8. Per ottenere un ENOB superiore a 8bit è opportuno eccedere con il numero di bit nominali del convertitore e utilizzare 10bit garantisce un buon margine di sicurezza. Le sorgenti di distorsione e le non idealità dei circuiti adibiti all'amplificazione e alla digitalizzazione saranno analizzate più in profondità nei capitoli relativi al progetto dei singoli stadi. Finora sono state introdotte le specifiche generali della conversione da analogico a digitale del segnale, esaminandone le prestazioni solo sulla base dell'informazione ricavabile dal confronto tra segnali d'ingresso e segnali d'uscita. E' opportuno approfondire alcune fondamentali caratteristiche specifiche dell'ADC cui non si può risalire a partire dai segnali ma che sono un riferimento in fase di progetto e che influenzano profondamente SNDR e quindi ENOB. Questi parametri sono gli indici di non linearità del convertitore ossia la DNL (non linearità differenziale) e la INL (non linearità integrale). La caratteristica statica di conversione di un ADC(Fig. 2.4) indica la codifica digitale di uscita (o la tensione corrispondente frutto di una ulteriore conversione D/A a valle) elaborata dall'ADC come funzione del valore di tensione fornito all'ingresso e



conseguente alla fase di campionamento di un segnale analogico; essa mostra la tipica morfologia a gradinata dovuta alla quantizzazione della dinamica di uscita.

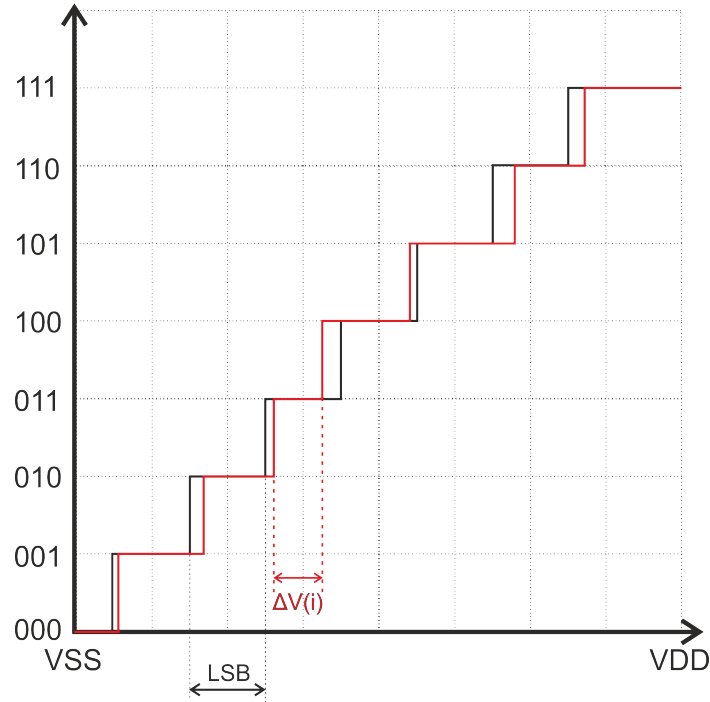


Figura 2.4: Esempio di caratteristica di un ADC a 3 bit (in rosso) a confronto con quella ideale (in nero).

La non linearità differenziale propriamente detta coincide con la massima differenza di dislivello tra le tensioni d'ingresso di due codifiche successive della conversione e il corrispondente valore teorico ( che si avrebbe nella caratteristica di un ADC ideale) pari a 1 LSB. La DNL si calcola come indicato dall'equazione 2.14:

$$DNL = \max(DNL(i)) = \max\left(\frac{\Delta V_{in, reale}(i) - \Delta V_{in, id}(i)}{\Delta V_{in, id}(i)}\right) \quad (2.14)$$

$\Delta V_{in, id} = LSB$  se l'architettura dell'ADC non è soggetta ad un errore intrinseco di guadagno. E' opportuno specificare che frequentemente con il termine DNL non si indica il valore massimo bensì la caratteristica che è possibile tracciare calcolando per ogni salto di livello la differenza rispetto al salto ideale. Per la salvaguardia della monotonicità della caratteristica è bene che la DNL non superi mai il valore assoluto di 0.5. La non linearità integrale è invece rappresentata dal massimo scostamento presente tra un punto della caratteristica dell'ADC e il corrispondente della caratteristica ideale. Come avviene per la DNL, anche per la INL è comune tracciare una *caratteristica della INL* che mostra il suo andamento per ogni codifica e per un suo calcolo accurato è necessario assicurarsi

di averla depurata da un eventuale errore di guadagno che potrebbe altrimenti essere erroneamente confuso con un fenomeno di non-linearità.

Al fine di poter confrontare le prestazioni dei vari convertitori è utile disporre di un indice o una figura di merito atta a normalizzarne le prestazioni secondo criteri di efficienza energetica. Il più importante di questi indici dipende dall'ENOB, e quindi indirettamente dalla DNL, ed è la figura di merito così definita:

$$FoM = \frac{P_{tot}}{2^{ENOB} f_{sampling}} \quad (2.15)$$

Dalla sua espressione si capisce che si tratta di un indicatore di efficienza che descrive quanto il convertitore è in grado di sostenere determinate prestazioni di risoluzione e velocità consumando poca energia. In figura 2.5 è indicata l'evoluzione delle FoM degli ADC nell'ultima decade e mostra come lo scaling tecnologico ed in particolare la possibilità di integrare capacità più piccole e di realizzare logiche più compatte e rapide sono stati elementi chiave nel determinare il miglioramento dell'efficienza dei convertitori. Nel capitolo relativo al progetto dell'ADC, dove saranno spiegati i dettagli del progetto e l'architettura scelta, particolare spazio sarà dedicato al legame tra le caratteristiche di quest'ultima e i parametri di qualità appena introdotti.

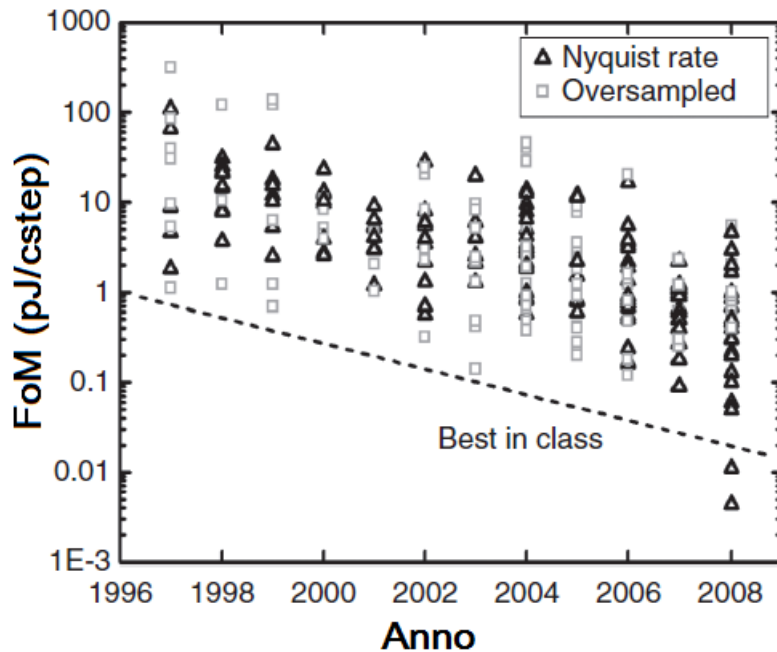


Figura 2.5: Evoluzione della FoM degli ADC nell'ultima decade [5].

### 2.1.4 Alimentazione

Esiste una intensa attività di ricerca che avanza parallelamente a quella sui sistemi di neural monitoring ed atta a sviluppare sistemi di alimentazione ad induzione elettromagnetica per garantire la possibilità di alimentare in modalità wireless dispositivi impiantabili nel corpo umano. I vantaggi di un tale sistema di trasferimento di energia sono legati all'abbandono delle batterie e sono essenzialmente tre: la riduzione dell'ingombro, la riduzione del peso e una durata di vita virtualmente infinita. Sistemi di questo tipo sono tipicamente costituiti da un modulo esterno con un amplificatore di potenza in classe E che pilota un'induttore, ed un ricevitore on chip a valle di un altro induttore. Quest'ultimo, il secondario, è elettricamente accoppiato con il primo e seguito da un circuito di ricezione. La frequenza di risonanza (accoppiamento) del secondario può essere appositamente regolata grazie all'impiego di un condensatore di shunt. Il circuito di ricezione oltre a convertire il segnale di energia in un'opportuna tensione di alimentazione, può anche essere progettato per raccogliere poche informazioni come eventuali bit di configurazione di sistema, trasmesse attraverso lo stesso collegamento. Il trasmettitore genera un campo in radiofrequenza nella banda dei megahertz che induce una tensione armonica ai capi del secondario. Un limitatore RF è utilizzato per proteggere da sovratensioni il circuito a valle del ricevitore, mentre un raddrizzatore a larga banda converte il segnale sinusoidale ricevuto in una tensione continua, infine regolata da un convertitore DC-DC.

Uno dei parametri più critici del sistema di alimentazione è la frequenza della portante. Le alte frequenze hanno la possibilità di trasferire un maggior quantitativo di potenza e permettono di ridurre le dimensioni di induttore e capacità di shunt. Lo svantaggio sta nella minore efficienza ottenibile a causa dell'aumentato consumo dei circuiti di ricezione, del minore range di trasmissione e del maggiore assorbimento di energia da parte dei tessuti biologici circostanti il ricevitore. Tipicamente la portante è scelta nella banda 1-10MHz, valore compatibile con la condivisione di clock integrati con altre unità funzionali del sistema (es. ADC).

Come accennato in precedenza, il link induttivo può essere utilizzato anche per la trasmissione di alcuni dati dall'esterno all'interno del chip. Questi potrebbero essere costituiti da bit per la selezione di specifici canali, per la programmazione del guadagno oppure per la regolazione della risoluzione dell'ADC e che possono essere trasmessi ad un bit-rate basso, compatibile con la banda della trasmissione. Nel sistema di Harrison [13] è presentato un blocco di alimentazione wireless operante a 2.56MHz e con modulazione ASK (Amplitude-Shift-Keying) che trasferisce 8mW consumando 1.5mW. Il margine di miglioramento nella realizzazione di sistemi di questo tipo è ancora grande ed i vantaggi che il loro impiego comporterebbe nel campo dei dispositivi impiantabili indicano con ogni probabilità un loro sviluppo crescente nei prossimi anni.

### 2.1.5 Trasmissione dei Dati

Per rispettare le specifiche termiche di un dispositivo impiantabile dalla dimensione verosimile di 4mmx4mmx1.5mm, il sistema complessivo non dovrà consumare più di 10-

12mW cioè  $187\mu W$ /canale assumendo la presenza di almeno 64 canali di acquisizione. In seguito ai progressi nello sviluppo di integrati per monitoraggio neurale a basso consumo di potenza, tuttavia, il limite per il confronto con le più recenti proposte dei gruppi di ricerca internazionali [18] è sceso al di sotto dei  $50\mu W$  per canale. La connessione wireless, in termini di tecnica di trasmissione e di elaborazione dei dati, è il fattore più influente sul consumo del chip. E' già stato spiegato come il link induttivo attraverso il quale è fornita l'alimentazione può essere utilizzato anche per il trasferimento dati, ma questa è una soluzione praticabile solo per bassi valori di bit-rate [33]. Una soluzione più efficace prevede la separazione di *downlink* (alimentazione, pochi bit di controllo), praticabile attraverso *inductive link* dall'*uplink* (trasferimento dei dati acquisiti) effettuato a frequenze superiori. La progettazione di trasmettitori a basso consumo ed alta capacità è una crescente sfida e costituisce una disciplina in rapida evoluzione. Per applicazioni a corto raggio esistono tre opzioni principali: trasmissione ultrawideband (UWB), trasmissione infrarossa (IR) e la classica trasmissione RF narrowband. La figura di merito di queste tecniche nel contesto di applicazioni impiantabili è l'energia spesa per ogni bit trasmesso.

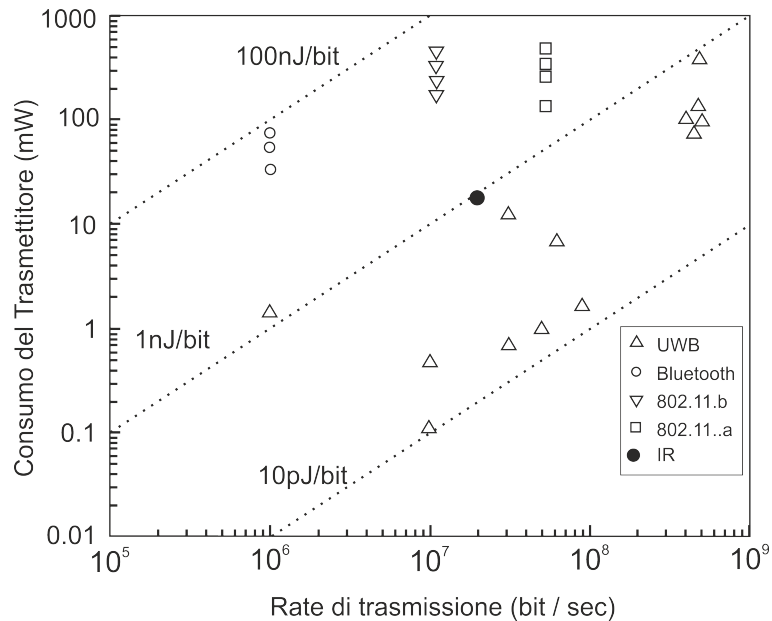


Figura 2.6: Confronto tra le prestazioni di alcuni recenti sistemi di comunicazione wireless

In figura 2.6 sono illustrate le prestazioni di diversi sistemi wireless recentemente pubblicati. I sistemi UWB dimostrano di avere le migliori figure di merito, di circa  $10pJ/bit$ , raggiungendo range di trasmissione superiori ai 10 metri [35] e proponendosi ai progettisti come l'alternativa più allettante. Tali circuiti infatti, trasmettono impulsi ampi nanosecondi con componenti in frequenza nella banda 3.1-10.6GHz, e sebbene la loro generazione richieda specifiche conoscenze progettuali, non necessita né di PLL né di particolari filtri e il power-amplifier può operare in modo intermittente per

risparmiare energia. Altre soluzioni come trasmettitori per protocolli WLAN (IEEE 802.11a/b) e Bluetooth non risultano sufficientemente efficienti, con consumi dell'ordine dei 100nJ/bit. Per quanto riguarda i sistemi IR, oltre a richiedere un'energia per singolo bit due ordini di grandezza superiore in media a quella richiesta dai sistemi UWB, necessitano dell'allineamento geometrico tra emettitore e ricevitore offrendo range di trasmissione limitati, nell'ordine dei centimetri [34]. Studi preliminari di applicabilità in campo biologico di sistemi UWB hanno dimostrato che i parametri di assorbimento di radiazioni a frequenza così alta da parte dei tessuti sono compatibili con l'impiego medico [36]. Recentemente i lavori di Chae et al. [18] e Meng et al. [31] presentano sistemi completi di neural recording che sfruttano la tecnica di trasmissione UWB.

Un fattore determinante il consumo richiesto dall'unità di trasmissione è il bit-rate che è necessario trasmettere e questo dipende a sua volta dalla quantità di informazione che si vuole catturare dal segnale sorgente e dalla modalità con cui la si codifica. Dal punto di vista del neuroscienziato avere a disposizione in tempo reale tutta l'informazione carpita dagli elettrodi di ogni canale sarebbe la soluzione migliore, se poi questi fossero disponibili a distanza di metri dal punto di rilevazione si avrebbero ulteriori vantaggi, tuttavia il progettista sa che il rischio di non rispettare i requisiti di consumo è molto alto. Al trade-off informazione-consumo-range i principali gruppi di ricerca internazionale hanno proposto negli anni soluzioni molto differenti. Come proposto da Harrison nel suo primo sistema, è possibile trasmettere solo l'informazione relativa all'istante di occorrenza degli spikes, ma in questo modo si perde la possibilità di capire, a partire dalla morfologia dei segnali, da quale neurone proviene la scarica. Anche riduzione della banda di acquisizione e della frequenza di campionamento e la trasmissione dei soli PA sono tecniche che permettono una riduzione del bitrate con una perdita solo parziale dell'informazione di morfologia. Un'ulteriore via percorribile è la ricerca di alternative alla tipica conversione A/D per ridurre i bit senza virtualmente perdere informazione, come cerca di fare Ghovanloo et al [30] nel suo sistema. Tutte queste soluzioni oltre a comportare la perdita di informazione, complicano il sistema introducendo la necessità di circuiti per la spike-detection ed in ogni caso sono difficilmente applicabili quando si desidera trasmettere anche le caratteristiche dei lenti LFPs. Una analisi dettagliata dei principali sistemi presentati in letteratura, con una particolare attenzione alle tecniche di trasmissione si può trovare nella prossima sezione. Prima di proseguire, è bene avere presente la relazione che intercorre tra alcuni parametri del sistema come il numero di canali, la frequenza di campionamento e la risoluzione della conversione A/D ed il data rate (o bit-rate). Esso infatti è una grandezza determinante per le prestazioni di consumo del sistema e, come si avrà modo di capire nel corso del capitolo, è spesso al centro di importanti scelte progettuali. In un sistema ad N canali campionati alla Nyquist ciascuno ad una frequenza  $f_{sampling}$ , il bit-rate è dato approssimativamente dal prodotto:

$$bitrate = f_{sampling} \cdot N_{channels} \cdot n_{bit} \quad (2.16)$$

Questo calcolo è evidentemente al lordo dei bit di overhead, cioè tutti quei bit che è necessario aggiungere allo stream per permettere al ricevitore di ricostruire la sequenza di trasmissione, il canale dal quale proviene ed altre eventuali informazioni complementari che si desidera trasmettere. La 2.16 non considera neppure l'effetto di eventuali tecniche di codifica o compressione dei dati ma rappresenta comunque una buona stima del bitrate che il trasmettitore deve sostenere.

## 2.2 I Sistemi Esistenti

In questa sezione sono brevemente presentati i SoCs che per le prestazioni ottenute e per le novità introdotte dal punto di vista circuitale e architetturale possono essere considerati i lavori più rappresentativi dello stato dell'arte nel campo dei sistemi di monitoraggio neuronale e trasmissione completamente integrati. Il metro di paragone tra gli esemplari proposti è costituito dal trade-off tra l'informazione fruibile dall'utilizzatore, la distanza alla quale essa è disponibile ed il consumo richiesto. La ricerca del compromesso ottimale si è spesso tradotta nello studio di originali procedure per la compressione dei dati: per questo i lavori a seguire presentano spesso soluzioni simili relativamente ai blocchi analogici di amplificazione e filtraggio, mentre la differenza cresce se si analizzano i blocchi di digitalizzazione e quelli di trasmissione.

### 2.2.1 Il Sistema di Harrison

Il sistema messo a punto da Harrison et al.[13] nel 2007, è il capostipite dei sistemi di monitoraggio neuronale; la sua struttura è illustrata in figura 2.7. Realizzato in tecnologia CMOS  $0.5\mu\text{m}$  (3 Metals, 2 Poly), è alimentato a 3V e progettato per elaborare il segnale proveniente da uno Utah Array di 100 microelettrodi. A valle di ciascun microelettrodo vi è un canale di acquisizione analogico la cui banda, compresa tra 300 Hz e 1kHz, permette di rilevare i soli potenziali d'azione (APs) e non i LFPs. Questo deriva dalla scelta a monte di operare la massima riduzione di informazione compatibile con l'applicazione al fine di rilassare i requisiti di efficienza del sistema di trasmissione. Anche l'informazione di morfologia dei potenziali d'azione infatti, a valle dell'acquisizione viene persa facendo passare i tracciati attraverso uno stadio di comparazione a soglia programmabile alla cui uscita è conservata unicamente l'informazione relativa al tempo di occorrenza dei singoli spikes.

Uno solo dei 100 canali può essere selezionato tramite un MUX per essere campionato alla nyquist da un SAR ADC a 10 bit allo scopo estrarne la morfologia. Il bit-rate complessivo con cui lavora il trasmettitore è pari a 330kbit/s ed il range di trasmissione si attesta intorno a 13cm. La frequenza di trasmissione scelta è di 433MHz per non soffrire dei fenomeni di assorbimento tissutale che si possono verificare al di sopra del GHz e per operare vicino allo standard di comunicazione per apparecchiature impiantabili (MICS, Medical Implant Communication System) la cui banda è 402-405MHz. Il consumo specifico complessivo ottenuto è di  $135\mu\text{W}/\text{canale}$ .

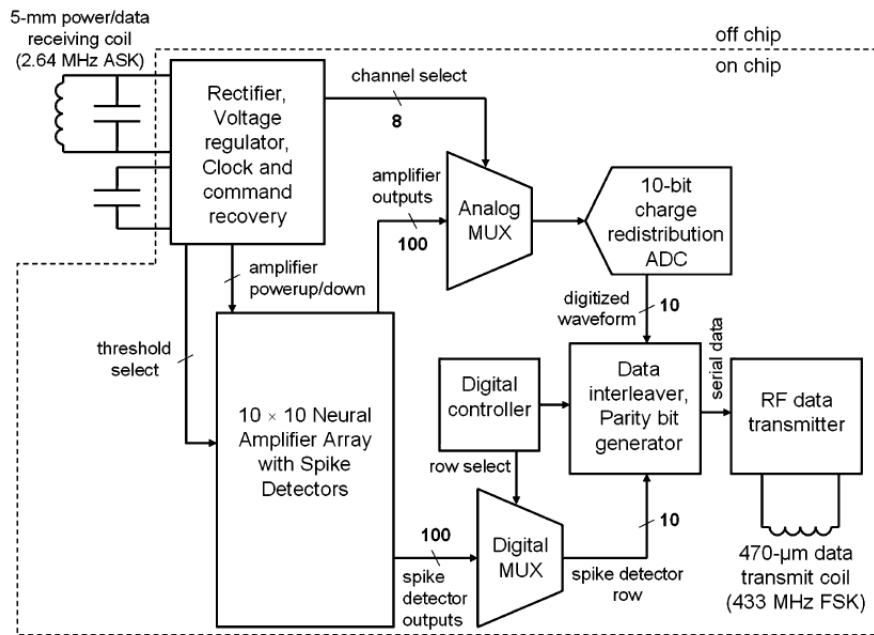


Figura 2.7: Il sistema in [13].

## 2.2.2 Il Sistema dell'Università della California Santa Cruz.

Chae et al. [18], della California University, Santa Cruz, realizza un sistema CMOS  $0.35\mu\text{m}$  a 128 canali, alimentato tra  $-1.65\text{V}$  e  $+1.65\text{V}$  il cui consumo complessivo è di  $6\text{mW}$ , pari a soli  $47\mu\text{W}$  per canale. Il suo Lo schema generale è mostrato in figura 2.8 Il chip permette di acquisire a discrezione l'intero tracciato di ogni canale ed in tempo reale, oppure di effettuare una selezione degli APs successiva alla digitalizzazione. Il data-rate trasmesso è in entrambi i casi molto elevato, superiore a  $80\text{Mbit/s}$ . A fronte di questi numeri il risparmio energetico appare considerevole ma è ottenuto a spese di un ridotto range di trasmissione, prossimo al centimetro. Una distanza così esigua compromette inevitabilmente la possibilità di impiego del chip e rappresenta il principale punto debole di un sistema che, al contrario, offre soluzioni interessanti ed efficienti per tutti gli altri blocchi funzionali.

Il Front-End di acquisizione e digitalizzazione garantisce una banda in grado di catturare l'informazione dei LFPs insieme agli APs e presenta una struttura parcellizzata in cui ciascuno degli 8 ADC SAR (la cui risoluzione è programmabile tra 6 e 9 bit) utilizzati serve 16 canali di acquisizione attraverso un analog-MUX. A valle della digitalizzazione può essere effettuata una *on-the-fly spike feature extraction* che consiste nell'individuazione degli spikes e nella loro classificazione per canale e neurone di provenienza possibile grazie ad un DSP che effettua queste operazioni in parallelo per garantire il funzionamento in tempo reale. Il flusso di informazioni a valle della classificazione consiste in una serie di bit che contengono sia informazioni di segnale che bits di overhead nel ca-

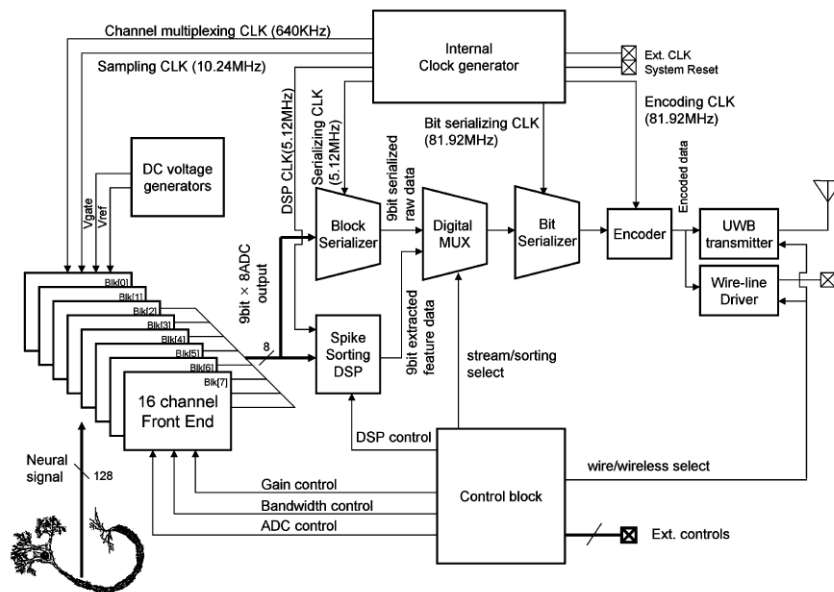


Figura 2.8: Sistema di neural monitoring di Chae et al. [18].

so di funzionamento con feature extraction, oppure nel completo tracciato dei segnali campionati a 40kSps e dai relativi bit di overhead (inizio sequenza, selezione del canale, ecc.). L'unità di trasmissione è di tipo IR-UWB e trasmette un data-rate di 90Mbit/s codificato alla Manchester, che può essere modulato in modalità OOK (on-off keying) oppure PPM (pulse-position modulation) per la generazione degli impulsi.

### 2.2.3 Il Sistema WINeR

Uno dei fattori che maggiormente incide sul bit-stream in sistemi il cui obiettivo è la trasmissione in tempo reale del tracciato *raw* di ogni canale è la risoluzione con cui ciascuno di questi è digitalizzato. Ghovanloo et al [30] propongono una procedura di Analog-to Time Conversion (ATC) alternativa alla digitalizzazione il cui risultato è una drastica diminuzione del data-rate in uscita, equivalente a 640kbaud/s. Battezzato con il nome WINeR (Wireless Neural Recording), il sistema è alimentato per induzione tra -1.5V e +1.5V e acquisisce a flusso continuo l'informazione di 32 canali provenienti da un MEA. L'amplificazione avviene nella banda compresa tra 0.1Hz e 1kHz (quindi interessa anche i LFPs) grazie a due stadi che realizzano un guadagno complessivo variabile da 66dB a 78dB e che possono essere messi in stand-by quando se non devono essere utilizzati.

La conversione A/T è effettuata comparando il segnale all'uscita del secondo amplificatore di ciascun canale con un'onda triangolare e rilevando gli istanti (quindi anche gli intervalli di tempo) in cui il primo è maggiore o minore della seconda. A valle della conversione un MUX ed un registro circolare sono utilizzati per completare la compressione



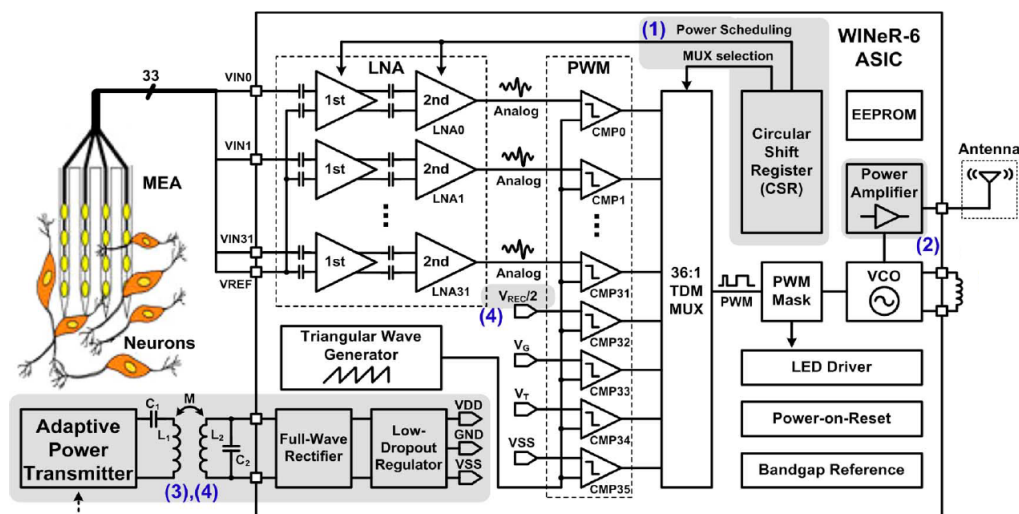


Figura 2.9: Il Sistema WINeR [30].

dell'informazione attraverso un'operazione di Time-Division-Multiplexing. La riduzione del numero di transizioni digitali rende l'intero sistema meno sensibile a fenomeni di cross-talk e feed-through. Il consumo complessivo del sistema è di  $5.85\text{mW}$ , cioè circa  $182\mu\text{W}/\text{canale}$  ed è evidentemente molto basso considerando la quantità di informazione trasferita. A queste due ultime considerazioni è bene accompagnare una precisazione: la procedura di compressione utilizzata complica notevolmente l'architettura del receiver, ponendo problemi critici rispetto alla qualità del segnale trasmissibile. Quest'ultimo infatti non deve solo effettuare una elaborazione digitale dei dati, come avverrebbe in un sistema classico, ma anche una elaborazione analogica atta a individuare con precisione gli intervalli temporali di modulazione alle varie frequenze. Tale informazione è facilmente soggetta a degrado durante la trasmissione (interferenze, rumore) tanto più quanto maggiore è la distanza alla quale si desidera trasmettere e potrebbe essere ischiosa già al di sopra di poche decine di centimetri. Nonostante ciò, il range di trasmissione dichiarato per il WINeR è di circa 1m.

#### 2.2.4 Il Sistema Genov-Abdelhalim.

Il SoCs presentato nel 2011 da R. Genov e K. Abdelhalim dell'Università di Toronto, come il precedente, merita particolare attenzione grazie all'originale metodo di compressione dei dati. A differenza del precedente, il sistema in [42] non rinuncia né al tradizionale campionamento alla nyquist né alla digitalizzazione pur ottenendo un bit stream limitato a  $1.5\text{Mbit/s}$  e contenente tutte le informazioni di interesse clinico (LFPs, APs). La chiave ma al tempo stesso il difetto di questo risultato sta nella parziale rinuncia alla conservazione della morfologia dei potenziali d'azione, con il taglio delle frequenze superiori ad  $1\text{kHz}$ . Questo filtraggio non è effettuato dal blocco analogico

a monte del convertitore, bensì durante la procedura stessa di digitalizzazione, in cui l'ADC SAR a redistribuzione di carica può essere utilizzato per realizzare un filtraggio digitale, attraverso in fase di campionamento. Con opportune modifiche della logica del convertitore è infatti possibile effettuare moltiplicazioni della tensione memorizzata ai capi dell'array di capacità tipico dell'architettura. Il Data-rate finale è trasmesso a 915MHz con modulazione FSK. Il consumo complessivo è di 5.03mW, ovvero  $79\mu\text{W}$  per ogni canale. Il sistema offre infatti la possibilità di registrare da 64 canali; il chip d'implementazione è realizzato in tecnologia CMOS  $0.13\mu\text{m}$ .

### 2.2.5 Il Sistema del Politecnico di Milano

Nel contesto dei sistemi completi per il monitoraggio neurale, il Politecnico di Milano ha contribuito realizzando un sistema di registrazione neurale a 64 canali alimentato a 3V in tecnologia CMOS AMS  $0.35\mu\text{m}$  [39] i cui dati sono riassunti e confrontati con gli altri principali sistemi pubblicati in letteratura nella tabella 2.1. Il sistema utilizza un microwire MEA commerciale prodotti da Tucker-Davis Tech. e garantisce l'amplificazione e il filtraggio grazie ad una serie di due amplificatori (LNA e secondo stadio) single ended il cui guadagno è fisso e pari a circa 70dB di cui rispettivamente di 37dB garantiti dall'LNA e i restanti 33dB dal secondo stadio. La banda di acquisizione è stata scelta tra 300Hz e 15kHz, pertanto solo gli APs sono rilevabili. Il front-end complessivo di acquisizione e digitalizzazione è stato progettato secondo lo schema di figura 2.2b, con un MUX analogico, un ADC SAR a redistribuzione di carica e un sistema di buffers in grado di pilotare sia la capacità parassita delle lunghe piste di connessione (buffers di linea, uno per canale) sia la capacità d'ingresso del convertitore (buffer d'ingresso all'ADC).

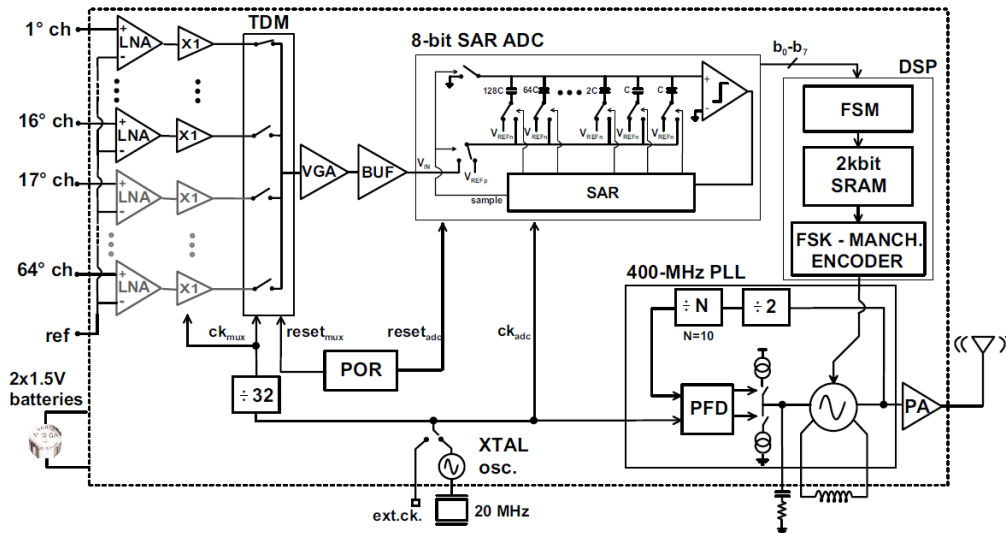


Figura 2.10: Il Sistema Polimi [39].

Un amplificatore a guadagno variabile è posto a valle del MUX, prima del buffer d'ingresso dell'ADC per permettere eventualmente un'incremento di guadagno (Fig. 2.10). I bit d'uscita del convertitore sono serializzati e elaborati da una logica per effettuare il riconoscimento degli spikes e la memorizzazione dei relativi bit, gli unici destinati alla trasmissione. Una volta modulati (FSK) i bit sono trasmessi con un sistema a banda stretta (3MHz) centrata a 400MHz (in pieno standard MICS) fino a 4m di distanza (bit-rate pari a 1.25Mbit/s). Il consumo specifico complessivo ottenuto, di  $269\mu\text{W}/\text{canale}$ , può apparire elevato se confrontato con i precedenti sistemi. In realtà ben  $164\mu\text{W}$  dei 269 per canale sono dovuti alla scelta di un sistema di trasmissione a raggio più lungo e più affidabile ( il BER è inferiore a  $10^{-5}$ ) di tutti gli altri sistemi presentati.

Tabella 2.1: Confronto tra le prestazioni dei principali sistemi integrati di monitoraggio neuronale

Autore	Harrison '07	Chae '08	Bonfanti '09	Ghovanloo '10	Genov '11
Ref.	[13]	[18]	[39]	[30]	[42]
Tecn.	0.5 $\mu$ m	0.35 $\mu$ m	0.35 $\mu$ m	0.5 $\mu$ m	0.13 $\mu$ m
VDD	3.3V	+/-1.65V	+/- 1.5V	+/- 1.5V	1.2V
Canali	100	128	64	32	64
$G_{tot}$	60dB	57-60dB	65-83dB	68-78dB	54-60dB
$V_{rms,noise,IN}$	5.1 $\mu$ V $_{rms}$	4.9 $\mu$ V $_{rms}$	3.05 $\mu$ V $_{rms}$	9.3 $\mu$ V $_{rms}$	6.5 $\mu$ V $_{rms}$
Freq. di TX	433MHz	4GHz	400MHz	915MHz	915MHz
Modulaz.	2FSK	UWB	2MC-FSK	2FSK (PWM)	FSK-OOK
Tipo	recording	str./rec.	recording	PWM	str./FIR
Data-rate	330kbit/s	90Mbit/s	1.25Mbit/s	640kbaud/s	1.5Mbit/s
TX BW	$\approx$ 0.8MHz	$\approx$ 38MHz	$\approx$ 3MHz	$\approx$ 3MHz	NA
Consumo	135 $\mu$ W/ch.	47 $\mu$ W/ch.	269 $\mu$ W/ch.	220 $\mu$ W/ch.	79 $\mu$ W/ch.
TX Range	13cm	NA	4m	$\approx$ 1m	$\approx$ 1m

I bit in uscita dall'ADC , una volta serializzati, sono inviati dapprima ad uno stadio digitale adibito all'individuazione ed alla selezione degli spikes ed infine ad un blocco di modulazione (FSK) e trasmissione narrowband operante con un data-rate a 1.25Mbit/s. Questo blocco, costituito da e da un Power Amplifier, permette una trasmissione a lunga distanza ma domina il consumo del chip richiedendo da solo 105 dei 269 $\mu$ W/ch. che caratterizzano il consumo specifico totale.

## 2.3 Motivazioni e Obiettivi per un Nuovo Progetto

I limiti dei risultati di ricerca fin qui riportati mette in evidenza come lo sviluppo di sistemi multicanale per la registrazione di segnali neuronali sia ancora all'inizio, e molta strada dev'essere fatta per arrivare ad una generazione di sistemi affidabili ed a lungo raggio realmente interessanti per applicazioni impiantabili. A partire dal 2005 sono state infatti proposte diverse realizzazioni, tuttavia l'impatto applicativo è stato limitato dalle difficoltà di coniugare bassi consumi, alta capacità di trasmissione e portate adeguate. Inoltre, scarsa attenzione è stata fin qui posta alla compatibilità tra le bande di frequenza impiegate e quanto ammesso alle normative, e questo è imputabile all'assenza di una linea guida di progettazione metodologicamente affermata. Il nuovo progetto ha l'ambizione di dimostrare la realizzazione di un nuovo sistema adeguato alla rilevazione di segnali neurali in-vivo , impiegabile sia nella pratica neurochirurgica che in esperimenti di neuroscienze. Il sistema , esattamente come la sua precedente versione, dovrà interfacciarsi con un MEA a 64 canali e dovrà garantire la trasmissione continua ed in tempo reale (modalità streaming) dei dati acquisiti e digitalizzati con 10 bit utilizzando un sistema di trasmissione a banda larga capace di trasmettere a un data rate che,

in accordo alla relazione 2.16 e alle specifiche sopraelencate, non deve essere inferiore a 19.2Mbit/s . Tale cifra, ben superiore all'equivalente del vecchio sistema, dovrà essere garantita pur limitando il consumo ad un valore inferiore ad 1mW, corrispondente a  $15\mu\text{W}/\text{canale}$  ed un range di trasmissione di qualche metro. La configurazione adottata come riferimento per il sistema prevede una alimentazione a batteria , mentre l'unità di ricezione sarà connessa ad alimentazione fissa.

La sfida proposta, oltre a costituire un progresso applicativo specifico, contribuisce all'investigazione di architetture e topologie circuitali efficienti per la sensoristica integrata, con specifico riferimento allo studio di:

- Stadi di amplificazione a bassa tensione e minimo rumore ( $\leq 5\mu V_{rms}$  in 10kHz ) e dissipazione di potenza ( $\leq 2\mu\text{W}$ );
- Stadi di campionamento e conversione a minimo consumo ( $\text{FoM} \leq 10\text{fJ}/\text{cstep}$ ) con velocità e risoluzione (10bit) intermedie;
- Trasmettitori ad alta efficienza ( $\leq 35\text{pJ}/\text{bit}$ ) per comunicazione ad alta capacità (20Mbps) e range dell'ordine di qualche metro;

Il motivo della ricerca di una trasmissione continua di dati risiede nell'elevato contenuto informativo contenuto nei LFPs, non rilevati né trasmessi in un sistema di selezione dei soli spikes e che sono indispensabili soprattutto nell'ottica dell'impiego neurochirurgico. L'integrazione di questo sistema dovrà essere effettuata in tecnologia CMOS 130nm con tensione di alimentazione non superiore a 1.2V, ben al di sotto di quella prevista dal precedente chip. In realtà, come vedremo nel prossimo paragrafo, l'alimentazione sarà di 0.5V per ridurre il più possibile il consumo. Il progetto di un chip con simili caratteristiche non può prescindere dall'affrontare le problematiche tipiche della cosiddetta elettronica *Ultra-Low Power*. L'obiettivo sarà perseguito sia attraverso l'ottimizzazione dell'architettura di sistema, sia migliorando rispetto allo stato dell'arte le prestazioni dei singoli blocchi.

### 2.3.1 Il Front-End di Acquisizione e Digitalizzazione

Sebbene il vecchio sistema presentasse globalmente ottime caratteristiche per quanto riguarda la qualità del segnale, l'efficienza del front-end di digitalizzazione e il range di trasmissione, il suo consumo non è dei più contenuti e non è in grado di catturare i Local Field Potentials. Alla luce dei sempre migliori risultati pubblicati nel panorama della bioelettronica internazionale e delle possibilità aperte dal consolidamento della ricerca nel campo dell'elettronica low-power, è stato scelto di progettare ex-novo un sistema in grado di proporre la migliore soluzione possibile al trade-off informazione-consumo-distanza, cercando di coniugare la trasmissione dell'intera informazione possibile, LFPs compresi , da 64 canali e con le migliori prestazioni di consumo e range di trasmissione

che si possono ottenere da un sistema UWB. Uno schema generale del sistema in cui il Front-End si colloca è mostrato in Fig. 2.11.

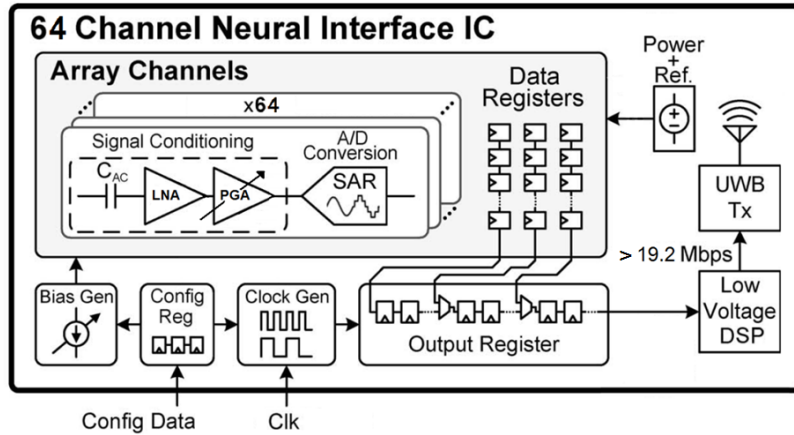


Figura 2.11: Il Nuovo Sistema.

La tensione di alimentazione del nuovo sistema sarà 0.5V, tale scelta è giustificata dalle prospettive di risparmio che introduce, essendo la potenza in prima approssimazione proporzionale a  $V_{DD}$ . A questo punto è lecito chiedersi se una scelta simile non porti in realtà verso l'inasprimento di alcune specifiche e problemi progettuali che non permettano in realtà di ottenere il risultato sperato senza penalizzare il SNR. Ammettendo di poter disporre di un trasmettitore in grado di operare efficientemente a 0.5V, ci si può concentrare sul front-end di acquisizione e digitalizzazione che essendo posto a monte della conversione è lo stadio più critico. In linea di principio un convertitore analogico digitale realizzato in tecnologica scalata e per lavorare a frequenze e risoluzioni intermedie o basse può solo beneficiare della riduzione di tensione alimentazione. Esiste in realtà il grande inconveniente di riuscire a sintetizzare un LSB inferiore ad 1mV, ma questa analisi non si propone di considerare le difficoltà progettuali, bensì di giustificare una scelta generale. La dinamica di uscita dello stadio amplificante è la specifica più penalizzata dalla riduzione dell'alimentazione, tuttavia ciò non ha ripercussioni sulla qualità del segnale digitalizzato. Al contrario, sia il rapporto SNR che il prodotto tra potenza e rumore riferito in ingresso non cambiano, come spiegato nel ragionamento che segue. Il guadagno complessivo di amplificazione, che è generalmente dimensionato per sfruttare tutta la dinamica ed è approssimativamente pari a:

$$G \approx \frac{V_{DD}}{A_{IN}} \quad (2.17)$$

con  $A_{IN}$  ampiezza del segnale all'ingresso. Si consideri ora il rumore di quantizzazione riportato all'ingresso dell'amplificatore, si può scrivere:

$$V_{QN,in} = \frac{LSB^2}{12G^2} \approx \frac{A_{IN}^2}{12 \cdot 2^{2n}} \quad (2.18)$$

Cioè in prima approssimazione esso non dipende dalla tensione di alimentazione scelta a patto di poter sempre dimensionare il guadagno in funzione dell'ampiezza  $A_{IN}$  del segnale utile alla sorgente. Il rumore di quantizzazione non è però l'unica sorgente di rumore, esiste infatti anche il rumore sovrapposto all'ingresso. Questo generalmente è la combinazione del rumore introdotto dallelettronica di amplificazione e del rumore sovrapposto alla sorgente. In entrambi i casi un guadagno dimensionato adeguatamente per non saturare la catena non peggiora il SNR. L'architettura scelta per il nuovo sistema è il risultato di un accettabile compromesso tra la complessità progettuale e la riduzione del consumo, prevedendo la soluzione di un ADC per canale (Fig. 2.2a) che elimina la necessità di multiplexer analogici ed anche quella dei buffers di linea e del buffer di ingresso dell'ADC. La ridotta dinamica offerta dall'alimentazione scelta ha influenzato la decisione finale di progettare l'intero canale con topologie completamente differenziali piuttosto che single-ended, come invece è accaduto nel precedente progetto. A queste condizioni si impone la minimizzazione del consumo dei tre blocchi successivi del canale, cioè rispettivamente un amplificatore a basso rumore (LNA) e guadagno fisso pari a 40dB, un amplificatore a guadagno programmabile (PGA) tra 0dB e circa 20dB, e un ADC (Fig. 2.12) SAR (ad approssimazioni successive) a redistribuzione di carica. La programmabilità del guadagno in questo sistema come in molti altri già presentati è consigliabile per evitare la saturazione del canale; l'ampiezza dei segnali al suo ingresso infatti varia molto dipendentemente dalla localizzazione nel tessuto neuronale del corrispettivo elettrodo. La scelta di un convertitore SAR a redistribuzione di carica è giustificata dalla singolare compatibilità di questa architettura con applicazioni che richiedono risoluzioni intermedie di 8-12bit, velocità medio-basse e soprattutto consumi molto contenuti. Assumendo il massimo consumo permesso per l'intero chip pari ad 1mw, e attribuendone al blocco di trasmissione il 70% , il front-end del singolo canale deve poter funzionare con meno di  $4\mu W$  e di questi si possono attribuire approssimativamente  $2\mu W$  all'LNA,  $1\mu W$  al PGA ,  $0.5\mu W$  all'ADC ed i restanti  $0.5\mu W$  per eventuali circuiti di servizio (reti di polarizzazione, clock).

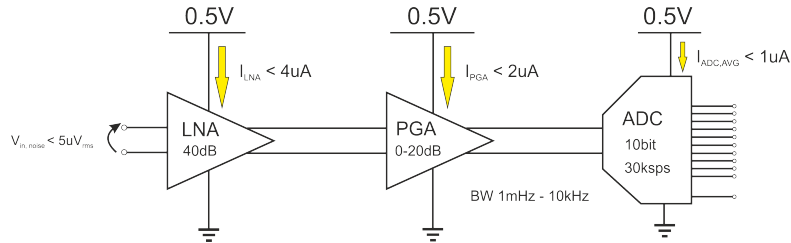


Figura 2.12: Architettura del Front-End di digitalizzazione del nuovo sistema e con le principali specifiche schematizzate.

Sebbene le specifiche di velocità del convertitore siano rilassate a 30kSps (per un buon campionamento alla Nyquist) dalla scelta dell'architettura di sistema, le sue specifiche di consumo risultano evidentemente più stringenti ed in ogni caso l'obiettivo del progetto è di realizzare per ogni blocco, ADC compreso, circuiti dalle prestazioni confrontabili con i migliori lavori presenti in letteratura. Maggiori dettagli circa lo stato dell'arte dei singoli blocchi saranno forniti nei capitoli dedicati al loro progetto. Insieme alle unità funzionali del canale, sono stati progettati i riferimenti delle correnti di polarizzazione e dei buffers differential-to-single-ended da utilizzare (degli instrumentation amplifiers , INA) per la lettura dei segnali intermedi della catena, fondamentali per la verifica delle prestazioni del front-end.



## Capitolo 3

# Lo stadio di Amplificazione e Filtraggio

In questo capitolo è illustrato il progetto dei due amplificatori del front-end, dalla fase di analisi delle topologie circuitali, attraverso la loro scelta e il dimensionamento, fino alla realizzazione del layout e all'analisi delle prestazioni complessive. Insieme alle relazioni ed ai risultati direttamente riguardanti il progetto, sono riportati, ove necessario, alcuni approfondimenti teorici utili a comprendere le caratteristiche dei circuiti ed alcuni chiarimenti circa la risoluzione di alcune problematiche di progetto non elementari.

### 3.1 Aspetti Fondamentali

Ogni canale del sistema è costituito da uno stadio di amplificazione e un convertitore analogico digitale. I requisiti fondamentali del primo sono stati già descritti nel precedente capitolo e si possono sintetizzare in tre punti:

- Banda passante tra 100mHz e 10kHz-15kHz circa;
- Consumo inferiore a  $4\mu\text{W}$  per 0.5V di alimentazione;
- Rumore riferito all'ingresso inferiore ai  $5\mu\text{V}_{rms}$

A questi si può aggiungere un requisito relativo al CMRR che dovrebbe essere almeno di 60dB per far sì che i disturbi di modo comune non corrompano il segnale utile all'uscita della catena; la questione del controllo e della reiezione del modo comune sarà chiarita più avanti in una sezione dedicata. Per far fronte alla riduzione di dinamica di uscita conseguente alla scelta di una tensione di alimentazione di soli 0.5V è stato deciso di improntare il progetto dell'intero front-end (stadio amplificante compreso) su topologie fully differential. Le basse frequenze alle quali si attesta l'informazione utile (in particolare i LFPs) inducono a ritenere il soddisfacimento dei requisiti di banda un obiettivo facile. Se si trattasse unicamente di realizzare un filtraggio passa-basso a 10-15kHz tale valutazione potrebbe considerarsi accettabile, tuttavia La necessità di

tagliare la banda al di sotto dei 100mHz potendo contare unicamente su componenti integrati (il cui valore è limitato dall'occupazione di area disponibile nel chip) costituisce un problema la cui risoluzione è tutt'altro che banale. Solo grazie all'utilizzo di componenti attivi polarizzati in sottosoglia, cui è dedicato un paragrafo poco avanti nel testo, è stato possibile ottenere resistenze dell'ordine dei  $T\Omega$ , funzionali a realizzare singolarità inferiori all'Hertz. Quanto al limite superiore di banda, il valore 10-15kHz rappresenta più un riferimento generale che un vincolo da rispettare in modo rigoroso. Il contenuto informativo del segnale infatti è concentrato per la maggior parte al di sotto dei 5kHz e sono molti i lavori in letteratura in cui la banda del front-end non supera 10kHz. Filtraggio e guadagno sono strutturati secondo lo schema di Fig. 2.3(c). Data la variabilità del segnale d'ingresso, l'amplificazione è ripartita tra due blocchi in cascata: un LNA, posto a monte, garantisce un guadagno di 40dB mentre il successivo PGA garantisce il restante, programmabile tra 0dB e circa 20dB. Questi valori di guadagno scelti permettono di sfruttare adeguatamente la dinamica di uscita, poco inferiore a 1V. La possibilità di separare la frequenza di taglio passa alto del LNA da quella del PGA offre anche una particolare occasione per minimizzare il rumore riferito in ingresso. Esso è dato principalmente da due contributi: il rumore termico dell'OTA del LNA e il rumore termico degli pseudoresistori. Il rumore flicker può invece essere reso trascurabile attraverso un adeguato dimensionamento dei MOS d'ingresso del primo amplificatore e grazie anche al fatto che il trasferimento complessivo del front-end filtra le basse frequenze. Il consumo degli amplificatori è essenzialmente dominato dal primo stadio dell'OTA utilizzato per il LNA, che deve garantire un basso contributo di rumore per non deteriorare il SNR di ingresso al di sotto del limite già imposto dal rumore alla sorgente. La ricerca di prestazioni in tal senso conduce spesso a polarizzare i MOS dello stadio di ingresso del LNA in regime di debole inversione. Scelte come queste sono frequenti in progetti le cui specifiche prevedono tensioni di alimentazioni 1.2V, 3V, 3.3V. Quando la tensione di alimentazione è meno di un quarto rispetto ai valori riportati poco sopra, come in questo caso, non solo i transistori dello stadio di ingresso ma tutti i transistori impiegati devono essere polarizzati in debole inversione o addirittura sottosoglia. Tale soluzione permette sia di massimizzare l'efficienza dei MOS sia di utilizzare stadi con 3 o 4 transistori in serie tra i due rail. Il prezzo da pagare è un limite alla corrente di polarizzazione con cui questi possono funzionare con dimensioni relativamente contenute. Per questi motivi e nella prospettiva di ottenere cifre di merito (NEF e PEF, introdotte nel capitolo 2) migliori rispetto al precedente sistema e confrontabili con quelle dei migliori esempi presentati in letteratura, anche l'indagine di topologie circuitali vantaggiose è stata un elemento centrale dell'iter progettuale. Nei prossimi due paragrafi si trovano due approfondimenti sull'utilizzo dei MOS in sottosoglia mentre i dettagli del progetto sono adeguatamente approfonditi nei prossimi paragrafi. L'unico aspetto che si desidera sottolineare, prima di procedere, è la presenza di tutti gli elementi (alimentazione minima, basso rumore, minimo consumo) che rendono questo progetto una sfida nel campo della progettazione di circuiti analogici *Ultra Low-Power*.

### 3.1.1 Circuiti CMOS Sottosoglia

La trattazione approfondita dei modelli utilizzati per il progetto di circuiti con transistori MOS in sottosoglia si trova in [48] e in [25] mentre alcuni fondamentali esempi di circuiti analogici con dispositivi MOS così polarizzati sono presentati in [49]. Lo scopo di questo paragrafo è riportare poche nozioni fondamentali che possano agevolare la comprensione del capitolo. Un MOS è un dispositivo a quattro terminali (Source, Gate, Drain, Bulk) la cui struttura è schematizzata in figura 3.1.

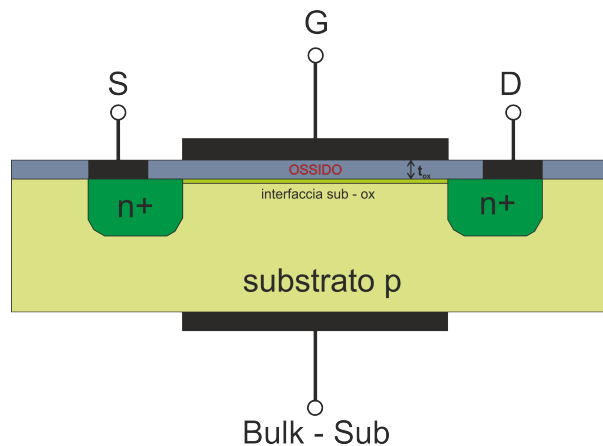


Figura 3.1: Schema elementare della struttura di un dispositivo N-MOSFET.

Una descrizione esaustiva del suo funzionamento si può trovare in [53], in questo contesto consideriamo il caso semplice di un NMOS con bulk cortocircuitato al source. La prerogativa fondamentale del MOS è la possibilità di modulare, attraverso la tensione imposta al morsetto di gate (in prima approssimazione galvanicamente isolato), la corrente che scorre tra Source e Drain. Tale corrente può essere convogliata su una opportuna impedenza di carico per produrre un segnale analogo al modulante ma più ampio; in questo modo il dispositivo funziona da amplificatore di potenza. Perché tutto questo sia possibile devono essere verificate alcune condizioni: il MOS deve essere innanzitutto polarizzato, cioè i suoi morsetti devono essere posti a precisi valori di tensione (non soggetti a grandi variazioni) che permettano di realizzare determinate condizioni fisiche nel materiale semiconduttore di cui è prevalentemente costituito. In particolare, quando la differenza di potenziale di polarizzazione tra Gate e Source (connesso al Bulk) supera una certa *tensione di soglia*  $V_{th}$ , all'interfaccia tra substrato di silicio e ossido di gate si accumula una concentrazione di portatori carichi (elettroni, in un NMOS) che costituisce un vero e proprio canale conduttivo che collega drain e source. In presenza del canale e di una d.d.p. non nulla  $V_{DS}$  tra drain e source, si instaura una corrente tra gli stessi le cui caratteristiche dipendono da numerosi fattori: tensione tra gate e source e la sua differenza rispetto a  $V_{th}$  (detta tensione di overdrive  $V_{OV}$ ), tensione tra drain e source, geometria del dispositivo. La relazione che lega queste grandezze è diversa a seconda che si tratti di un dispositivo a canale lungo ( $L > 1\mu\text{m}$ ) o di uno a canale corto

( $L < 1\mu\text{m}$ ), ma in entrambi i casi piccole variazioni della tensione  $V_{GS}$  (minori di pochi multipli della tensione termica  $U_T$  o di  $2V_{OV}$ ) generano piccole variazioni di corrente nel canale. Ciò significa che da un *piccolo segnale* di tensione al gate dipende una piccola *corrente di segnale*  $i_{DS}$ , attraverso un trasferimento che prende il nome di transconduttanza e si indica con  $g_m$ . In generale la transconduttanza di un MOSFET polarizzato sopra soglia (cioè con  $V_{GS} \geq V_{th}$ ) è il risultato di una approssimazione lineare della sua caratteristica ed è circa pari a [48]:

$$g_m \approx \frac{I_D}{nU_T} \frac{2}{1 + \sqrt{1 + 4 \cdot IC}} \quad (3.1)$$

con  $n \approx 1.4 - 1.7$  [48] e IC prende il nome di *coefficiente di inversione*. Quest'ultimo si ricava dall'equazione:

$$IC = \frac{I_D}{2\mu C'_{ox} \frac{W}{L} U_T^2} \quad (3.2)$$

Una espressione della transconduttanza più compatta e agevole da utilizzare si può ottenere definendo il parametro  $\bar{n}$ :

$$\bar{n} = \frac{1}{n} \frac{2}{1 + \sqrt{1 + 4 \cdot IC}} \quad (3.3)$$

ed è:

$$g_m \approx \frac{I_D}{\bar{n}U_T} \quad (3.4)$$

A seconda del valore di IC si distinguono tre regioni di funzionamento del dispositivo: forte inversione ( $IC > 10$ ), moderata inversione ( $0.1 < IC < 10$ ) e debole inversione ( $IC < 0.1$ ). Per IC molto superiore a 1 la 3.1 tende a coincidere con la classica espressione della transconduttanza del MOSFET in forte inversione:

$$g_m = \frac{2I_D}{V_{OV}} = \sqrt{2\mu C'_{ox} \frac{W}{L} I_D} \quad (3.5)$$

Fissata la corrente nel dispositivo, è possibile portare quest'ultimo da un funzionamento in forte inversione ad un regime di debole inversione aumentando il suo rapporto di forma. A questo cambiamento coincide una diminuzione della tensione  $V_{OV}$  che da valori generalmente superiori a 100mV-200mV si porta fino a poche decine di mV. In queste condizioni il canale non è formato ma è comunque presente una concentrazione

residua di portatori all'interfaccia ossido di gate-substrato compatibile con il passaggio di corrente. L'espressione della corrente di drain in debole inversione e sottosoglia è [48]:

$$I_D = 2n\mu C'_{ox} \frac{W}{L} U_T^2 \exp \frac{V_G - V_{T0}}{nU_T} \left( \exp \frac{-V_S}{U_T} - \exp \frac{-V_D}{U_T} \right) \quad (3.6)$$

Il principale vantaggio dell'utilizzo del MOS sottosoglia è l'aumento, a pari corrente, della transconduttanza e quindi della sua efficienza di utilizzo. Essa è definita come rapporto tra  $g_m$  e la corrente  $I_D$  che scorre nel dispositivo ed illustrata in figura 3.2. In debole inversione e sottosoglia l'efficienza del mosfet aumenta avvicinandosi al valore massimo ideale dell'equivalente transistor bipolare.

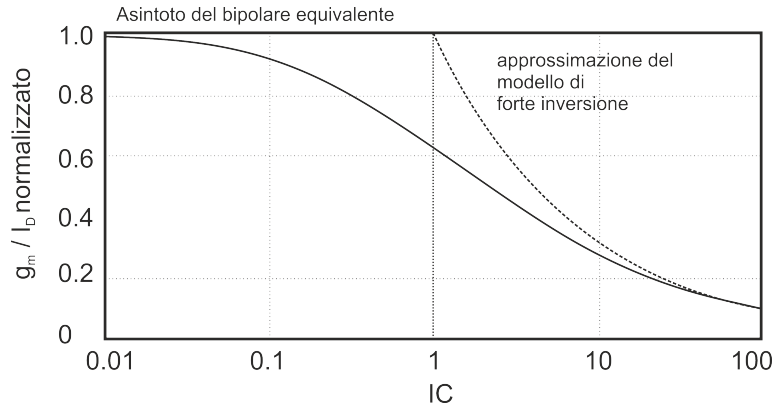


Figura 3.2: Curva di efficienza (normalizzata rispetto al massimo teorico) del dispositivo MOSFET rispetto al coefficiente di inversione.

L'utilizzo di MOS in debole inversione o sottosoglia è molto frequente nelle coppie di ingresso di amplificatori differenziali, questo permette di ridurre il rumore riferito in ingresso, specialmente quello dovuto agli altri transistori del circuito. Gli inconvenienti di pratiche come questa sono prevalentemente riconducibili ad un aumento delle dimensioni del dispositivo, in particolare della  $W$ , che si traduce in una maggiore occupazione di silicio disponibile nel chip ed in una diminuzione della frequenza di taglio. L'aumento di dimensione può risultare critico già in presenza di correnti di pochi  $\mu A$  con MOS in debole inversione ed è anche superiore quando dal regime di debole inversione si passa al funzionamento con una tensione tra gate e source inferiore alla soglia.

### 3.1.2 Pseudoresistori MOS

Un altro impiego dei MOS in sottosoglia è la realizzazione di valori di resistenza così elevati da renderne praticamente impossibile l'implementazione con normali resistori integrati. La denominazione di source e drain in un tipico MOSFET di segnale è puramente formale perchè i due terminali non sarebbero di per sè identificabili a partire

dalla sola struttura fisica, perfettamente simmetrica, del dispositivo. Sia per un PMOS che per un NMOS potremmo dunque chiamare i morsetti di Source e Drain in modo generico, ad esempio A e B, come illustrato in figura 3.3:

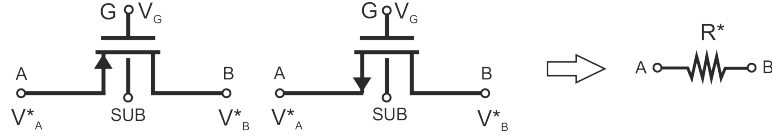


Figura 3.3: NMOS e PMOS configurati a pseudo-resistori [49].

Coerentemente con questa convenzione, ed assumendo il terminale di bulk connesso al substrato, la corrente tra A e B in un MOS in debole inversione (o sottosoglia) assume l'espressione:

$$I_{AB} = \pm 2n\mu C'_{ox} \frac{W}{L} U_T^2 \exp \frac{V_G - V_{T0}}{nU_T} \left( \exp \frac{-V_A}{U_T} - \exp \frac{-V_B}{U_T} \right) \quad (3.7)$$

che è una relazione lineare tra la corrente e la funzione esponenziale delle tensioni ai terminali. Si definiscano ora le grandezze chiamate *pseudo-tensioni* associate ai terminali A e B come segue:

$$V_{A,B}^* = \pm V_0 \exp \frac{-V_{A,B}}{U_T} \quad (3.8)$$

e la *pseudo-conduttanza*:

$$G^* = \frac{1}{R^*} = \frac{2n\mu C'_{ox} \frac{W}{L} U_T^2}{V_0} \exp \frac{V_G - V_{T0}}{nU_T} \quad (3.9)$$

dove  $V_0$  è una generica tensione positiva. Allora si può riscrivere la 3.7 nel seguente modo:

$$I_{AB} = G^* (V_A^* - V_B^*) = \frac{(V_A^* - V_B^*)}{R^*} \quad (3.10)$$

cioè una relazione lineare simile alla legge di ohm. Per analogia, ogni rete ottenuta interconnettendo transistori modellizzati come bipoli source-drain (A,B) è lineare per correnti e pseudo-tensioni (non tensioni) e risponde pertanto alle legge di Kirchoff ai nodi. Ogni rete di normali resistori lineari può essere ricostruita sostituendo a questi dei MOS in debole inversione. Inoltre, il valore delle pseudo-resistenze è controllabile attraverso la tensione di gate, come si deduce dalla 3.9. Se differenti transistori in debole

inversione e in un substrato comune hanno la stessa tensione di gate, allora i valori dei corrispondenti pseudoresistori sono proporzionali alle correnti specifiche  $I_{spec}$ . Se invece esiste una differenza  $\Delta V_G$  tra i dispositivi allora il rapporto tra le loro pseudoresistenze cambia esponenzialmente, in accordo alla 3.9.

Questo tipo di analisi può risultare utile per comprendere il principio di funzionamento di reti a pseudoresistori ed è sufficiente ad analizzare reti in cui le tensioni dei nodi di drain e source sono circa costanti e con una differenza di potenziale molto piccola tra loro. Come vedremo in seguito, nel progetto di amplificatori una analisi di questo tipo non è utile a comprendere le caratteristiche dei circuiti perchè spesso uno dei due terminali A o B (Drain o Source) è soggetto a significative variazioni di tensione. Occorre formalizzare le relazioni in modo più adatto all'applicazione, a questo scopo si riscrive la 3.7 come segue, considerando uno pseudoresistore PMOS:

$$I_{SD} \approx I_{spec} \exp \frac{V_{GS} - V_{T0}}{nU_T} (1 - \exp \frac{-V_{SD}}{U_T}) \quad (3.11)$$

dove  $I_{spec}$  è pari a  $2n\mu C'_{ox} \frac{W}{L} U_T^2$  e in cui è stato raccolto il termine esponenziale della maggiore tensioni tra  $V_S$  e  $V_D$ , che assumiamo convenzionalmente come  $V_S$ . Si noti che per  $V_D$  pari a  $V_S$  non scorre corrente nel dispositivo, ma anche solo per piccole d.d.p. tra i due nodi la corrente non è nulla e si può associare al dispositivo una pseudoresistenza che può essere ricavata come segue:

$$R^* = \left( \frac{dI_{SG}}{dV_{DS}} \right)^{-1} = \frac{1}{I_{spec} \exp(\frac{V_{SD}}{U_T}) \exp(\frac{V_{GS} - V_{T0}}{nU_T})} \quad (3.12)$$

espressione equivalente al reciproco della 3.9. Si osserva dalla 3.12 che il valore della resistenza dipende da diversi fattori e che fissata la tensione di gate e il rapporto di forma del transistor (cioè fissata  $I_{spec}$ ) allora la resistenza è comunque soggetta alle variazioni associate alla tensione  $V_{DS}$ . Se la tensione di drain supera la tensione di source la relazione 3.12 diventa speculare, con l'interscambio delle tensioni di gate e source.

Di questa proprietà si deve tenere conto ogni volta che si desidera sintetizzare una resistenza se essa serve ad imporre caratteristiche di un circuito, come il guadagno o una singolarità, che dovrebbero essere mantenute il più possibile costanti. In figura 3.4 è mostrato l'andamento della resistenza sintetizzata da uno pseudoresistore PMOS per una data tensione di gate ed al variare della tensione di source con una analisi DC in Cadence *Virtuoso*<sup>TM</sup>. Anche con simulazioni AC è possibile stimare il valore della resistenza sintetizzata da uno pseudoresistore, la figura 3.5 mostra il risultato di una di queste simulazioni. Anche trascurando il ruolo della tensione tra drain e source, il valore assoluto della loro resistenza mostra comunque una dipendenza esponenziale dalla tensione di gate (precisamente da  $V_{GS}$ ). Per questo motivo e nell'ottica di dover realizzare resistori variabili, è sconsigliabile ricercare una regolazione accurata della resistenza attraverso l'imposizione diretta della tensione di Gate. Un approccio migliore consiste

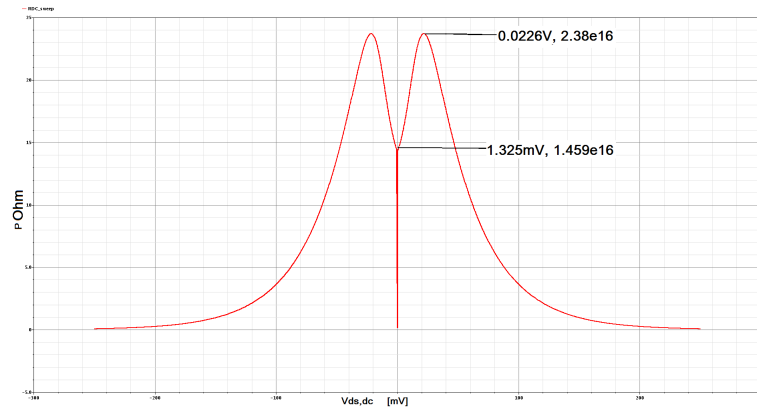


Figura 3.4: Resistenza di uno pseudoresistore MOS HG UMC130nm, analisi DC al variare della tensione Drain Source.

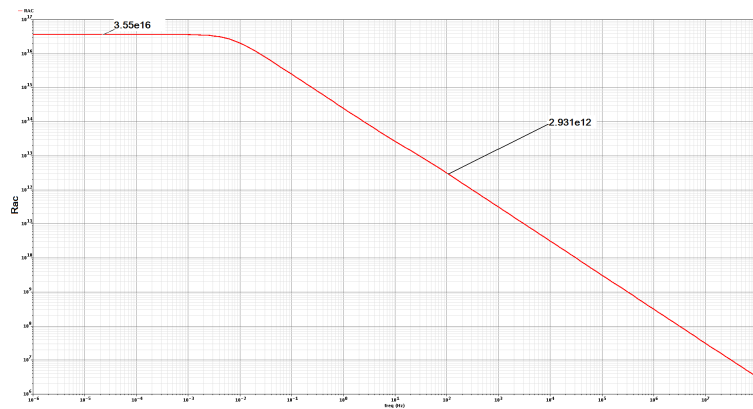


Figura 3.5: Resistenza di uno pseudoresistore MOS HG UMC130nm, analisi AC.

nella regolazione in corrente attraverso un transistor di controllo delle stesse dimensioni dello pseudoresistore, secondo lo schema di figura 3.6:

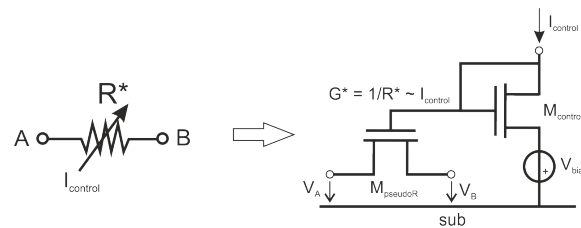


Figura 3.6: Controllo in corrente della pseudoresistenza.



In questo modo la tensione di gate è il risultato di una compressione logaritmica della corrente  $I_{control}$  e la dipendenza della pseudo-resistenza dalla grandezza di controllo è lineare.

In ogni caso, quando non sussiste il bisogno di una selezione accurata del valore di  $R^*$  e si desidera semplicemente regolarne l'ordine di grandezza, anche il controllo diretto della tensione di gate può rappresentare una via percorribile. La necessità di realizzare poli a frequenze vicine o inferiori all'Hertz rende questi dispositivi particolarmente interessanti per i sistemi integrati di neural monitoring, come testimoniano le numerose soluzioni proposte in letteratura. Già nel primo sistema di neural monitoring, cioè quello di Harrison [13], si trova la loro prima proposta di utilizzo nel blocco di amplificazione.

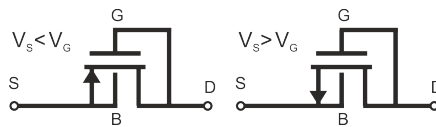


Figura 3.7: NMOS e PMOS configurati a pseudo-resistori come nel sistema di Harrison (PMOS).

La configurazione proposta per gli pseudo-resistori di questo sistema (Fig. 3.7) è leggermente diversa da quella spiegata finora perchè presenta drain e gate cortocircuitati tra loro, così come source e bulk. In realtà questa non rappresenta una differenza sostanziale: i transistori sono comunque polarizzati in debole inversione e i nodi sono posti alla stessa tensione di polarizzazione. In queste condizioni le proprietà resistive dei bipoli precedentemente descritte sono conservate. L'unica limitazione rispetto al caso precedente è dovuta al cortocircuito tra gate e drain che rende impossibile la regolazione in tensione del valore di resistenza, dipendente solo dalle proprietà fisiche e geometriche del dispositivo. Un discorso analogo vale anche per altre configurazioni proposte ma che per questioni di sintesi evitiamo di riportare; le si potranno in ogni caso esaminare nei lavori cui si farà riferimento nel corso del testo.

## 3.2 Caratteristiche di Sistema

In questa sezione è illustrata l'architettura generale dei blocchi costituenti lo stadio di amplificazione e filtraggio accompagnata dalla spiegazione di alcune scelte circa la modalità di implementazione delle singolarità del trasferimento e le relazioni che, scelta l'architettura, permettono di valutare i parametri di interesse (rumore, guadagno, banda) del circuito.

### 3.2.1 Architettura Generale

L'architettura complessiva dello stadio di amplificazione e filtraggio è illustrata in figura 3.8. Sia LNA che PGA sono realizzati utilizzando degli OTA completamente differenziali simmetricamente reazionati grazie a reti capacitive e pseudoresistori PMOS. L'accoppiamento in AC con la sorgente è garantito dalle due capacità d'ingresso  $C_{1_{LNA}}$  del LNA di 15pF ciascuna. Queste, oltre a offrire in banda utile una impedenza molto superiore (oltre  $60M\Omega$  ad 1kHz) a quella dell'elettrodo, contribuiscono al guadagno dell'amplificatore. Per effetto della retroazione infatti esso è dato dal rapporto tra l'impedenza loro associata e l'impedenza associata al ramo di reazione. Sia nel PGA che nel LNA questa è costituita dal parallelo di uno pseudoresistore PMOS regolabile in tensione e una capacità  $C_2$  di 150fF.

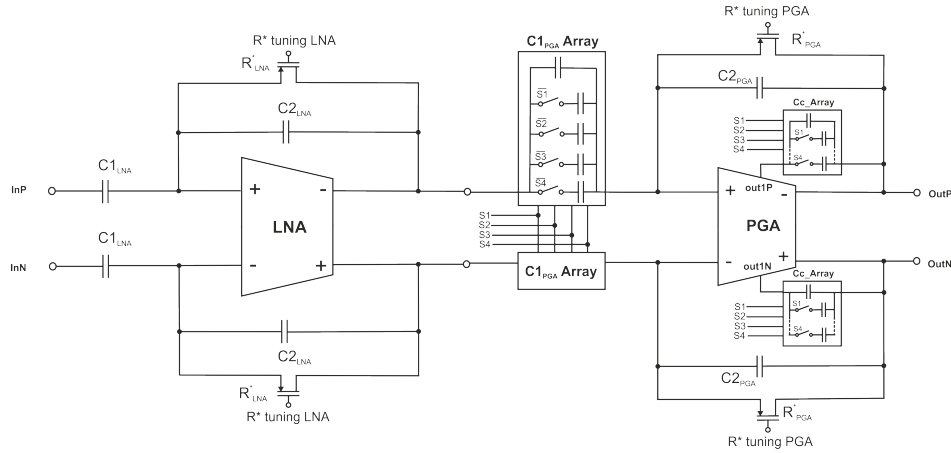
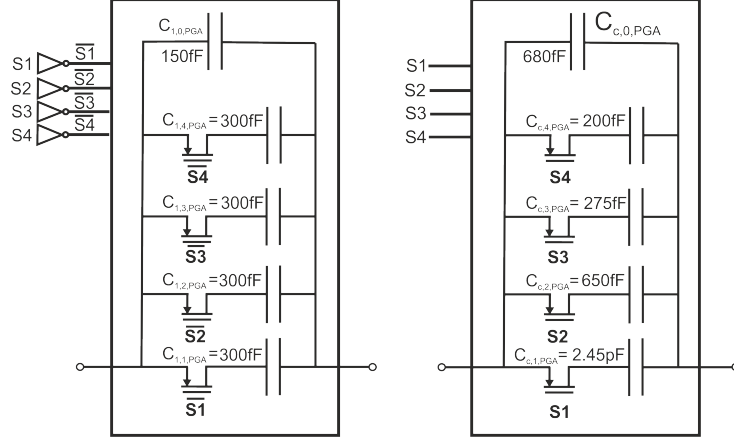


Figura 3.8: Catena di amplificazione del front-end.

Essendo gli pseudoresistori MOS dei due rami dimensionati per ottenere resistenze  $R^*$  superiori al  $T\Omega$ , in banda utile l'impedenza associata alla capacità dei reazioni  $C_{2_{LNA}}$  (uguale a  $C_{2_{PGA}}$ ) è molto inferiore e pertanto si ottiene un guadagno complessivo costante pari a:

$$G_{LNA} = \frac{C_{1_{LNA}}}{C_{2_{LNA}}} = 100 \quad (3.13)$$

corrispondente a 40dB. La variabilità del guadagno del PGA è stata ottenuta grazie a due coppie simmetriche di array capacitivi dotati di interruttori attivabili con quattro segnali logici di selezione (0V-0.5V): S1, S2, S3 e S4.



(a) Capacità di ingresso  $C_{1,PGA}$  del PGA. (b) Capacità di compensazione dell'OTA del PGA.

Figura 3.9: I due Arrays di capacità del PGA.

Una coppia di array del tipo mostrato in Fig. 3.9(a) è controllata dai negati dei bit di selezione e serve alla regolazione del guadagno attraverso la capacità di ingresso del PGA (selezionabile tra 150fF, 450fF, 750fF, 1.05pF, 1.35pF). Con questa soluzione e data l'analogia con l'architettura generale del LNA il guadagno del PGA può essere ricavato dalla relazione:

$$G_{PGA} = \frac{C_{1,0,PGA} + \sum_{i=1}^4 S_{\langle i \rangle} \cdot C_{1,i,PGA}}{C_{2,PGA}} \quad (3.14)$$

La seconda coppia di array capacitivi, del tipo in Fig. 3.9(b), è controllata dagli stessi segnali logici  $S_{\langle 1:4 \rangle}$  e serve per regolare, coerentemente al guadagno selezionato, la capacità di compensazione dell'OTA. In questo modo è possibile garantire la stessa banda e la stessa stabilità qualsiasi sia il guadagno scelto. I valori di guadagno selezionabili corrispondono ai 5 possibili rapporti tra  $C_{1,PGA}$  e  $C_{2,PGA}$  e sono riassunti dalla tabella 3.1. Maggiori dettagli sul loro dimensionamento si trovano nella sezione dedicata al PGA.

Il trasferimento del LNA è di tipo passa banda e il valore del polo passa alto è pari a circa:

$$f_{HP} \approx \frac{1}{2\pi C_{2,LNA} R^*} \quad (3.15)$$

Tabella 3.1: Relazioni tra bit di controllo, capacità d'ingresso, capacità di compensazione e guadagno del PGA.

S1	S2	S3	S4	$C_{1,PGA}$	$G_{PGA}$	$C_{C,PGA}$
0	0	0	0	150fF	0dB	4.25pF
1	0	0	0	450fF	9dB	1.805pF
1	1	0	0	750fF	13dB	1.155pF
1	1	1	0	1.05pF	16dB	880fF
1	1	1	1	1.35pF	19dB	680fF

Gli pseudoresistori MOS di reazione di entrambi gli amplificatori sono stati dimensionati con rapporti di forma molto ridotti e tali da prevenire lo spostamento dei poli passa alto al di sopra del limite inferiore di banda anche in presenza di ampi segnali all'uscita. Il polo passa basso sarà circa pari al  $GBWP_{LNA}$  (prodotto guadagno banda del LNA, tenendo conto di OTA e reazione) e dovrebbe essere compreso tra 10kHz e 15kHz. Un discorso analogo si può fare per il PGA ed è opportuno sottolineare la possibilità di regolare in modo specifico attraverso la tensione di gate l'ordine di grandezza della resistenza degli pseudoresistori MOS sia del LNA che del PGA. In questo modo è possibile separare la frequenza di taglio passa alto del LNA dalla omologa del PGA traendone un significativo vantaggio in termini di riduzione del rumore. La giustificazione di quanto appena affermato è illustrata nel prossimo paragrafo.

### 3.2.2 Analisi del Rumore

Le prestazioni di rumore dell'intero front-end dipendono prevalentemente dal progetto del LNA e dal contributo degli pseudoresistori MOS di reazione. Si consideri inizialmente il rumore  $\bar{E}_{n,OTALNA}^2$  introdotto dall'OTA come se fosse l'unico contributo al rumore, tralasciando quello dovuto agli pseudoresistori. Esso può essere espresso in modo compatto come somma del termine termico  $\bar{E}_{n,OTA_{th,LNA}}^2$  e di quello relativo al rumore flicker  $\bar{E}_{n,OTA_{f,LNA}}^2$  (le espressioni dettagliate degli addendi saranno studiate nella successiva sezione). Indicando con  $C_p$  le capacità parassite di ciascun nodo di ingresso dell'OTA del LNA e considerando il circuito nella banda di interesse (fig. 3.10), si può verificare che se da un lato queste non influenzano in modo significativo il guadagno dell'amplificatore, dall'altro hanno un effetto peggiorativo sul contributo di rumore riferito in ingresso dell'OTA.

Per effetto della retroazione esso è infatti pari a:

$$\bar{E}_{n,INLNA}^2 = \bar{E}_{n,OTALNA}^2 \left(1 + \frac{Z_2}{Z_1 || Z_P}\right)^2 \frac{1}{G_{LNA}^2} \quad (3.16)$$

E tenendo conto che all'interno della banda utile il comportamento di tutte le impedenze della rete di reazione esterna è capacitivo, si ottiene:

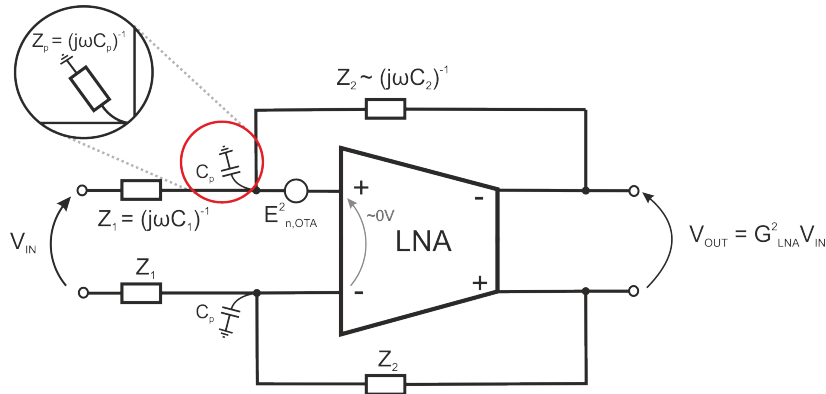


Figura 3.10: Schematizzazione dell'amplificatore e ruolo della capacità parassita.

$$\bar{E}_{INLNA}^2 = (\bar{E}_{n,OTA_{th,LNA}}^2 + \bar{E}_{n,OTA_{\frac{1}{f},LNA}}^2) \cdot \left( \frac{C_p + C_{1LNA} + C_{2LNA}}{C_{1LNA}} \right)^2 \quad (3.17)$$

$$\bar{E}_{INLNA}^2 = (\bar{E}_{n,OTA_{th,LNA}}^2 + \bar{E}_{n,OTA_{\frac{1}{f},LNA}}^2) \cdot \left( 1 + \frac{C_p}{C_{1LNA}} + \frac{1}{G_{LNA}} \right)^2 \quad (3.18)$$

L'utilizzo di una capacità  $C_{1LNA}$  grande permette di massimizzare il guadagno e ridurre l'impatto del parassita  $C_p$ . Si può comprendere come questo parassita possa essere significativo se si considera la necessità di avere dei transistori di ingresso con grande superficie sia per mantenerli sottosoglia che per abbassare il più possibile la *corner frequency* dello spettro di rumore riferito in ingresso e portare fuori dalla banda i contributi di rumore a frequenza più bassa. In ogni caso almeno in fase di layout è stata ricercata una soluzione che minimizzasse il contributo al parassita delle connessioni e le capacità di ingresso sono state connesse al nodo di gate per il loro top plate. La relazione 3.18 potrebbe suggerire di aumentare il guadagno associato al primo blocco di amplificazione, esso tuttavia è limitato dall'impedenza offerta dall'elettrodo, dominata da un contributo capacitivo di 150-350pF ed inoltre con segnali ampi 1-2mV in ingresso un guadagno superiore a 100 potrebbe facilmente saturarne la dinamica di uscita. I 15pF scelti per  $C_{1LNA}$  dovrebbero essere comunque sufficienti a limitare il deterioramento del rumore dovuto ai parassiti pur garantendo una impedenza d'ingresso adeguata; in ogni caso è stato considerato come limite per  $C_p$  il valore di 500fF per mantenere intorno all'1% il deterioramento delle prestazioni di rumore. La descrizione dettagliata dei contributi termico e flicker dovuti all'OTA dell'LNA dipende dalla topologia circuitale selezionata e pertanto sarà presentata nel relativo paragrafo.

A questo punto si può considerare anche il rumore associato agli pseudoresistori MOS. Sperimentalmente è stato verificato [5] che la loro densità spettrale di potenza è quella tipica del rumore termico ed è legata alla resistenza di piccolo segnale  $R^*$  da essi sintetizzata:

$$\bar{E}_{n,R^*}^2 = 4kTR^* \quad (3.19)$$

La possibilità di separare le frequenze  $f_{HP_{LNA}}$  e  $f_{HP_{PGA}}$  permette una significativa diminuzione di questo contributo. La densità spettrale di potenza del rumore degli pseudoresistori in uscita deriva dall'interazione con le capacità  $C_{2_{LNA}}$  per effetto della retroazione:

$$\bar{E}_{n,R^*,IN}^2 = 2 \frac{4kT}{R^*} \frac{1}{(2\pi C_{2_{LNA}} f)^2} \quad (3.20)$$

e riportata in ingresso (si divida per il quadrato del guadagno) risulta pari a:

$$\bar{E}_{n,R^*,IN}^2 = 2 \frac{4kT}{R^*} \frac{1}{(2\pi C_{1_{LNA}} f)^2} = \frac{8kT f_{HP_{LNA}}}{2\pi G_{LNA}^2 C_{2_{LNA}} f^2} \quad (3.21)$$

mostrando un andamento inversamente proporzionale al quadrato della frequenza. Tale andamento suggerisce la possibilità di filtrare il loro contributo di rumore attraverso l'opportuna scelta delle frequenze di taglio passa alto. Integrando tale espressione nella banda passante del trasferimento complessivo della catena di amplificazione, cioè calcolando:

$$\int_{f_{HP_{PGA}}}^{f_{LP}} \bar{E}_{n,R^*,IN}^2 df = \frac{8kT f_{HP_{LNA}}}{2\pi G_{LNA}^2 C_{2_{LNA}}} \int_{f_{HP_{PGA}}}^{f_{LP}} \frac{1}{f^2} \quad (3.22)$$

si ottiene:

$$\frac{8kT f_{HP_{LNA}}}{2\pi G_{LNA}^2 C_{2_{LNA}} f_{HP_{PGA}}} \quad (3.23)$$

Da cui risulta evidente il vantaggio di massimizzare il rapporto tra la frequenza di taglio passa alto del PGA (vero limite inferiore di banda del trasferimento complessivo) e quella del LNA. I valori di resistenza degli pseudoresistori misurati sperimentalmente si attestano tipicamente un ordine di grandezza al di sotto di quelli simulati [38]. In questo progetto si è scelto di dimensionare gli pseudoresistori PMOS di entrambi gli amplificatori per ottenere (in simulazione) valori resistivi oltre un ordine di grandezza superiori ai valori necessari, lasciando però la possibilità di regolarne il valore (almeno per il chip prototipo) attraverso l'imposizione esterna della tensione di gate (mediante un opportuno pin del chip). Per rispettare i vincoli di banda e per non abbassare troppo il limite inferiore della banda del LNA (ed evitare di raccogliere troppo rumore flicker dell'OTA), si è previsto di regolare le resistenze  $R^*$  degli pseudoresistori per ottenere

$f_{HP_{LNA}}$  pari a 10mHz e  $f_{HP_{PGA}}$  pari a 100mHz. Data la capacità di reazione di 150fF le resistenze sintetizzate dovranno essere pari a 10T $\Omega$  per l'LNA e 1T $\Omega$  per il PGA. Il contributo degli pseudoresistori al rumore complessivo riferito in ingresso per la catena di amplificazione dimensionata secondo i parametri finora descritti dovrebbe teoricamente essere pari a circa 600nV<sub>rms</sub>, valore quasi dieci volte inferiore al limite previsto dalle specifiche.

Dall'analisi appena svolta risulta evidente che ad assumere il ruolo più rilevante nel determinare il rumore riferito in ingresso sia l'OTA costituente il LNA. Nel precedente sistema la topologia circuitale utilizzata per questo amplificatore era un telescopic cascode con ingresso a PMOS ed una NEF misurata dell'intero blocco di filtraggio e amplificazione di circa 2.41. La massima NEF teorica per uno stadio differenziale a bipolari è pari a  $\sqrt{2}$ . Uno stadio differenziale CMOS, la cui efficienza  $g_m/I_D$  è sempre inferiore a quella del bipolare, non potrà che avere una NEF superiore a tale limite ed in particolare il più basso valore possibile per la NEF di uno stadio differenziale CMOS è di circa 2.02 [5]. Nella tabella 3.11 sono riportate le NEF di alcuni recenti lavori pubblicati nell'ambito di amplificatori per segnali biologici.

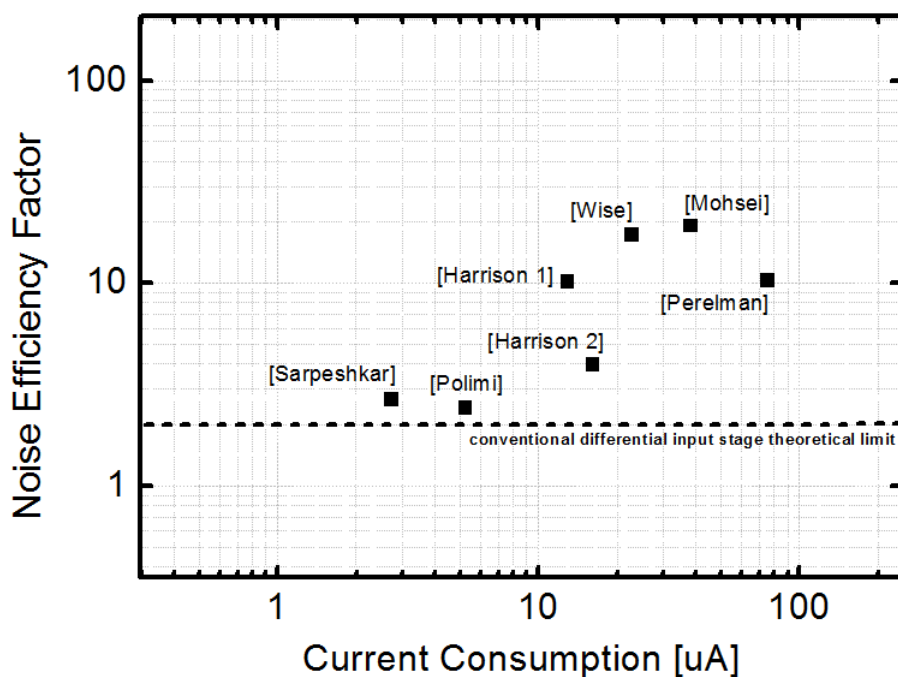


Figura 3.11: NEF di lavori recentemente pubblicati( Polimi [39], Sarpeshkar [17], Harrison 2003 (1) [14] e 2007 (2) [13], Wise [15], Perelman [16] .

Per ridurre la NEF al di sotto del valore 2 è necessario affidarsi a topologie circuitali differenti dai classici stadi d'ingresso a singola coppia differenziale e carico attivo. La

topologia utilizzata e la soluzione delle problematiche associate al suo impiego saranno adeguatamente spiegate nella prossima sezione. Solo per dare alcuni valori di riferimento, in base alla 3.24, per ottenere un valore di NEF inferiore a due ammettendo di ottenere un rumore di ingresso pari a  $5\mu V_{rms}$  e con una banda di 10kHz, la massima corrente assorbita dall'alimentazione per l'intera catena dovrebbe essere ricavabile dalla relazione:

$$I_{TOT} \leq \frac{NEF^2}{(5\mu V_{rms})^2} \cdot \pi U_T 4kTBW \approx 2.14\mu A \quad (3.24)$$

Questo valore è ben inferiore ai  $6\mu A$  necessari all'adempimento delle specifiche. Va infine ricordato che la NEF non tiene conto del risparmio energetico derivante dalla riduzione dell'alimentazione perchè considera unicamente la corrente che da quest'ultima è assorbita. Con le specifiche tipiche dell'applicazione e considerando topologie circuitali classiche il consumo di corrente è destinato a rimanere sempre nell'ordine di  $3-4\mu A$  per canale [5]. Sebbene meno utilizzata, la PEF è un indice più adeguato a rappresentare l'efficienza energetica e per questo motivo verrà sempre presa come riferimento, insieme alla NEF, nelle analisi conclusive.

### 3.3 Low Noise Amplifier

La sezione è dedicata all'analisi dettagliata del low noise amplifier, comprendente OTA, reazione esterna e rete di reazione di modo comune. Al controllo del di modo comune è dedicato un paragrafo in cui sono introdotti alcuni elementi teorici ed altri più relativi all'applicazione necessari a chiarire alcune scelte circa le topologie circuitali adottate. Alla fine della sezione sono riportate le prestazioni dello stadio ottenute come risultato di simulazioni montecarlo.

#### 3.3.1 Progetto e Dimensionamento del Primo Stadio

Migliorare la NEF per portarla al di sotto di 2 impone l'abbandono delle architetture differenziali comunemente utilizzate per gli stadi d'ingresso di OTA CMOS a favore di nuove topologie. La migliore dal punto di vista dell'efficienza, comunemente definita *tail-current-reuse topology* (TCR), è stata presentata per la prima volta da Holleman et al. [47] ed è illustrata schematicamente in figura 3.12.

Se confrontato con un normale stadio differenziale, questo stadio di ingresso utilizza un solo transistor in più offrendo una transconduttanza equivalente doppia. Chiamando genericamente  $g_{m_n}$  e  $g_{m_p}$  le transconduttanze degli NMOS e dei PMOS rispettivamente, per facilitare la stima degli ordini di grandezza assumeremo come ulteriore semplificazione:

$$g_{m_n} = g_{m_p} = g_m \quad (3.25)$$



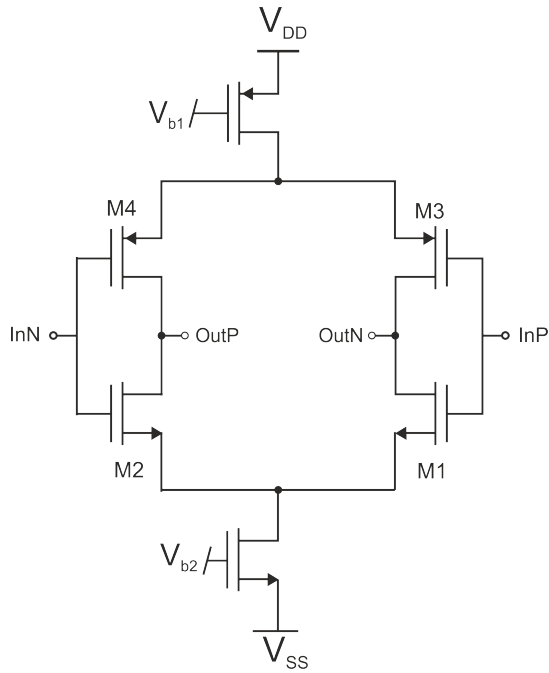


Figura 3.12: Stadio di ingresso Tail-Current-Reuse.

e indicheremo con  $r_0$  la resistenza di Early dei MOS. Si ha così:

$$g_{m_{TOT}} = \frac{i_{out_{DM}}}{v_{in,diff}} = \frac{g_{m,n} + g_{m,p}}{2} = g_m \quad (3.26)$$

e anche:

$$R_{out_{DM}} \approx 2r_{0n} || 2r_{0p} \approx r_0 \quad (3.27)$$

Per un guadagno differenziale dello stadio pari circa a:

$$G_{Diff} \approx g_m r_0 \quad (3.28)$$

Per valutare il rumore riferito all'ingresso si sfrutta invece il teorema di norton. In questo modo è possibile ricavare il contributo alla corrente di rumore in uscita di ciascun MOS e dividendo quest'ultimo per il guadagno al quadrato dello stadio si ottiene anche il relativo rumore riferito in ingresso. Quindi:

$$\bar{E}_{noiseOUT}^2 = \bar{I}_{nOUT}^2 R_{out}^2 \quad (3.29)$$

$$\bar{E}_{noiseIN}^2 = \frac{\bar{E}_{nOUT}^2}{G_{Diff}^2} \quad (3.30)$$

Esplicitando i termini supponendo sempre l'uguaglianza delle transconduttanze e considerando il caso generale di MOS in forte inversione, si può dimostrare che il contributo alla densità spettrale di corrente di rumore in uscita dei singoli transistori di ingresso è pari a:

$$\bar{I}_{n_{th},OUT}^2|_{single} = 2kT\gamma g_m \quad (3.31)$$

e conseguentemente si ha:

$$\bar{E}_{n,OTA_{th},OUT}^2 = 4 \cdot (\bar{I}_{n_{th},OUT}^2)_{single} R_{out}^2 = 8kT\gamma g_m R_{out}^2 \quad (3.32)$$

$$\bar{E}_{n,OTA_{th},LNA}^2 = \frac{8kT\gamma}{g_m} \quad (3.33)$$

Queste relazioni sono state scritte immaginando che i transistori operino in forte inversione. Se invece i transistori operano sottosoglia, allora le densità spettrali di rumore di corrente sono di tipo shot, ovvero per un singolo MOS in saturazione vale la seguente espressione per la densità spettrale di rumore in uscita (in corrente):

$$\bar{I}_{n_{th},MOS_{single}}^2 \approx 2qI_D \left(1 + \exp \frac{-V_{DS}}{U_T}\right) \approx 2qI_D \quad (3.34)$$

mentre per l'intero stadio, applicando Norton e tenendo conto della 3.4, si ha:

$$\bar{E}_{n,OTA_{th},OUT}^2 = 4 \cdot \frac{1}{2} 2qI_D R_{out}^2 = 4qI_D R_{out}^2 \quad (3.35)$$

$$\bar{E}_{n,OTA_{th},LNA}^2 = \frac{4qI_D}{g_m} \approx \frac{4q\bar{n}^2 U_T^2}{I_D} \quad (3.36)$$

Questo risultato coincide con il rumore termico riferito all'ingresso di uno stadio differenziale classico in cui si considerano come sorgenti di rumore soltanto i MOS della coppia differenziale di ingresso; non ci sono altri contributi. Lo stadio riduce quindi

i contributi di rumore ai minimi termini con un ovvio impatto positivo sulle cifre di merito. Per il rumore flicker è possibile effettuare una analisi analoga e ottenere:

$$\bar{E}_{n,OTA\frac{1}{f},LNA}^2 = \frac{k_n^{\frac{1}{f}}}{C'_{ox}W_nL_n f} + \frac{k_p^{\frac{1}{f}}}{C'_{ox}W_pL_p f} \quad (3.37)$$

La corrente del primo stadio può essere dimensionata trascurando quest'ultimo contributo e considerando solo quello termico. Data la banda, per avere un rumore riferito in ingresso minore di  $5\mu V_{rms}$ , deve essere rispettata la relazione:

$$\sqrt{\bar{E}_{n,OTA_{th},LNA}^2 \cdot BW} \leq 5\mu V_{rms} \quad (3.38)$$

da cui, considerando una banda di 15kHz e un fattore  $\bar{n}$  di circa 1.6, si può ottenere la corrente minima per il singolo ramo:

$$I_D \geq \frac{4q\bar{n}^2 U_T^2 BW}{(5\mu V_{rms})^2} \approx 409nA \quad (3.39)$$

Per avere un migliore margine è stata scelta una corrente complessiva per il primo stadio  $I_1$  di 900nA (450nA per ramo) cui corrisponde un rumore riferito all'ingresso di circa  $4.26\mu V_{rms}$ .

Resta tuttavia da considerare il comportamento del circuito sul modo comune. A tal scopo si consideri lo stadio di figura 3.13, dove semplicemente sono evidenziate le resistenze di early dei generatori MOS di coda, ciascuna assunta pari a  $r_{0T}$ .

In presenza di un segnale di modo comune all'ingresso è possibile separare i due rami in parallelo e, sempre assumendo l'uguaglianza delle transconduttanze, è facile ricavare:

$$R_{OUT_{CM}} \approx g_m r_{01} r_{0T} \quad (3.40)$$

Questa resistenza è molto più elevata della corrispettiva differenziale e contribuisce ad un aumento del guadagno di modo comune, che è pari a :

$$G_{1st,CM} = \frac{g_m R_{OUT_{CM}}}{1 + 2g_m r_{0T}} \approx g_m r_0 \quad (3.41)$$

Un guadagno di modo comune così elevato implica il rischio di trovarsi con un pessimo CMRR in presenza di mismatch e la possibilità di sacrificare troppa dinamica di amplificazione anche per piccoli disturbi in ingresso. Inoltre, come si sarebbe già potuto osservare dalla figura 3.12, la serie di due generatori di corrente dello stesso valore non

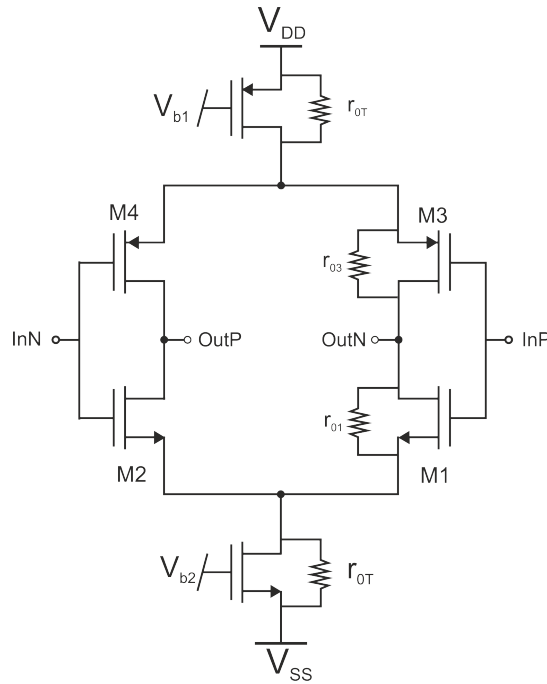


Figura 3.13: Stadio di ingresso Tail-Current-Reuse con generatori di coda reali.

è praticamente mai realizzabile, ed uno stadio così costruito non sarebbe polarizzabile senza adottare una opportuna rete di modo comune che ne stabilizzi il punto di lavoro. Per questi motivi fondamentali lo stadio effettivamente utilizzato per il front-end presenta alcune differenze che lo rendono adeguato al suo impiego. Esso è illustrato in figura 3.14.

Una struttura simile è stata presentata in [50]. La particolare interconnessione dei transistori permette di fissare la tensione di uscita attraverso la tensione  $V_{GS}$  dei due NMOS di coda. A differenza di alcune sue varianti [55], questa struttura mantiene il suo funzionamento su variazioni di processo e senza significative alterazioni della corrente di polarizzazione. La fig. 3.15 riporta la struttura che polarizza ciascuno dei due rami. Essa costituisce un sistema retroazionato che tende ad annullare qualsiasi variazione di tensione ai nodi d'anello, compreso quello d'uscita, dovute a relative iniezioni di corrente di segnale di modo comune.

Al contrario, su segnale differenziale il nodo Y, cui afferiscono entrambi i source degli NMOS di ingresso dei due rami, si comporta come un nodo fermo in tensione ed esclude l'azione della rete di retroazione. Il Guadagno di anello della rete locale coincide con il guadagno di anello del circuito di fig. 3.15 ed è pari a:

$$G_{loop_{CM}} \approx -g_{m2} r_{o2} g_{m5} r_{o5} \approx g_m^2 r_0^2 \quad (3.42)$$

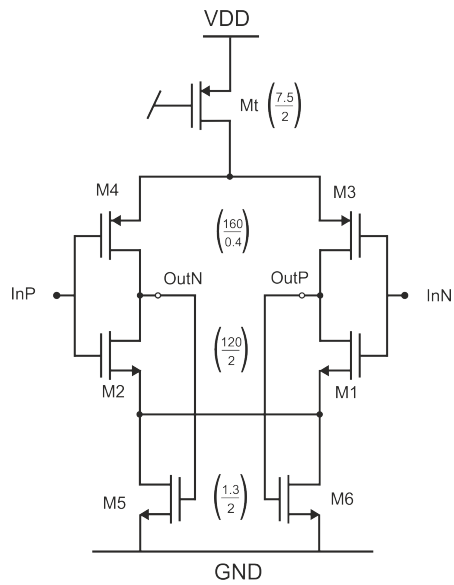


Figura 3.14: Stadio di ingresso dell'LNA.

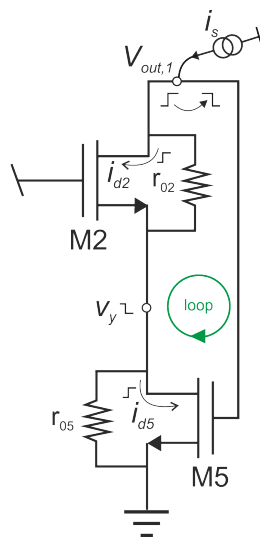


Figura 3.15: Nucleo del circuito di reazione intrinseco del primo stadio dell'LNA.

Studiando il modello di mezzo circuito su segnale di modo comune si ottiene per il primo stadio un guadagno definito dalla relazione:

$$G_{CM-CM_I} = \frac{G_{CM-CM_I}^0}{1 - G_{loop_{CM}}} \approx \frac{2g_{m_1}R_{out1_{CM}}}{1 + g_{m_1}r_{01}} \cdot \frac{1}{g_{m_1}r_{01}g_{m_5}r_{05}} \quad (3.43)$$

In cui  $G_{CM-CM_I}^0$  indica il trasferimento che si avrebbe in assenza di retroazione. Con tutti i MOS in sottosoglia le transconduttanze si possono considerare praticamente identiche tra loro e sostituendo il valore della resistenza di uscita di modo comune con l'espressione 3.40, si ha:

$$G_{CM-CM_I} \approx \frac{2}{g_m r_0} \quad (3.44)$$

Ottenendo così la desiderata attenuazione del guadagno di modo comune. Questa struttura non presenta inoltre problemi di stabilità perchè l'effetto Miller dovuto alla capacità di compensazione tra primo e secondo stadio introduce un polo dominante nell'anello locale che abbassa il prodotto guadagno banda ben al di sotto della seconda singolarità, dovuta a piccoli parassiti al nodo di source dei MOS di ingresso. Una alternativa alla rete di modo di comune scelta è quella illustrata in Figura 3.16 che prevede una rete esterna.

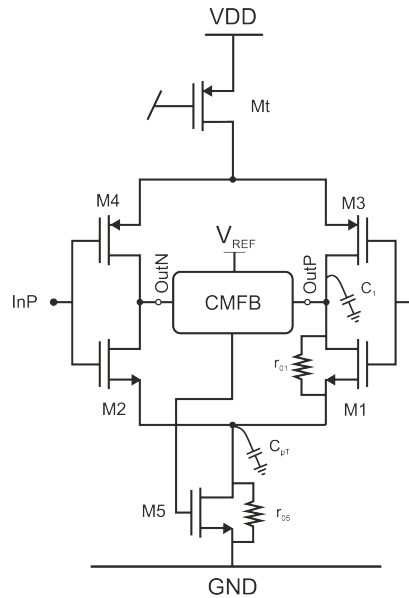


Figura 3.16: Stadio di ingresso dell'LNA con rete di CMF esterna.

Questa soluzione permetterebbe di controllare attraverso una tensione di riferimento il modo comune in uscita ma al prezzo di una serie di problematiche relative alla

compensazione, resa difficile dalla elevata impedenza di modo comune dei nodi di uscita dello stadio. Inoltre è da sottolineare che la polarizzazione di una rete esterna implica un consumo aggiuntivo di corrente che non è desiderabile in applicazioni ultra-low-power. Per questo motivo è stata scelta l'architettura di figura 3.14, vantaggiosa sia in termini di consumo che di semplicità.

L'adozione di un approccio specifico per i singoli stadi nella realizzazione delle reti di controllo del modo comune trae fondamento nella necessità di evitare problemi di polarizzazione che si avrebbero utilizzando una rete di common mode feedback globale. Tali problemi derivano dall'utilizzo della topologia TCR in amplificatori fully differential a due stadi e hanno rivestito un ruolo determinante anche in relazione al progetto del secondo stadio. La trattazione approfondita della questione è onerosa e richiede ancora l'introduzione di troppi elementi per essere svolta a questo punto ed è stata presentata nel paragrafo di questa sezione dedicato al controllo della tensione di modo comune. Come in questo caso, si chiederà talvolta al lettore una certa pazienza e la temporanea accettazione di alcune scelte progettuali che sembreranno non trovare una piena ed immediata giustificazione, semplicemente perchè questa sarà chiarita in seguito.

In figura 3.14 sono visibili tutti i rapporti di forma dei transistori. Per i MOS M1, M2, M3 ed M4 sono stati utilizzati i transistori a ossido spesso del designkit ( $t_{ox}$  pari a 6.5nm e  $L_{min}$  di 0.35 $\mu m$ ) per evitare problemi di saturazione dell'interno blocco amplificatore dovuti a eventuali correnti di gate iniettate su pseudoresistori da 10T $\Omega$  mentre per tutti gli altri sono stati utilizzati MOS ad ossido sottile ( $t_{ox}$  pari a 2.2nm con  $L_{min}$  di 0.13 $\mu m$ ). Tutti i transistori sono polarizzati al di sotto della loro soglia  $V_{th}$ , pari a 249mV per quelli ad ossido sottile e 274mV per quelli ad ossido spesso. I rapporti di forma dei MOS di ingresso sono elevati allo scopo di garantirne sempre il funzionamento in regime di sottosoglia: la tensione  $V_{GS}$  è di 112mV per  $M_{1,2}$  e di 148mV (valore assoluto) per  $M_{3,4}$ . Il PMOS e i due NMOS di coda hanno invece un ossido più sottile ed il rapporto di forma di questi ultimi è stato dimensionato per avere una  $V_{GS}$  di circa 250mV portando la corrente di ramo stabilita dalle specifiche di rumore. Con la scelta della corrente del primo stadio è anche fissato il valore della capacità di compensazione alla Miller con cui si desidera stabilizzare il circuito. Avendo infatti la necessità di garantire un limite superiore di banda di almeno 10kHz ed essendo questo fissata dal rapporto tra il prodotto guadagno banda dell'OTA, definito dalla relazione :

$$GBWP_{OTALNA} \approx \frac{g_{m1}}{2\pi C_c} \quad (3.45)$$

ed il guadagno del LNA, si ha:

$$C_c \approx \frac{g_{m1}}{2\pi G_{LNA} BW} \approx 2pF \quad (3.46)$$

dove con  $g_{m1}$  si intende la transconduttanza equivalente del primo stadio, che nella semplificazione di transconduttanze identiche coincide con la  $g_{m_I}$  di dei singoli MOS di ingresso.

### 3.3.2 Progetto e Dimensionamento del Secondo Stadio

Il secondo stadio dell'OTA dev'essere progettato per offrire la massima dinamica possibile al segnale e dev'essere in grado di pilotare un carico capacitivo  $C_{LLNA}$  di circa 1.5-1.6pF comprensivo di capacità di ingresso del PGA (1.35pF), parassiti al nodo di uscita e carico della rete di reazione. Esso può essere genericamente schematizzato come uno stadio transconduttore  $g_{m2}$  con relativa resistenza di uscita e integrato in un sistema reazionato negativamente e compensato alla Miller come quello indicato in figura 3.17:

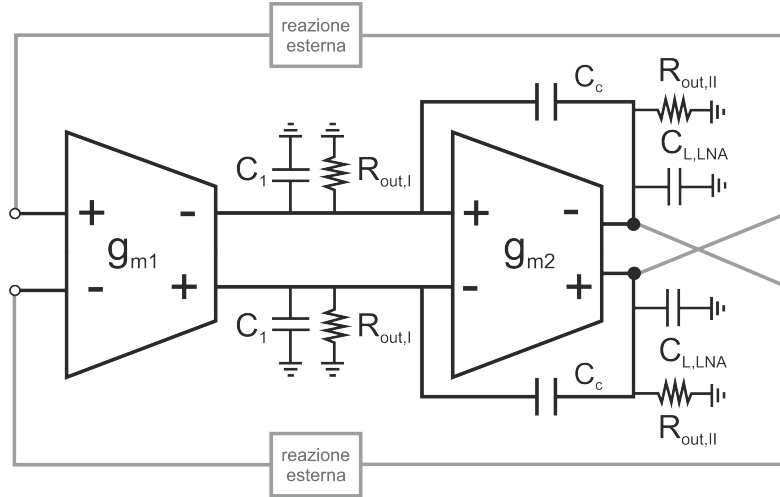


Figura 3.17: Schema equivalente di un amplificatore fully differential a due stadi.

Innanzitutto è importante chiarire quale relazione intercorra tra la transconduttanza del secondo stadio (ancora da definire) a quella del primo stadio allo scopo di approntare un dimensionamento in grado di garantire la stabilità del sistema. Si consideri che lo stadio in esame compensato alla Miller presenta un polo a bassa frequenza, uno ad alta frequenza ed uno zero destro definiti con buona approssimazione dalle equazioni:

$$f_{PLF} \approx \frac{1}{2\pi C_{CLNA} g_{m2} R_{out,II} R_{out,I}} \quad (3.47)$$

$$f_{PHF} \approx \frac{g_{m2}}{2\pi C_{LLNA}} \quad (3.48)$$

$$f_{zRP} \approx \frac{g_{m2}}{2\pi C_{CLNA}} \quad (3.49)$$

$$(3.50)$$



La relazione che si sta ricercando è imposta da una precisa condizione sul margine di fase, da mantenere superiore a 60 gradi. Tenendo conto che esso si valuta alla frequenza pari al rapporto tra il prodotto guadagno banda dell'OTA intero  $GBWP_{LNAOTA}$  e il guadagno dell'intero blocco  $G_{LNA}$ , e indicando questa frequenza con  $GBWP_{LNA}$ , si ha :

$$\phi_{PM} = \pi - \operatorname{atan}\left(\frac{GBWP_{LNA}}{f_{PLF}}\right) - \operatorname{atan}\left(\frac{GBWP_{LNA}}{f_{PHF}}\right) - \operatorname{atan}\left(\frac{GBWP_{LNA}}{f_{zRP}}\right) \quad (3.51)$$

Considerando che il primo polo interviene a frequenze basse e le espressioni di secondo polo e zero derivanti dal pole splitting per effetto Miller, si ottiene la condizione:

$$\frac{\pi}{2} - \operatorname{atan}\left(\frac{C_{LNA}g_{m1}}{G_{LNA}C_{CLNA}g_{m2}}\right) - \operatorname{atan}\left(\frac{g_{m1}}{G_{LNA}g_{m2}}\right) \leq \frac{\pi}{3} \quad (3.52)$$

Risolvendo la disequazione dati i parametri del sistema e per una capacità di compensazione di 2pF (quella scelta) si ottiene la condizione:

$$g_{m2} \geq \frac{g_{m1}}{64} \quad (3.53)$$

Con  $g_{m1}$  e  $g_{m2}$  si sono indicate le transconduttanze equivalenti di primo e secondo stadio, pertanto a seconda della architettura scelta la 3.53 imporrà un vincolo specifico per la corrente di polarizzazione dei MOS del secondo stadio.

Per offrire al segnale la massima dinamica possibile e ridurre al minimo la distorsione anche in presenza di segnali di ingresso di pochi mV la prima topologia che è stata presa in considerazione è stata la common source (una per ramo, dato che è fully differential) direttamente connessa al primo stadio e compensata alla Miller, illustrata in figura 3.18.

Con una opportuna rete di reazione dedicata per controllare il modo comune dell'uscita uno stadio del genere apparirebbe adeguato all'impiego. Successivamente, a causa di problemi di polarizzazione associati al comportamento sul segnale di modo comune che saranno chiariti nel prossimo paragrafo, si è deciso di utilizzare uno stadio differente, precisamente come quello mostrato in figura 3.19 e inserito nella struttura complessiva dell'amplificatore in figura 3.20.

La struttura è simile a quella presentata in [54] ed costituita da quattro rami con due coppie connesse da specchi di corrente PMOS a rapporto unitario. La reazione di modo comune è realizzata con una coppia di NMOS afferenti ai nodi d'uscita. I transistori utilizzati sono ad ossido sottile e i loro rapporti di forma in micron sono: (3/50) per  $M_7$  e  $M_{11}$ , (20/35) per i MOS dello specchio, (1.5/50) per  $M_8$ ,  $M_{12}$  e per i MOS di common mode feedback  $M_{15}$  e  $M_{16}$ . Con l'uscita del primo stadio a circa 250mV i

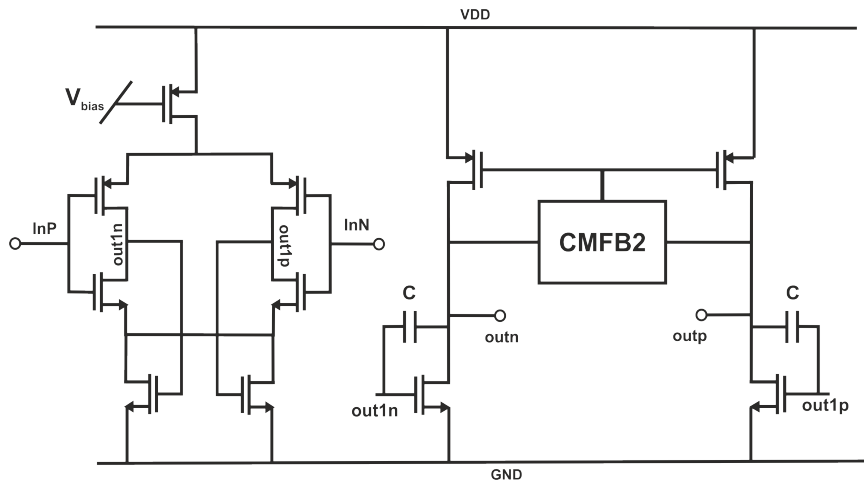


Figura 3.18: Esempio di OTA con stadio di uscita Common Source reazionato in modo dedicato.

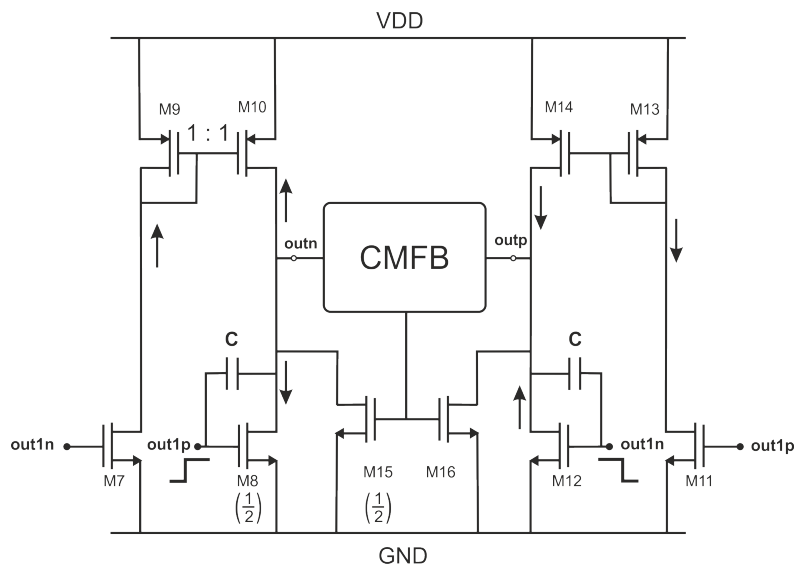


Figura 3.19: Secondo stadio dell'OTA utilizzato per l'LNA.

transistori del secondo stadio lavorano tutti in debole inversione. Si noti che tutti gli NMOS ai nodi di uscita hanno un rapporto di forma pari a un mezzo rispetto agli altri transistori. In assenza di questo splitting, a differenza del source a massa, lo stadio avrebbe un trasferimento di modo comune (CM to CM) nullo ai singoli nodi di uscita ma non sarebbe possibile controllarne esternamente la tensione media. Con la soluzione adottata si ottengono invece i seguenti trasferimenti:

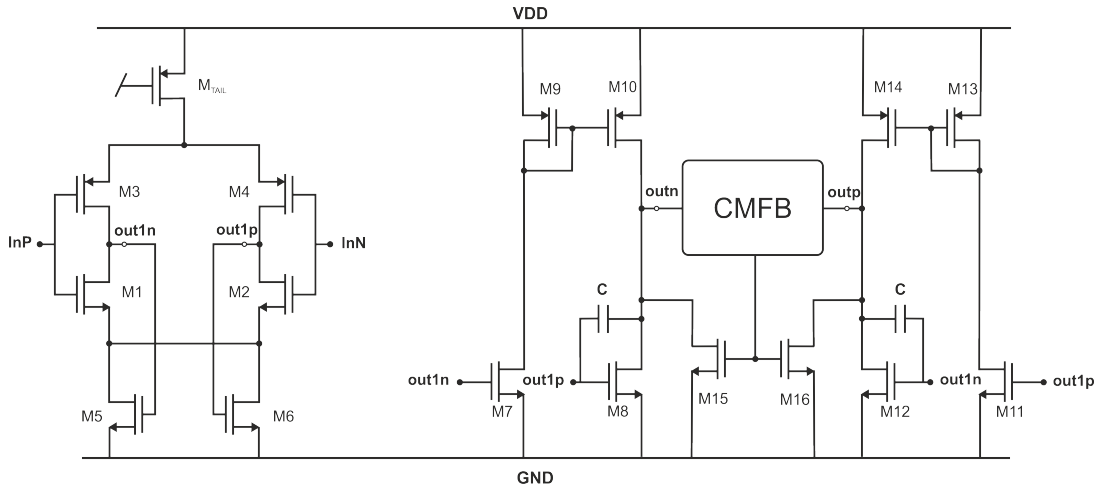


Figura 3.20: Struttura completa di primo e secondo stadio dell'OTA utilizzato per l'LNA.

$$G_{CM-CM_{II}} = \frac{\frac{1}{2}g_{m_{II}}R_{out}}{1 - G_{loop,CM}} \quad (3.54)$$

$$G_{DM-DM_{II}} = \frac{3}{2}g_{m_{II}}R_{out} \quad (3.55)$$

Il guadagno da modo comune a modo comune è generalmente di scarso interesse in amplificatori fully differential perchè non corrompe in modo diretto il segnale utile all'uscita. Esso è tuttavia riportato perchè con una alimentazione molto ridotta aumenta il rischio di riduzione della dinamica del segnale differenziale che può verificarsi in seguito ad uno spostamento del modo comune delle uscite. Con i MOS in debole inversione e assumendo che essi abbiano identica  $g_{m,II}$ , la transconduttanza equivalente dello stadio (su segnale differenziale) è pari a  $\frac{3}{2}g_m$ . Ciò significa che la corrente necessaria a polarizzare ognuno dei 4 rami della struttura dev'essere, in base alla 3.53, pari almeno a 5nA. Data l'esiguità di un simile valore, la sua conseguente scarsa incidenza sul consumo e la difficoltà di generare in modo affidabile una corrente così ridotta è stato deciso un confortevole sovradimensionamento per ottenere un margine di fase superiore, pari a circa 75°. Imponendo una corrente di 50nA in ciascun ramo, la corrente totale assorbita dal secondo stadio è di 200nA. Il guadagno di anello dell'intero amplificatore comprende tuttavia anche la rete di reazione capacitiva esterna il cui effetto sostanziale è una attenuazione inversamente proporzionale al guadagno dell'LNA. Le principali singolarità ed il prodotto guadagno banda dell'amplificatore (comprendente OTA e retroazione) possono essere stimate tenendo conto della compensazione alla Miller effettuata (con 2pF) attraverso le relazioni:

$$f_{PLF} \approx \frac{1}{2\pi C_{CLNA}(\frac{3}{2})g_{m_{II}}R_{out_{II}}R_{out_I}} \approx 280Hz \quad (3.56)$$

$$f_{PHF} \approx \frac{3g_{m_{II}}}{4\pi C_{LLNA}} \approx 215kHz \quad (3.57)$$

$$f_{z_{RP}} \approx \frac{3g_{m_{II}}}{4\pi C_{CLNA}} \approx 150kHz \quad (3.58)$$

$$GBWP_{LNA} \approx \frac{g_{m_I}}{2\pi C_{CLNA}G_{LNA}} \approx 10kHz \quad (3.59)$$

dove con  $g_{m_I}$  si indica la transconduttanza dei MOS del primo stadio (assunta uguale per pmos e nmos), con  $g_{m_{II}}$  quella dei mos del secondo e  $R_{out_I}$  e  $R_{out_{II}}$  le resistenze di uscita di primo e secondo stadio, rispettivamente vicine a  $6.5M\Omega$  e  $16M\Omega$ . Con il dimensionamento effettuato il margine di fase è pari a circa  $75^\circ$ . Considerando il rapporto tra guadagno differenziale (DM to DM) e corrente consumata l'efficienza di questa struttura risulta inferiore del 25% rispetto a quella dell'equivalente (pari corrente totale) stadio source a massa, ma ha un minore trasferimento da modo comune a modo comune. E' qui interessante notare come con i bassi valori di corrente richiesti la scelta della topologia non è guidata da motivi di consumo ma dal controllo della tensione di modo comune dei nodi di uscita.

### 3.3.3 Il Controllo della Tensione di Modo Comune

A differenza di amplificatori di tipo single ended, gli stadi amplificatori completamente differenziali sono caratterizzati da quattro diversi guadagni:

- Un guadagno da modo differenziale a modo differenziale (cui in genere ci si riferisce con il semplice termine differenziale)  $G_{DM-DM}$ ;
- Un guadagno da modo comune a modo comune  $G_{CM-CM}$ ;
- Un guadagno da modo comune a modo differenziale  $G_{CM-DM}$ ;
- Un guadagno da modo differenziale a modo comune  $G_{DM-CM}$ ;

Nel caso di un generico OTA privo di rete di reazione di modo comune si indicano in genere queste grandezze con i simboli  $A_{DD}$ ,  $A_{CC}$ ,  $A_{CD}$ ,  $A_{DC}$ . Mentre i primi due sono generalmente caratteristici anche di topologie ideali di amplificatori, il secondo e il terzo assumono valori non nulli solo in presenza di asimmetrie dello stadio, ad esempio dovute al mismatch. Il rapporto di reiezione del modo comune (CMRR, common mode rejection ratio) in questo tipo di amplificatori si valuta secondo la relazione:

$$CMRR = \frac{A_{DD}}{A_{CD}} \quad (3.60)$$

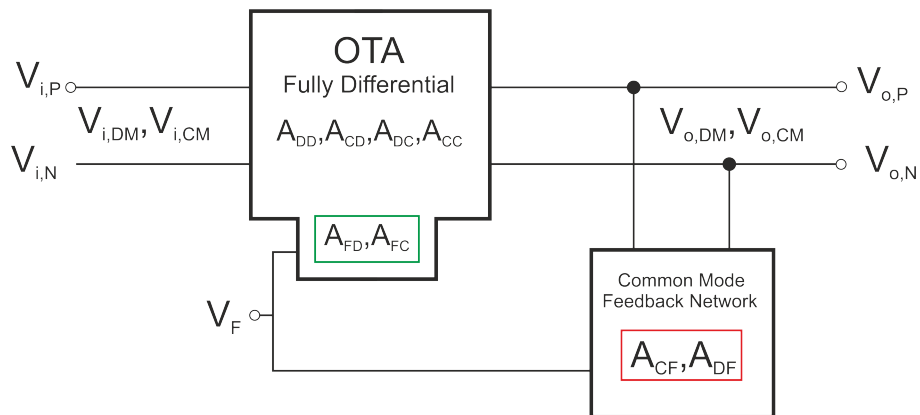
perchè il segnale utile in uscita è letto in modalità differenziale e pertanto sono principalmente i disturbi differenziali in uscita a degradarne la qualità; alla luce di quanto mostrato è chiaro che il principale trasferimento da minimizzare è quello da modo comune a differenziale. I trasferimenti  $A_{DC}$  e  $A_{CC}$  apparentemente non rientrano nel calcolo del CMRR. Come vedremo tra poco, ciò in realtà non è del tutto vero, in ogni caso, è bene che siano minimizzati per non ridurre la dinamica dell'amplificatore verso uno dei due rail e per mantenere stabile il punto di polarizzazione. E' per questi motivi e eventualmente per controllare accuratamente la tensione di polarizzazione di modo comune delle uscite di un amplificatore che si utilizzano reti di reazione di modo comune. Qualsiasi sia l'architettura della rete scelta, essa sarà caratterizzata da un anello di reazione ed un relativo guadagno  $G_{loopCMF}$  che entra in gioco nel modificare i trasferimenti degli amplificatori rispetto alla situazione priva di CMFB, per ottenerne di nuovi:  $A_{DD}^*$ ,  $A_{CC}^*$ ,  $A_{DC}^*$  e  $A_{CD}^*$ . In presenza di rete di reazione di modo comune il CMRR si modifica e si può valutare attraverso la formula:

$$CMRR^* = \frac{A_{DD}^*}{A_{CD}^*} \quad (3.61)$$

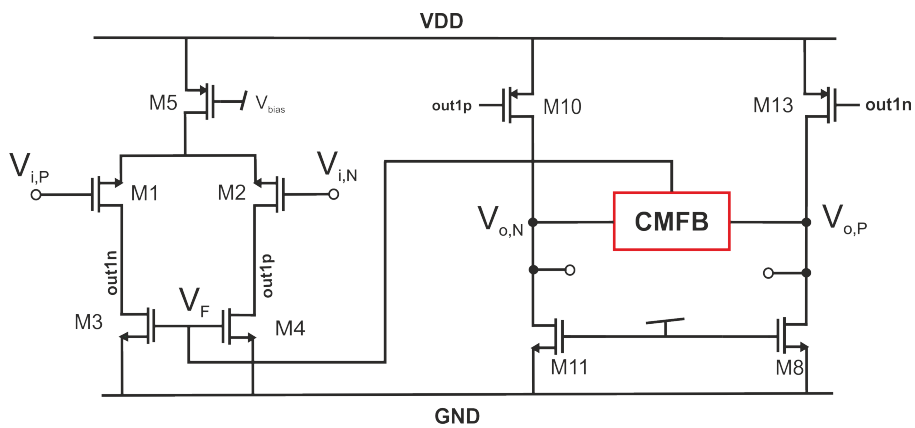
E' utile stabilire la relazione che lega l'espressione del rapporto di reiezione di modo in comune in presenza della relativa rete di reazione, con il CMRR di partenza. A questo scopo si consideri lo schema di amplificatore fully differential di fig. 3.21(a) (con una sua possibile implementazione in 3.21(b)).

Idealmente la rete di reazione evidenziata in rosso legge il segnale modo comune e genera attraverso il trasferimento  $A_{CF}$  una tensione di reazione  $V_F$  attraverso la quale reagisce sull'amplificatore per annullare le variazioni del segnale iniziale. In realtà, a causa dei suoi mismatch interni essa è sensibile anche a segnali di uscita  $V_{o,DM}$  di natura differenziale, che attraverso il trasferimento  $A_{DF}$  concorrono a determinare il valore di  $V_F$ . Inoltre, i mismatch propri dell'amplificatore e in generale (ciò dipende dalle topologie circuitali) di quei componenti dell'amplificatore coinvolti nella rete di CMFB alterano l'effetto della reazione, per cui  $V_F$  non influenza soltanto il modo comune di uscita, come avverrebbe nel caso ideale, ma agisce anche sul modo differenziale attraverso il trasferimento  $A_{FD}$ . L'effetto della retroazione di modo comune non è quindi soltanto quello di sopprimere i disturbi di modo comune e permettere il controllo della tensione media ai nodi di uscita. A causa delle sue non idealità infatti essa produce anche un disturbo del segnale utile in uscita che in generale concorre a degradare il  $CMRR^*$ .

In anello aperto, le relazioni ingresso-uscita rispetto ai segnali differenziali e di modo comune per i blocchi di uno stadio amplificante schematizzati in fig 3.21(a) possono essere descritte dal sistema [46]:



(a) Schema a blocchi di un amplificatore con rete di reazione di modo comune.



(b) Esempio di OTA con rete CMFB riconducibile allo schema in (a).

Figura 3.21: Descrizione schematica di un amplificatore fully-differential con rete di CMFB.

$$V_{o,DM} = A_{DD}V_{i,DM} + A_{CD}V_{i,CM} + V_F A_{FD} \quad (3.62)$$

$$V_{o,CM} = A_{DC}V_{i,DM} + A_{CC}V_{i,CM} + V_F A_{FC} \quad (3.63)$$

$$V_F = A_{DF}V_{o,DM} + A_{CF}V_{o,CM} \quad (3.64)$$

Resta da vedere in che modo, quando il circuito è in anello chiuso, i vari effetti si combinino nel determinare le prestazioni del circuito e il suo  $CMRR^*$ . Coerentemente a quanto spiegato finora è possibile scrivere l'espressione del guadagno d'anello della rete di reazione di modo comune secondo la relazione:

$$G_{loop_{CMF}} = A_{CF}A_{FC} \quad (3.65)$$

E' anche possibile definire il trasferimento che dal modo comune in uscita produce, attraverso la rete di reazione, un disturbo differenziale agli stessi nodi:

$$\Delta G_{loop_{CMF}} = A_{CF}A_{FD} \quad (3.66)$$

Per una descrizione del circuito in anello chiuso il sistema 3.68 può essere riscritto in modo più compatto secondo le equazioni:

$$V_{o,DM} = A_{DD}^*V_{i,DM} + A_{CD}^*V_{i,CM} \quad (3.67)$$

$$V_{o,CM} = A_{DC}^*V_{i,DM} + A_{CC}^*V_{i,CM} \quad (3.68)$$

in cui non compare più la tensione interna  $V_F$  e si sono utilizzate le espressioni dei trasferimenti in anello chiuso. Le espressioni di questi ultimi si possono ricavare dal confronto dei due sistemi in funzione dei trasferimenti in anello aperto e dei guadagni  $G_{loop_{CMF}}$  e  $\Delta G_{loop_{CMF}}$ , e sostituendo le espressioni ottenute per  $A_{CD}^*$  e  $A_{DD}^*$  nella 3.61 si ottiene:

$$CMRR^* = CMRR \cdot \frac{A_{DD}(1 - G_{loop_{CMF}}) + A_{DC}\Delta G_{loop_{CMF}}}{A_{DD}(1 - G_{loop_{CMF}}) + CMRR \cdot A_{CC}\Delta G_{loop_{CMF}}} \quad (3.69)$$

relazione che lega il  $CMRR^*$  ( in anello chiuso ) con il CMRR in assenza di reazione di modo comune. Se si considerano CMRR (senza reazione) e  $G_{loop_{CMF}}$  sufficientemente elevati e un guadagno  $A_{DC}$  trascurabile, la 3.69 può essere approssimata a:

$$CMRR^* = \frac{A_{DD}(1 - G_{loop_{CMF}})}{A_{CC}\Delta G_{loop_{CMF}}} \quad (3.70)$$

Spiegazioni più dettagliate circa il ruolo delle reti di CMFB nell'influenzare il CMRR di amplificatori differenziali si trovano in [46] e [45]. In questo contesto è sufficiente sottolineare come dall'ultima espressione del CMRR si deduca l'importanza della scelta di una rete di common-mode feedback forte, lineare e poco sensibile ai mismatch ma anche l'importanza di contenere il guadagno  $A_{CC}$  nella scelta della topologia di OTA. Questo non solo per stabilizzare la tensione di uscita in polarizzazione e per massimizzare la dinamica di uscita ma anche per non degradare il CMRR.

Dopo questa breve introduzione necessaria a ricordare alcuni fondamentali degli amplificatori fully differential e ad illustrare soprattutto i possibili esiti di non idealità statistiche, si può procedere ad affrontare alcune problematiche di carattere completamente deterministico che sono spesso trascurate ma che in questo contesto acquistano rilevanza. Il controllo o la regolazione della tensione di modo comune dei nodi di uscita dell'amplificatore sono funzionali all'ottenimento del massimo swing possibile. Anche i nodi di uscita del primo stadio, sebbene non siano soggetti a oscillazioni ampie, è bene che siano polarizzati a una tensione che garantisca il buon funzionamento dell'intero amplificatore. Attraverso opportune reti di retroazione di modo comune è possibile forzare la tensione media di ciascuna coppia differenziale di questi nodi ad un valore di riferimento desiderato. Nel caso di questo progetto tale valore è rappresentato da metà della alimentazione, cioè 0.25V. Almeno in fase di prototipazione, si desidera poter regolare questo valore dall'esterno attraverso un pin.

L'OTA del LNA è un fully differential a due stadi con una retroazione esterna ( capacità e pseudoresistori MOS ) che impone l'uguaglianza in polarizzazione tra i potenziali dei nodi di uscita del secondo stadio e i nodi di ingresso del secondo. Mentre su segnale differenziale questa reazione può sempre essere resa negativa semplicemente incrociando adeguatamente le connessioni tra i rami, su segnale di modo comune il segno della retroazione dipende dalla topologia degli stadi interni. Un amplificatore fully differential costituito dalla cascata di due normali stadi invertenti, ad esempio il primo seguito da un source a massa, su segnale di modo comune ha un trasferimento non invertente e la retroazione costruita esternamente per il modo differenziale costituisce al contrario un feedback di modo comune positivo. Il principale problema di una configurazione di questo tipo non è tanto la stabilità in frequenza. In genere il guadagno dell'anello è inversamente proporzionale al guadagno differenziale dell'intero blocco di amplificazione, nel caso in esame  $G_{LNA}$ , pertanto una rete interna di common mode feedback basterebbe da sola a imporre una retroazione globale negativa.

Gli elementi critici di una tale configurazione risiedono nel legame tra la polarizzazione, il segno del trasferimento di modo comune dell'intero stadio ed il tipo di reazione di modo comune realizzata. Come evidenziato in [44] questo tipo di amplificatori ( completamente differenziali a due stadi invertenti ) possono andare incontro a problemi di polarizzazione quando la rete di reazione di modo comune agisce dalle uscite del secondo stadio (common-mode sense) ad un transistor di coda del primo stadio, ed in particolare



quando il trasferimento di modo comune complessivo è non invertente. Allo scopo di comprendere meglio il problema si consideri inizialmente lo stadio in figura 3.22. Il punto di lavoro uno stadio così configurato dipende dall'interazione tra l'anello di reazione negativa di modo comune interno e quello positivo esterno.

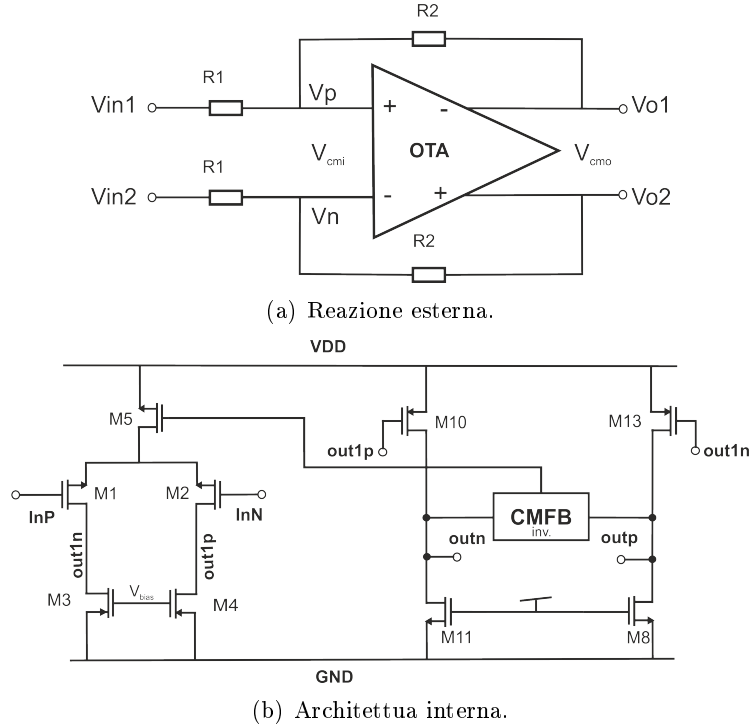


Figura 3.22: Generico amplificatore fully differential reazionato in modo resistivo (da [44]).

Immaginando di portare ai nodi di ingresso  $V_{in1}$  e  $V_{in2}$  (Fig. 3.22(a)) un segnale costante pari alla tensione di riferimento di metà dinamica  $V_{REF}$ , e di chiamare  $V_{cmi}$  e  $V_{cmo}$  rispettivamente la tensione in continua di modo comune ai gate di M1 e M2 e all'uscita, vale la relazione:

$$V_{cmo} = \left(1 + \frac{R_2}{R_1}\right)V_{cmi} - \frac{R_2}{R_1}V_{REF} \quad (3.71)$$

La rappresentazione grafica di questa relazione per due valori differenti del rapporto  $\alpha$  tra  $R_1$  e  $R_2$  è mostrata in figura 3.23. Nella stessa figura è mostrato un tipico andamento della curva ingresso-uscita dell'amplificatore in anello aperto realizzata ponendo all'ingresso (nodi di gate di M1 e M2) una polarizzazione di modo comune variabile tra massa e alimentazione. L'insieme delle curve continue (una delle due) e tratteggiate costituiscono le condizioni imposte da anello di reazione esterno e trasferimento interno e ogni loro punto di incontro rappresenta un possibile punto di lavoro. La curva tratteggia-

ta mostra un andamento crescente per via della natura non invertente dell'amplificatore su segnale di modo comune. Essa inoltre non sono lineari perchè in corrispondenza di tensioni di ingresso molto alte il PMOS di coda va in zona ohmica, compromettendo la reazione di modo comune e portando ad alimentazione le uscite, mentre per tensioni di ingresso al limite inferiore della dinamica la reazione resta attiva mantenendo l'uscita vicino alla tensione di riferimento.

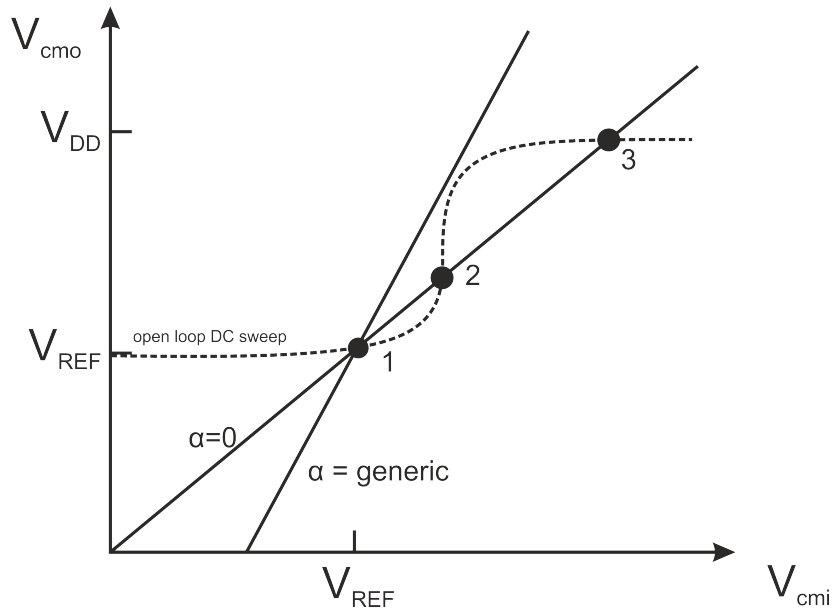
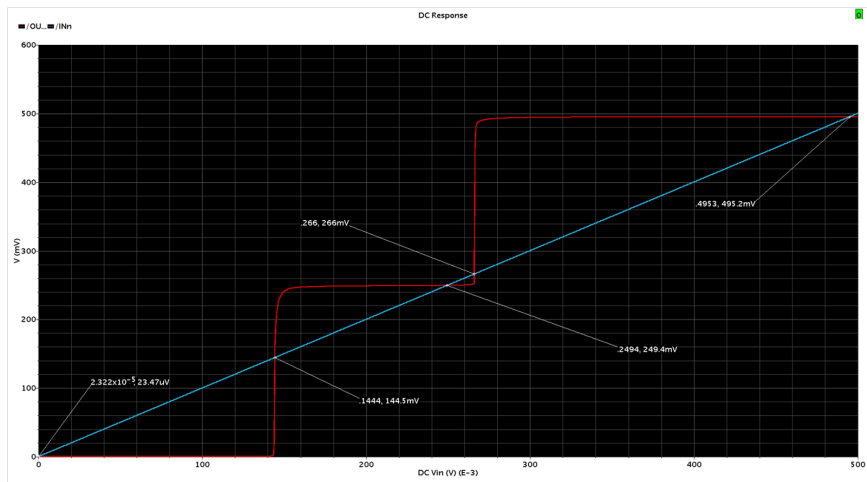


Figura 3.23: Individuazione grafica dei punti di lavoro di un amplificatore fully differentiale semplice a due stadi (da [44]).

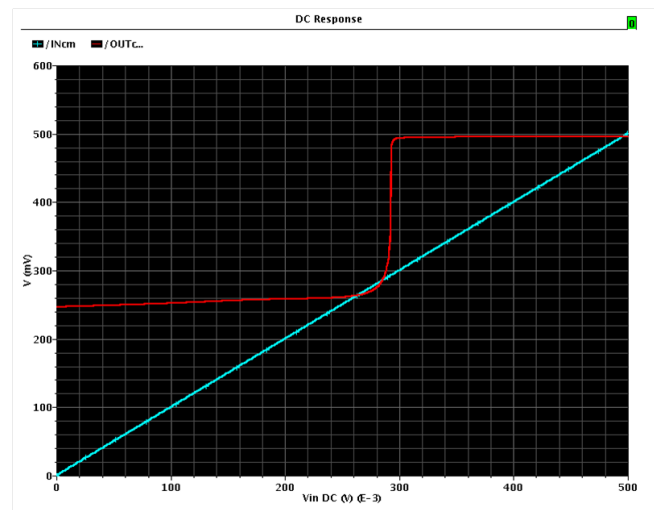
Si nota immediatamente dal grafico che per una configurazione a buffer, come quella del LNA in continua, esistono punti di lavoro multipli. Le simulazioni su questo tipo di circuiti confermano l'analisi. In particolare in figura 3.24(a) e 3.24(b) sono riportate curve analoghe per topologie di amplificatori del tipo di Fig. 3.22(b) ma rispettivamente con ingresso a NMOS e a PMOS (sempre reazionati sul transistore di coda), alimentati tra 0 e 0.5V.

Anche le simulazioni in DC e transitorio (in accensione) confermano come il circuito si stabilizzi spesso in un punto di lavoro inadatto al funzionamento. Raggiunto tale indesiderato equilibrio, lo stadio non è più in grado di uscirne nemmeno se si forza la polarizzazione imponendo una tensione di riferimento attraverso la rete di reazione di modo comune. Non potendo cambiare la struttura della reazione esterna l'unica possibilità è agire sulla topologia dell'OTA. Sempre in 3.22 è indicata come soluzione al problema lo spostamento della reazione di modo comune interna dal transistore di coda ai gate dei transistori di carico del primo stadio, secondo lo schema di fig 3.25.

In questo modo l'entrata in zona ohmica del transistore di coda non comprometterebbe l'azione della rete di modo comune. Tale soluzione però non è praticabile nel



(a) Topologia con ingresso a nmos.



(b) Topologia con ingresso a pmos.

Figura 3.24: Risultati di simulazioni in Cadence Virtuoso che mostrano punti di lavoro multipli.

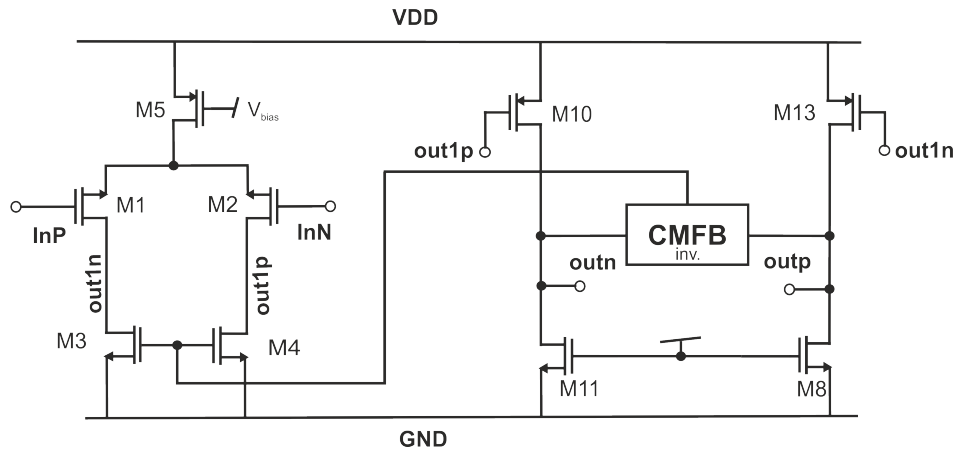


Figura 3.25: Soluzione di CMFB agente sui MOS di carico del primo stadio, proposta in [44] per garantire la polarizzazione dello stadio.

caso dell'OTA di fig. 3.13 dove l'NMOS della coda inferiore ed il PMOS della superiore costituiscono gli unici gate sui quali è possibile reagire. Ciò che invece si può realizzare è una doppia reazione di modo comune, una per stadio attraverso due differenti reti esterne. Come già spiegato in uno dei precedenti paragrafi, la compensazione dell'anello di reazione di modo comune del primo stadio sarebbe stata complicata a causa della sua elevata impedenza di uscita e per questo motivo le è stata preferita una reazione locale ( Fig. 3.14). Al contrario nel secondo stadio è stata utilizzata una rete di retroazione esterna e per escludere completamente la possibilità di punti operativi multipli si è scelta l'architettura già mostrata in fig. 3.19, differente dal source a massa. Questa infatti ha un trasferimento di modo comune (CM-CM) non invertente, pertanto il trasferimento complessivo di modo comune dell'OTA progettato è invertente ed è esclusa ogni possibilità di punti operativi multipli derivanti dall'entrata in zona ohmica dei transistori di coda del primo stadio, come illustrato dal grafico di fig. 3.26.

La rete di reazione di modo comune utilizzata per il secondo stadio è costituita da un singolo OTA ed è mostrata in Fig. 3.27 con il relativi dimensionamento dei transistori (le cifre sono in micron).

La tensione di modo comune di uscita è rilevata al nodo di interconnessione di due pseudoresistori PMOS rappresentati dai transistori  $M_{X_1}$  e  $M_{X_2}$  ed è riportata direttamente al gate di uno stadio differenziale CMOS per il confronto con la tensione di riferimento  $V_{REF}$ , regolabile esternamente. La tensione di reazione  $V_F$  è applicata ai gate dei due mos di reazione  $M_{15}$  e  $M_{16}$  di fig. 3.19 Con questa configurazione è possibile leggere bene il modo comune anche in presenza di grandi segnali differenziali senza che la rete esca dalla zona di funzionamento lineare. La reazione tende ad annullare il segnale errore inteso come la differenza tra la tensione media tra le uscite e la tensione di riferimento, per questo motivo è possibile regolare la prima al valore desiderato semplicemente imponendo  $V_{REF}$  attraverso un pin esterno (nel chip prototipo) o un riferimento integrato. Lo stadio differenziale lavora con transistori sottosoglia ed una corrente di

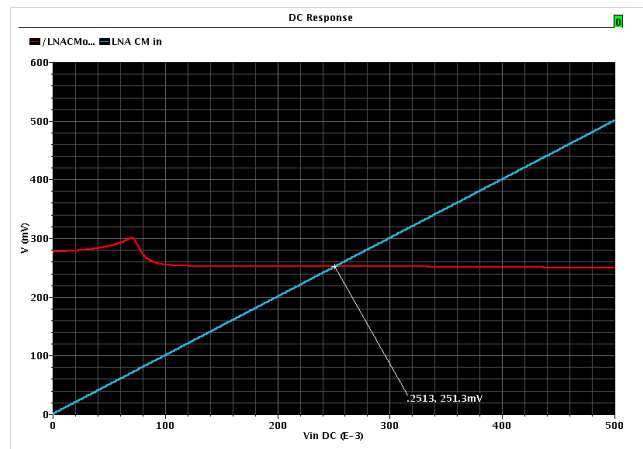


Figura 3.26: Grafico che mostra il singolo punto operativo del LNA.

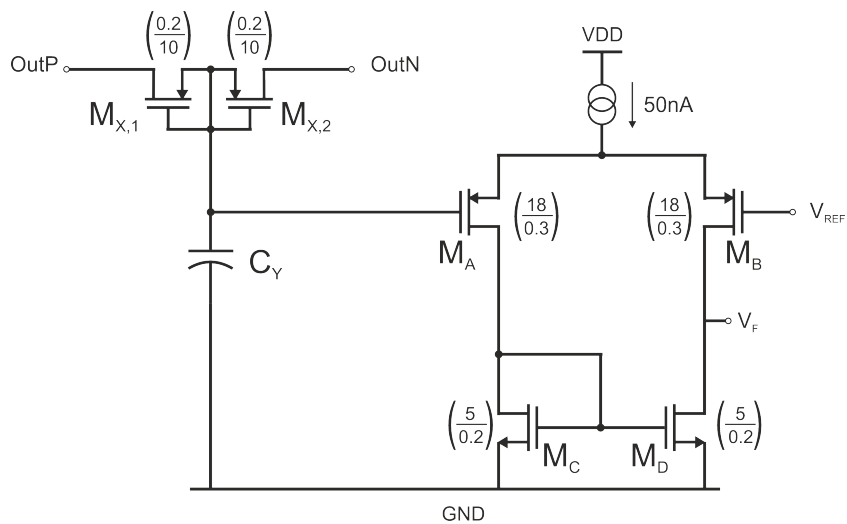


Figura 3.27: Rete di reazione di modo comune del secondo stadio del LNA.

coda di 50nA, cioè 25nA per ramo. Il Modello equivalente del secondo stadio del LNA reazionato rispetto ai segnali di modo comune è schematizzato in figura 3.28.

Si esamini per cominciare la rete che permette la lettura della tensione utile: ciascuno pseudoresistore è realizzato con un MOSFET ad ossido sottile e può essere modellizzato come il parallelo della pseudoresistenza  $R_X$  che sintetizza, pari a circa  $11G\Omega$ , e di una capacità  $C_X$  dovuta ai parassiti. La funzione della capacità  $C_Y$  e la scelta del suo valore possono essere comprese grazie all'analisi della stabilità del sistema retroazionato; si consideri inizialmente il caso in cui essa modellizzi semplicemente il parassita al nodo di gate di  $M_A$ .  $R_{OUT_{LNA}}$  ( precedentemente definita  $R_{OUT_{II}}$ ) e  $C_{OUT_{LNA}}$  rappresentano resistenza e capacità di uscita dell'amplificatore, sono pari rispettivamente a circa 1.5pF e

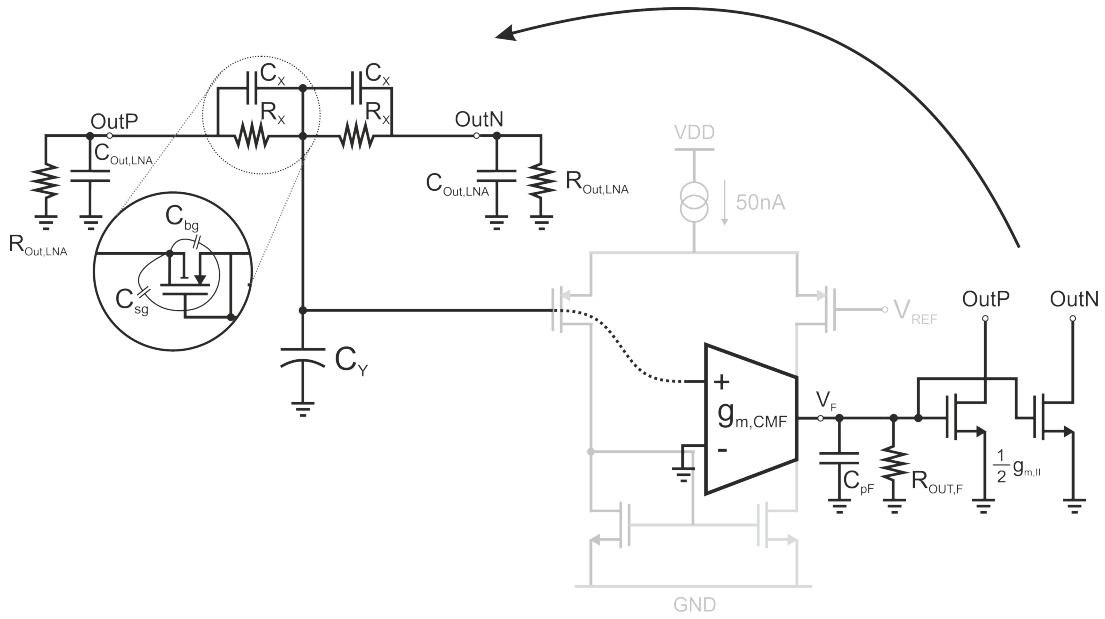


Figura 3.28: Rete di reazione di modo comune del secondo stadio del LNA.

600M $\Omega$ e sono responsabili del polo a più bassa frequenza del guadagno di anello  $G_{CMloop}$ :

$$f_{PLF,CM} = f_{p1} \approx \frac{1}{2\pi C_{Out,LNA} R_{out,LNA}} \approx 180Hz \quad (3.72)$$

Considerando che la capacità di uscita dell'amplificatore è oltre un ordine di grandezza superiore a  $C_Y$  e  $C_X$  ed entra in gioco a frequenze molto inferiori, a queste ultime due capacità sono associati un polo ed uno zero nel semipiano sinistro definiti con buona approssimazione dalle relazioni:

$$f_{PIF,CM} = f_{p2} \approx \frac{1}{2\pi(C_X + \frac{1}{2}C_Y)R_X} \approx 2.8kHz \quad (3.73)$$

$$f_{zCM} = f_z \approx \frac{1}{2\pi(C_X)R_X} \approx 1.6kHz \quad (3.74)$$

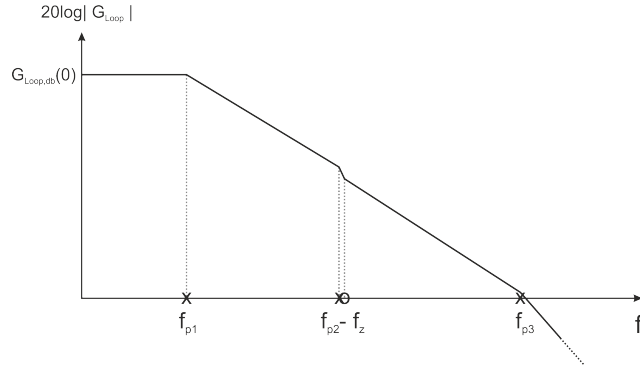
Il polo a frequenza superiore è dovuto alla presenza di un parassita capacitivo all'uscita dell'OTA di reazione, cioè al nodo  $V_F$ , e alla sua interazione con la resistenza a quel nodo, pari circa a 250M $\Omega$ ; la sua posizione è:

$$f_{PHF,CM} = f_{p3} \approx \frac{1}{2\pi C_{pF} R_{outF}} \approx 180kHz \quad (3.75)$$

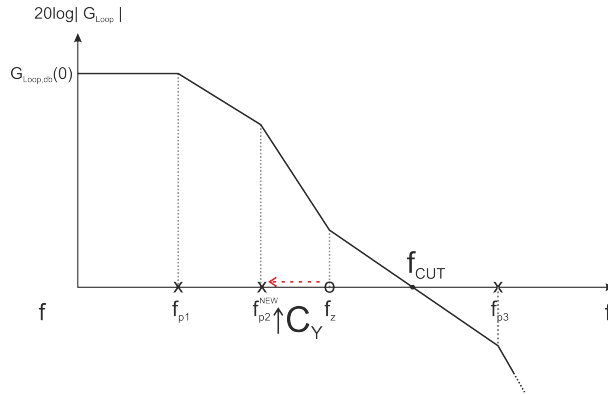
Il valore in continua del guadagno di anello è 72dB ed è così definito:

$$G_{CM_{Loop}} = \frac{g_{m_{II}} R_{out_{LNA}} g_{m_{CMF}} R_{out_F}}{2} \quad (3.76)$$

dove  $g_{m_{CMF}}$  è pari alla metà della transconduttanza dei mos  $M_{A,B}$ . In questa condizione il diagramma di bode del modulo può essere schematizzato come in figura: 3.29(a): Il doppietto polo zero si elide quasi perfettamente e la frequenza alla quale il guadagno attraversa l'asse 0dB cade in vicinanza del polo ad alta frequenza, determinando instabilità.



(a) Situazione con  $C_Y \approx C_X$ .



(b) Situazione con  $C_Y \gg C_X$

Figura 3.29: Schematizzazione del ruolo della capacità  $C_Y$ .

Aumentando il valore della capacità  $C_Y$  è possibile abbassare selettivamente la frequenza del polo intermedio, senza alterare lo zero, per ottenere la situazione di figura 3.29(b) e compensare il sistema. Affinchè il margine di fase del sistema sia superiore a  $45^\circ$  occorre che la frequenza di taglio  $f_{cut}$  sia circa una decade al di sopra della frequenza

cui è localizzato lo zero ma sufficientemente al di sotto del polo a frequenza maggiore. Una stima del suo valore è ottenibile dalla relazione:

$$f_{CUT} = G_{CM_{Loop}}(0) \frac{C_Y}{C_X} f_{p,CM_{LF}} \quad (3.77)$$

I diagrammi di bode di modulo e fase dell'anello di reazione di modo comune del secondo stadio così come è stato realizzato sono riportati in figura 3.30. Per localizzare l'attraversamento a circa 45kHz è stata scelta una capacità  $C_Y$  di 100fF, grazie alla quale è stato anche possibile ottenere un margine di fase prossimo ai 60°, come indicato in figura.

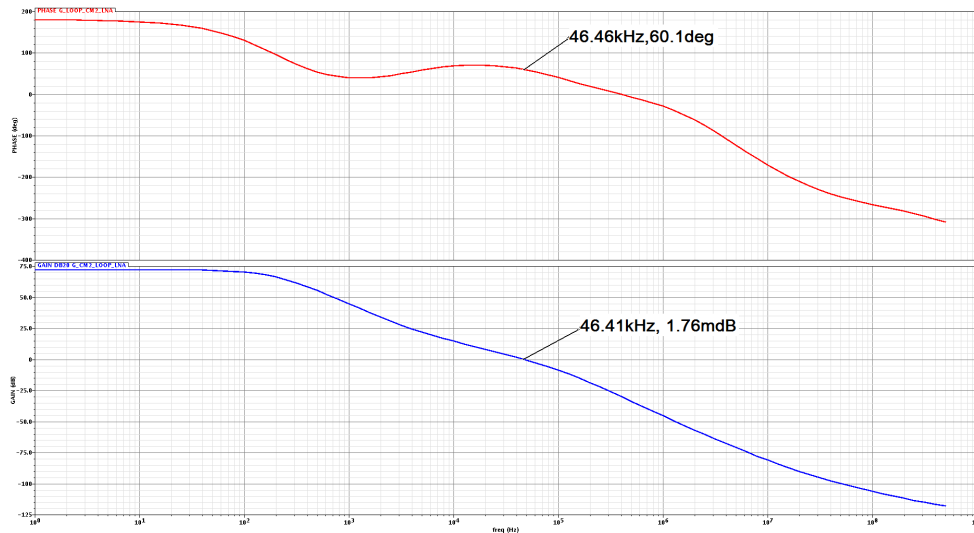


Figura 3.30: Modulo e fase del guadagno di anello della reazione di modo comune del secondo stadio del LNA.

### 3.3.4 Risultati

Il guadagno di anello da modo differenziale a modo differenziale dell'LNA è mostrato in figura 3.31 per parametri tipici dei transistori utilizzati.

Simulazioni montecarlo di variazioni di processo e di mismatch con deviazione standard maggiorata del 10% ( $\sigma = 3.3$  invece del valore 3 consigliato dai manuali di processo) sono state effettuate alla temperatura di 27 gradi centigradi per 1000 realizzazioni allo scopo di verificare l'affidabilità del circuito e definirne le prestazioni; i risultati sono riassunti dalla tabella 3.2.

Va specificato che nel calcolo del consumo, quasi quattro volte inferiore alle richieste iniziali, sono state incluse le correnti di riferimento, considerando tuttavia di utilizzare



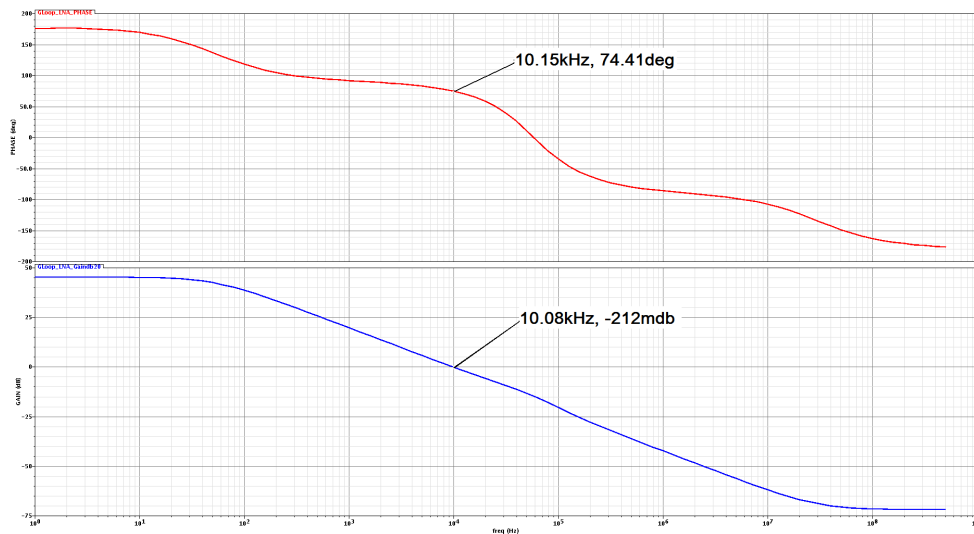


Figura 3.31: Modulo e fase del trasferimento d’anello da modo differenziale a modo differenziale dell’LNA

un riferimento per ciascuna grandezza di corrente su tutti i 64 canali, il loro contributo al consumo è stato suddiviso tra i canali e risulta poco influente sul risultato complessivo. La corrente equivalente assorbita dall’amplificatore è di circa  $1.165\mu\text{A}$ , e se anche si attribuisse l’intero consumo dei riferimenti al singolo amplificatore essa sarebbe pari a  $2.1\mu\text{A}$ : comunque ampiamente inferiore alle richieste.

Non è stato invece possibile simulare le prestazioni di rumore a causa della scarsa accuratezza dei modelli di rumore per i transistori della tecnologia fornita (UMC 130nm) in regime di sottosoglia. Per non appesantire un capitolo già sufficientemente denso, la descrizione di questo problema è riportata in appendice. Si tenga semplicemente presente che i risultati delle simulazioni relative al rumore dello stadio e dei singoli transistori mostravano spesso cifre superiori di un fattore 2 o anche di più al limite superiore teorico. Problemi analoghi e simili conclusioni sono state riportate anche da altri gruppi di ricerca che hanno lavorato con questa tecnologia [52]. In attesa delle misure che potranno essere effettuate successivamente alla fabbricazione del chip prototipo, il rumore riferito in ingresso può essere stimato per via teorica. Il dimensionamento effettuato per il primo stadio dell’amplificatore indicava un limite massimo di rumore pari a  $4.26\mu V_{rms}$ . Esso era tuttavia calcolato su una banda di 15kHz, mentre la banda attuale dell’amplificatore (ed anche la banda dell’intera catena di amplificazione) ha un valore inferiore, di circa 11.5kHz. Tenendo conto di questa banda il rumore riferito in ingresso del LNA non dovrebbe essere superiore a  $4\mu V_{rms}$ . Tale valore potrà essere ulteriormente elevato a causa del parassita ai nodi di gate, ma per questo è necessario attendere il completamento del layout dell’intera catena di amplificazione e filtraggio. Analisi di corner evidenziano sempre un buon funzionamento dello stadio, così come quelle in temperatura comprese tra  $0^\circ\text{C}$  e  $70^\circ\text{C}$ .

Tabella 3.2: Tabella riassuntiva delle prestazioni del Low Noise Amplifier.

Parametri	Specifiche	Risultati
$V_{DD}$	0.5V	0.5V
$G_{LNA}$	40dB	39.91dB $\pm$ 9.89m dB
$f_{HP_{LNA}}$	10mHz	Regolabile
$f_{LP_{LNA}}$	10-15kHz	11.45kHz $\pm$ 989.15Hz ( $f_{-3dB}$ )
CMRR	NA	86.21dB $\pm$ 7.85dB
$P_{tot}$	$\leq 2\mu W$	$\approx 583nW$
$V_{out_{CM}}^{DC}$	250mV	251.507mV $\pm$ 3.892mV
$V_{OS}$	NA	36.281 $\mu V$ $\pm$ 650.67 $\mu V$
Ph. Margin	$\geq 60^\circ$	74.5 $^\circ$

### 3.4 Programmable Gain Amplifier

La sezione è dedicata all'analisi dettagliata del programmable amplifier, comprendente OTA, reazione esterna e rete di reazione di modo comune. Alla fine della sezione sono presentate le prestazioni dello stadio ottenute come risultato di simulazioni montecarlo.

#### 3.4.1 Progetto e Dimensionamento dell'OTA

Al contrario del Low Noise Amplifier, l'amplificatore a guadagno programmabile non presenta vincoli restrittivi rispetto al rumore e pertanto non è stato necessario adottare topologie particolari per il primo stadio dell'OTA interno. Esso è quindi costituito da un normale stadio differenziale con ingresso a NMOS. Essendo a valle di un guadagno di quasi 60dB, l'uscita del PGA dovrebbe essere tale da garantire la massima dinamica possibile. Data anche la ridotta alimentazione, per il secondo stadio dell'OTA è stata scelta una topologia elementare con il minimo di numero di transistori tra i nodi di uscita e i due rail, cioè il source a massa. In figura 3.32 è illustrata l'intera struttura dell'OTA del PGA, con il dimensionamento dei transistori.

Il primo stadio è stato polarizzato con una corrente di 50nA per ridurre al minimo il consumo e rilassare le specifiche sulla transconduttanza del secondo stadio, data anche l'assenza di limitazioni dovute allo slew rate (i nodi di gate degli NMOS di ingresso non sono mai soggetti a grandi segnali). Come nel LNA anche in questo caso i MOS di ingresso sono ad ossido spesso per scongiurare problemi legati allo scorrimento di correnti di gate nei rami degli pseudoresistori e come il LNA anche il PGA lavora con tutti i MOS sottosoglia. In particolare i MOS di ingresso sono polarizzati con una  $V_{GS}$  di circa 200mV. La capacità equivalente dell'array dell'ADC è di circa 1.15fF, ma per motivi di sicurezza il dimensionamento è stato effettuato riferendosi ad una capacità di carico  $C_{L_{PGA}}$  pari a 2 pF e la compensazione realizzata è alla Miller. A questo punto va ricordato che il PGA è progettato per avere un guadagno programmabile tra 0 e 19dB mantenendo la stessa banda, perciò allo switching della capacità esterna di ingresso  $C_{1_{PGA}}$  (funzionale a sintetizzare il guadagno) deve coincidere un'opportuna regolazione

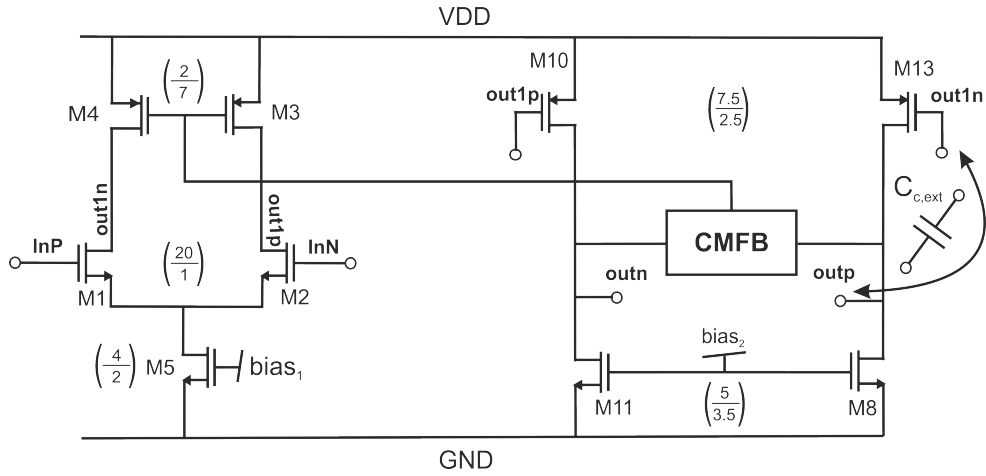


Figura 3.32: OTA utilizzato per il PGA.

della capacità di compensazione. Si consideri inizialmente il caso in cui il guadagno  $G_{PGA}$  è pari a 0dB, praticamente l'equivalente di una configurazione a buffer. In questa situazione il prodotto guadagno banda dell'OTA è approssimativamente:

$$GBWP_{PGA} \approx \frac{g_{mI}}{2\pi C_{CPGA}} \quad (3.78)$$

con  $g_{mI}$  transconduttanza dei MOS di ingresso e  $C_{CPGA}$  la capacità di compensazione (per ora indefinita). Questa grandezza dovrebbe corrispondere all'omologa dell'LNA, insieme alla quale contribuisce a determinare la frequenza passa basso  $f_{LP_{PGA}}$  dell'intero blocco amplificatore quindi va regolata accuratamente. Dato che la frequenza  $f_{LP}$  complessiva della catena è valutata nel punto in cui il trasferimento è di 3dB inferiore al suo valore in banda, per fare in modo che la banda dell'intera catena abbia un limite superiore tra 10kHz e 15kHz,  $GBWP_{PGA}$  (come  $GBWP_{LNA}$ ) dovrebbe essere almeno un fattore  $\sqrt{2}$  superiore. Ciò si traduce in valori di  $GBWP_{PGA}$  compresi tra 14.2kHz e 21.2kHz. Adottando una capacità pari a 4.25pF si ha un  $GBWP_{PGA}$  di 18.7kHz: questo valore corrisponde alla capacità totale dell'array di fig. 3.9(b), che è stata riportata in precedenza nella tabella 3.1. Per avere secondo polo e zero destro a frequenze maggiori del GBWP e scongiurare problemi di stabilità, il source a massa del secondo stadio è stato polarizzato con 200nA (per ramo). Restando nel caso in cui il PGA non guadagna e con la capacità di Miller scelta, le singolarità che contribuiscono a degradare il margine di fase dell'amplificatore sono le seguenti:

$$f_{PLF} \approx \frac{1}{2\pi g_{m_{II}} R_{out_{II}} R_{out_I} C_c} \approx 10Hz \quad (3.79)$$

$$f_{PHF} \approx \frac{g_{m_{II}}}{2\pi C_{L_{PGA}}} \approx 385kHz \quad (3.80)$$

$$f_{z_{RP}} \approx \frac{g_{m_{II}}}{2\pi C_{c_{PGA}}} \approx 180kHz \quad (3.81)$$

con  $g_{m_{II}}$  transconduttanza dei MOS del secondo stadio e  $R_{out_I}$ ,  $R_{out_{II}}$  resistenze di uscita dei due stadi. Secondo queste stime il margine di fase dell'amplificatore dovrebbe essere vicino a  $80^\circ$ , dando quindi garanzia di stabilità. Modulo e fase del trasferimento del PGA in anello aperto per la configurazione a guadagno minimo sono mostrati in figura 3.33.

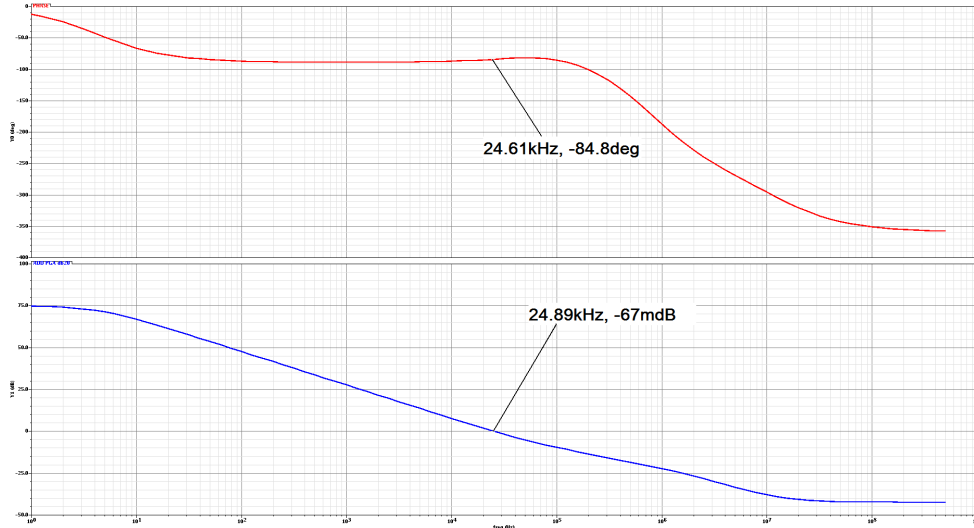


Figura 3.33: Modulo e fase del trasferimento da modo differenziale a modo differenziale in anello aperto del PGA nella sua configurazione a minimo guadagno.

Il margine di fase è ampiamente al di sopra di  $60^\circ$ . Per gli altri possibili guadagni dell'amplificatore la capacità di compensazione è corrispondentemente scalata per garantire banda costante e stabilità; i valori delle capacità sono riportati sempre in tabella 3.1 e la struttura dell'array si può apprezzare in figura 3.9(b). Nella sezione relativa ai risultati si possono trovare i risultati delle simulazioni che descrivono le prestazioni ottenute per il PGA e per l'intera catena di amplificazione.

### 3.4.2 Il Controllo della Tensione di Modo Comune

Il controllo della tensione di modo comune nel PGA assume un'importanza superiore rispetto a quella rivestita nel progetto dell'LNA, in cui il mantenimento della tensione

media delle uscite a 250mV serviva solo per non perdere dinamica. Questo amplificatore si trova infatti a monte del convertitore analogico-digitale. Il convertitore ha un'architettura SAR a redistribuzione di carica e si interfaccia con le uscite del PGA attraverso due interruttori che permettono il campionamento della tensione ai capi di un array capacitivo. Per il buon funzionamento della cascata PGA-ADC, e non solo per la salvaguardia dello swing, occorre che il modo comune delle uscite si trovi a metà della dinamica di alimentazione. La reazione di modo comune agisce dal primo stadio al secondo stadio secondo lo schema di figura 3.25, contattando i gate dei transistori di carico per evitare problemi di punti operativi multipli. Il punto di polarizzazione è evidenziato dall'intersezione della retta di carico con la relazione ingresso uscita (DC sweep) dell'OTA in anello aperto, evidenziate in fig. 3.34.

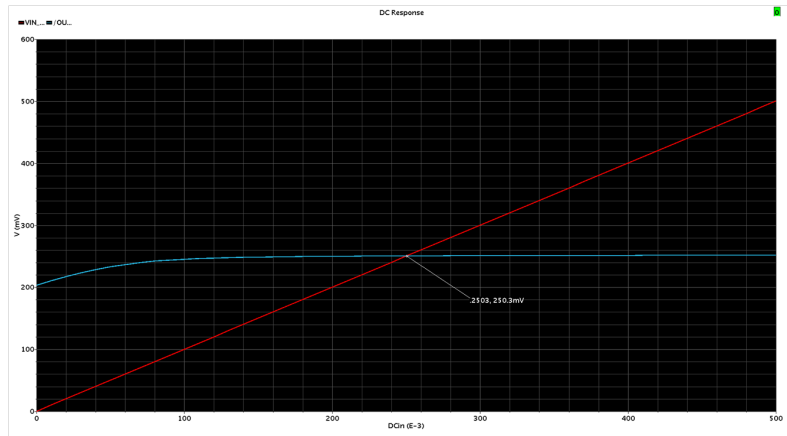


Figura 3.34: Grafico che mostra il singolo punto operativo del PGA.

Il blocco di reazione utilizzato è illustrato in figura 3.35. Analogamente all'omologa rete del secondo stadio dell'LNA essa legge il modo comune al nodo intermedio di una rete di due pseudoresistori PMOS ma la struttura della rete passiva che precede l'OTA è differente. Al nodo di gate di  $M_E$  stavolta è presente solo la capacità parassita mentre in parallelo agli pseudoresistori è posta una piccola capacità  $C_Z$  il cui ruolo sarà chiarito tra poco.

L'anello di reazione su segnale di modo comune può essere schematizzato come in figura 3.36. Esso condivide le stesse singolarità principali del trasferimento da modo differenziale a modo differenziale in anello aperto dell'OTA, con la differenza di avere metà del guadagno in continua. A differenza del trasferimento differenziale tuttavia, esso presenta un polo in più dovuto all'interazione del parassita  $C_{parCM}$  di gate dell'OTA di common mode feedback con i circa  $10G\Omega$  degli pseudoresistori:

$$f_{p,CMHF} \approx \frac{1}{2\pi C_{parCM} R_Z} \quad (3.82)$$

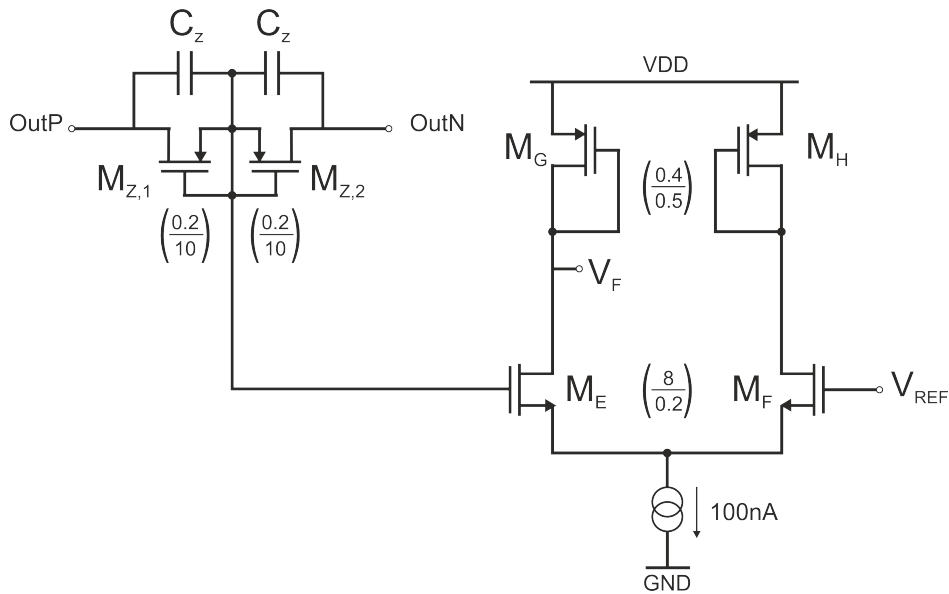


Figura 3.35: Rete di reazione di modo comune del secondo del PGA.

Questa singolarità è nell'ordine della decina di kHz e concorre a determinare la potenziale instabilità del circuito. Per compensare lo stadio anche sul modo comune è stata aggiunta una capacità  $C_Z$ , superiore al parassita  $C_{par_{CM}}$  e che contemporaneamente introduce uno zero sinistro e abbassa la frequenza del polo a valori prossimi a quella dello zero. In questo modo si ottiene una compensazione polo-zero del polo dovuto parassita e la stabilità è di nuovo resa possibile. E' stato sufficiente scegliere  $C_Z$  di 20fF per ottenere un margine di fase ampiamente superiore a  $60^\circ$ .

In figura 3.37 sono mostrati modulo e fase del trasferimento di anello di modo comune dell'OTA del PGA.

### 3.4.3 Risultati

Come fatto in precedenza per l'LNA, anche per il PGA è stata redatta una tabella (Tabella 3.3) che ne riassume le prestazioni simulate confrontandole con le specifiche.

I parametri forniti nella tabella sono il risultato di simulazioni montecarlo con deviazione standard maggiorata del 10% rispetto al valore consigliato dai manuali di processo. Come si può constatare tutte le specifiche di banda e guadagno sono state rispettate mentre il consumo ottenuto è oltre tre volte inferiore a quello richiesto delle specifiche. Nel calcolo del consumo sono state incluse le correnti di riferimento, dividendo opportunamente il loro contributo per il numero di canali.

Come per il low noise amplifier, se anche si attribuissero interamente le correnti di riferimento all'amplificatore, la corrente totale assorbita sarebbe pari a 900nA, la metà della massima permessa dalle specifiche. La frequenza del polo passa basso del PGA in anello chiuso è stata calcolata mediando i valori medi e le deviazioni standard della

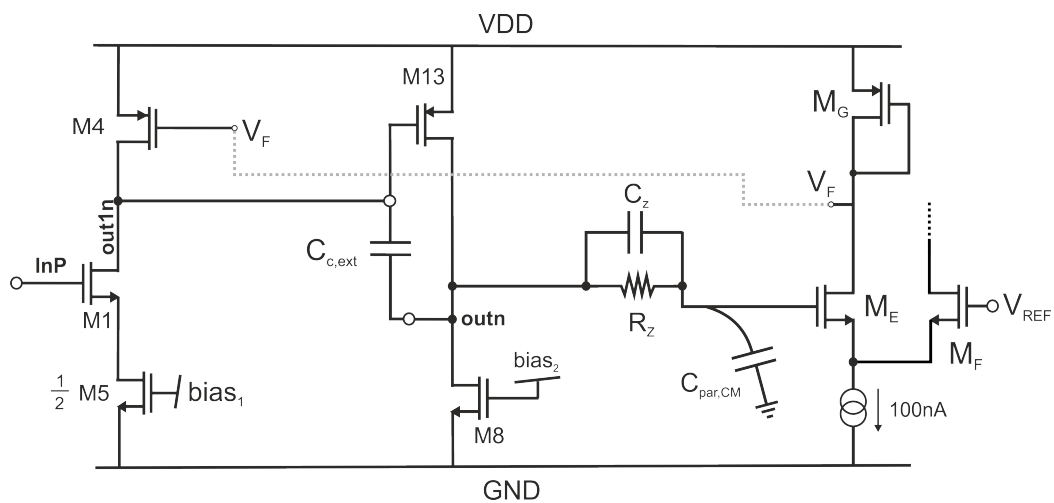


Figura 3.36: Rete di reazione di modo comune del secondo stadio del PGA.

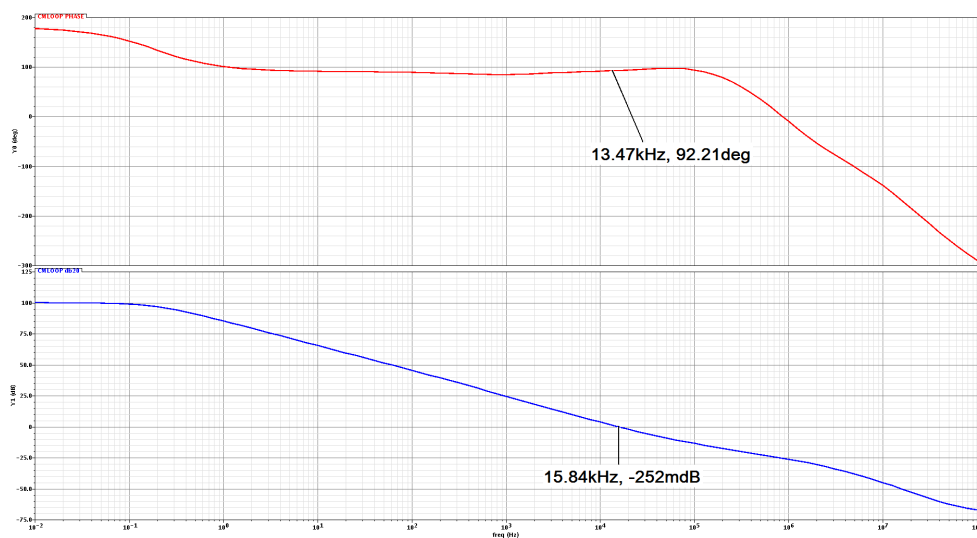


Figura 3.37: Modulo e fase del trasferimento di reazione di modo comune del PGA.

sua posizione per ciascuno dei cinque guadagni ed in ogni caso non si osservano sue significative variazioni tra le configurazioni. Essa è stata dimensionata per essere superiore di circa un fattore  $\sqrt{2}$  rispetto a 10kHz, in realtà, come già spiegato all'inizio del capitolo, questa specifica non è stata presa in considerazione in modo rigoroso e per avere più sicurezza in termini di stabilità si è accettato di ridurla fino a 10.5kHz. Ciò si traduce, come vedremo tra poco, in un valore di frequenza passa basso  $f_{LP-6dB}$  di circa 11.5kHz ma in un decremento del valore della frequenza  $f_{LP-3dB}$  poco al di sotto di 10kHz. Sebbene esso possa apparire un problema, in realtà non lo è per diversi motivi:

Tabella 3.3: Tabella riassuntiva delle prestazioni del PGA.

Parametri	Specifiche	Risultati
$V_{DD}$	0.5V	0.5V
$G_{PGA_0}$	0dB	886mdB $\pm$ 1.3dB
$G_{PGA_1}$	NA	9.2dB $\pm$ 642mdB
$G_{PGA_2}$	NA	13.67dB $\pm$ 9.18mdB
$G_{PGA_3}$	NA	16.57dB $\pm$ 11.9mdB
$G_{PGAMAX}$	20dB	18.75dB $\pm$ 14.7mdB
$f_{HPGA}$	10mHz	Regolabile
$f_{LPGA}$	10-15kHz	10.43kHz $\pm$ 711Hz ( $f_{-3dB}$ )
$CMRR @ G_{PGAMAX}$	NA	54.75 $\pm$ 6.4dB
$P_{tot}$	$\leq 1\mu W$	$\approx 278nW$
$V_{outCM}^{DC}$	250mV	251mV $\pm$ 12mV
$V_{OS}$	NA	157 $\mu V \pm 2.53mV$
Ph. Margin	$\geq 60^\circ$	84 $^\circ$

innanzitutto la gran parte del contenuto informativo del segnale non supera i 5kHz e anche su analisi di corner l'intero stadio mantiene una banda superiore a questo valore e quindi compatibile con l'applicazione; inoltre un restringimento della banda non ha ripercussioni sulla NEF perchè con essa si riduce anche il rumore riferito all'ingresso. La tensione di offset non rappresenta un problema per la conversione analogico digitale per due motivi: innanzitutto a valle della digitalizzazione qualsiasi errore di offset può essere eliminato con opportune elaborazioni digitali, e in secondo luogo non è così elevato da rappresentare un limite per la dinamica. Per quanto riguarda invece la frequenza di taglio passa alto del front-end di amplificazione, essa non è stata riportata accuratamente e per ogni guadagno in quanto si tratta di una grandezza regolabile dall'esterno. In ogni caso in tutte le simulazioni è sempre stato possibile trovare una opportuna tensione (tra 0 e 0.5V) da imporre al gate degli pseudoresistori per ottenere una  $f_{LP}$  di 100mHz. Le prestazioni dell'intera catena di amplificazione e filtraggio sono riportate nella tabella 3.4.

La NEF è stata calcolata considerando un rumore riferito all'ingresso coerente con la banda effettiva ottenuta e calcolato con stime teoriche e considerando un parassita al nodo di gate dei MOS di ingresso dell'LNA pari a quello ottenuto dopo il layout, cioè circa 600fF; la PEF è stata calcolata in base alla NEF ottenuta. Come si può osservare dai numeri elencati nelle tabelle ogni specifica di progetto, non solo è stata soddisfatta ogni specifica, ma sono stati ottenute prestazioni di consumo ampiamente inferiori alle richieste. Continuando a considerare la specifica di consumo massimo per canale di  $4\mu W$ , con i risultati finora ottenuti si potrebbe anche utilizzare un ADC che consuma  $3\mu W$  senza violare i requisiti. Come vedremo nel prossimo capitolo tuttavia, anche l'ADC è stato progettato per ottimizzare il più possibile il consumo e nell'ottica di realizzare un sistema complessivo il più efficiente possibile. Anche le stime di rumore permetto-



Tabella 3.4: Tabella riassuntiva delle prestazioni dello stadio di amplificazione e filtraggio.

Parametri	Specifiche	Risultati
$V_{DD}$	0.5V	0.5V
$G_{TOT_{MIN}}$	40dB	$39.53\text{dB} \pm 42.525\text{mdB}$
$G_{TOT_{MAX}}$	60dB	$58.47\text{dB} \pm 22\text{mdB}$
$f_{HP_{TOT}}$	100mHz	Regolabile
$f_{LP_{TOT}}$	10-15kHz	$11.370\text{kHz} \pm 1.024\text{kHz}$
$CMRR @ G_{TOT_{MAX}}$	NA	$86.4\text{dB} \pm 8.2\text{dB}$
THD	NA	$<1.83\%$
$I_{tot}$	$\leq 6\mu A$	$\approx 1.72\mu A$
$P_{tot}$	$\leq 3\mu W$	$\approx 860nW$
$V_{out_{CM}}^{DC}$	250mV	$251\text{mV} \pm 12\text{mV}$
$V_{OS}$	NA	$36.281\mu V \pm 650.67\mu V$
$V_{noise,in}$	$< 5\mu V_{rms}$	$\approx 3.75\mu V_{rms}$ (stima teorica)
NEF	$< 2$	$\approx 1.8$
PEF	$< 2$	$\approx 1.62$

no un certo ottimismo circa il risultato finale, tuttavia a causa dell'inadeguatezza dei modelli al regime di funzionamento in cui sono stati utilizzati i transistori nessuna conclusione potrà essere tratta prima delle misure. Successivamente alla fase di layout, le simulazioni sono state ripetute per verificare che ogni parametro non subisse variazioni significative e che il circuito funzionasse correttamente. Anche l'esito di queste simulazioni è stato soddisfacente e nella prossima sezione verrà illustrato il layout del front-end di amplificazione comprendente sia LNA che PGA.

### 3.5 Layout

L'obiettivo del progetto di cui questo lavoro è parte consiste nella realizzazione di un intero sistema a 64 canali comprendente front-end di conversione analogico digitale e blocco di trasmissione. La tecnologia utilizzata mette a disposizione un pitch quadrato di  $25\text{mm}^2$  includendo i pads, ossia circa  $20\text{mm}^2$  di effettiva superficie integrabile con unità funzionali del sistema. Considerando la complessità del blocco di trasmissione UWB e la necessità di una vasta gamma di circuiti di servizio, è ragionevole pensare che l'insieme dei 64 canali del front-end non debba superare il 60% di tale area, cioè il limite di  $12\text{mm}^2$ . Per questo motivo può essere considerata come specifica massima di area utilizzabile dal singolo canale (LNA, PGA e ADC) la cifra di riferimento di  $0.1875\text{mm}^2$ . Queste cifre saranno richiamate con più nella sezione relativa al layout dell'ADC, in cui verrà mostrato anche il layout dell'intero front-end, si tenga però presente che approssimativamente ciascun blocco non dovrebbe superare troppo una frazione pari ad un terzo dell'area dedicata al canale, cioè non più di  $0.0625\text{mm}^2$ .

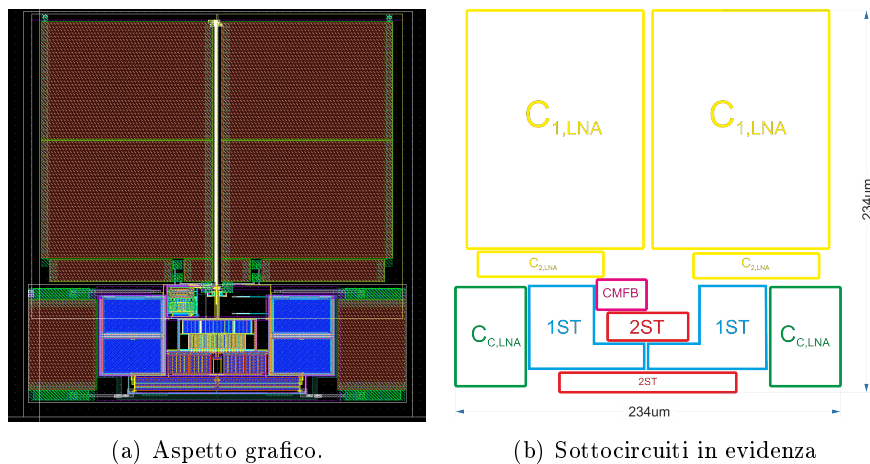


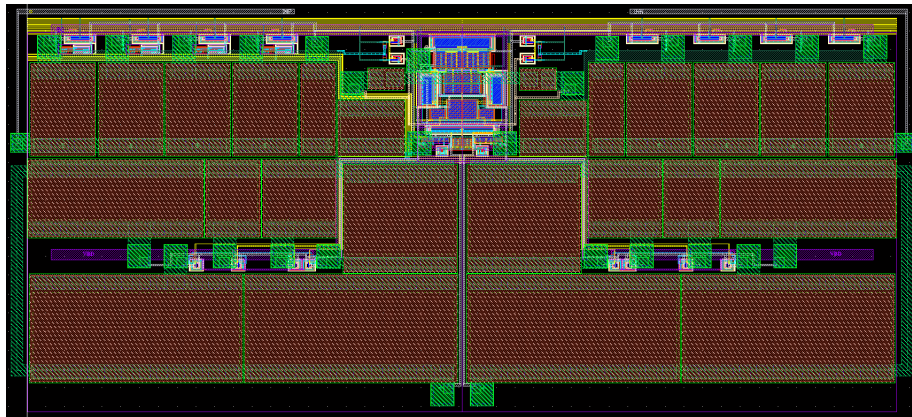
Figura 3.38: Layout del LNA.

Il figura 3.38(a) è mostrato il layout del LNA mentre a fianco, la 3.38(b) ne riproduce la struttura evidenziandone i componenti e le aree funzionali. Da notare è la simmetria della struttura, volta a garantire la minima interazione possibile tra segnali differenziali e di modo comune. La dimensione dell'amplificatore è di  $234\mu\text{m} \times 234\mu\text{m}$  per un'area di  $0.055\text{mm}^2$  occupata per più del 50% dalle grandi capacità di ingresso  $C_{1,LNA}$ . Il contributo delle connessioni alla capacità parassita all'ingresso del primo stadio è complessivamente di  $80\text{fF}$ . Questa quantità si somma ai circa  $500\text{fF}$  di capacità parassita già modellizzata nello schematico per un totale vicino al valore preventivato di circa  $500\text{fF}$ , con scarse conseguenze sulle prestazioni di rumore.

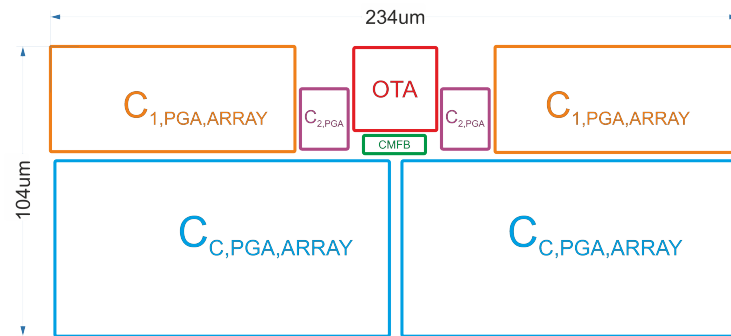
Il layout del PGA è illustrato e schematizzato nelle figure 3.39(a) e 3.39(b). Esso ha dimensioni più piccole rispetto al LNA, pari a  $234\mu\text{m} \times 105\mu\text{m}$ .

I risultati delle simulazioni post layout non evidenziano significativi scostamenti rispetto ai dati delle simulazioni montecarlo presentati nelle precedenti sezioni e i parassiti sommati ai nodi di guadagno non determinano situazioni critiche rispetto alla stabilità delle configurazioni retroazionate di entrambi gli amplificatori. In figura 3.40 è rappresentato il risultato della simulazione post layout dell'intera catena di amplificazione e filtraggio. L'immagine è stata ottenuta simulando il sistema con parametri tipici per i transistori e le capacità della tecnologia ma non sono stati utilizzati generatori ideali se non per modellizzare le tensioni imposte esternamente al chip. Per tutte le altre tensioni e correnti di riferimento sono stati utilizzati i circuiti di servizio progettati appositamente per l'integrato e che verranno descritti nell'ultimo capitolo. Sono evidenti cinque distinti trasferimenti ciascuno corrispondente ad un'opportuna selezione delle capacità di ingresso e di compensazione del PGA.

In figura 3.41 è mostrato il segnale di uscita della catena di amplificazione in corrispondenza di ciascun guadagno selezionabile e dato un segnale sinusoidale di ingresso a  $2\text{kHz}$  con una ampiezza di  $250\mu\text{V}_{pp}$ . La dinamica di uscita ottenuta è di  $\text{mV}_{pp}$  accettando una distorsione armonica totale dell'1.83%.



(a) Aspetto grafico.



(b) Sottocircuiti in evidenza

Figura 3.39: Layout del PGA.

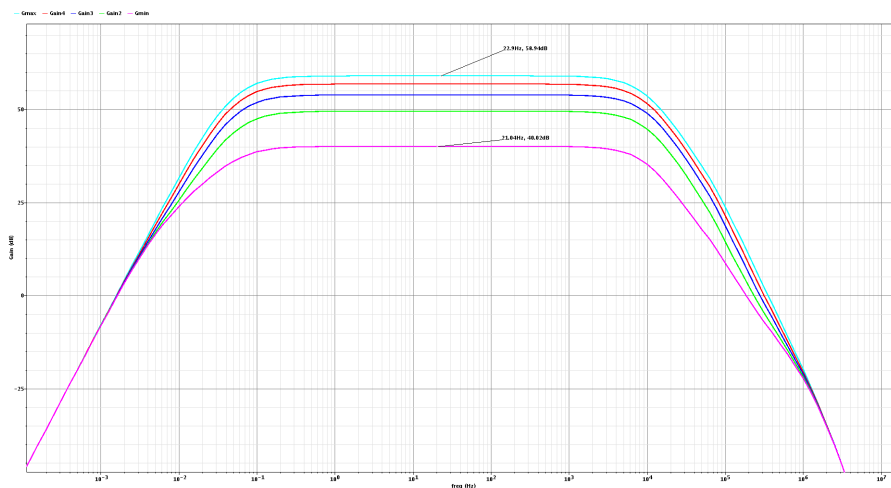


Figura 3.40: Trasferimenti programmabili della catena di amplificazione e filtraggio.

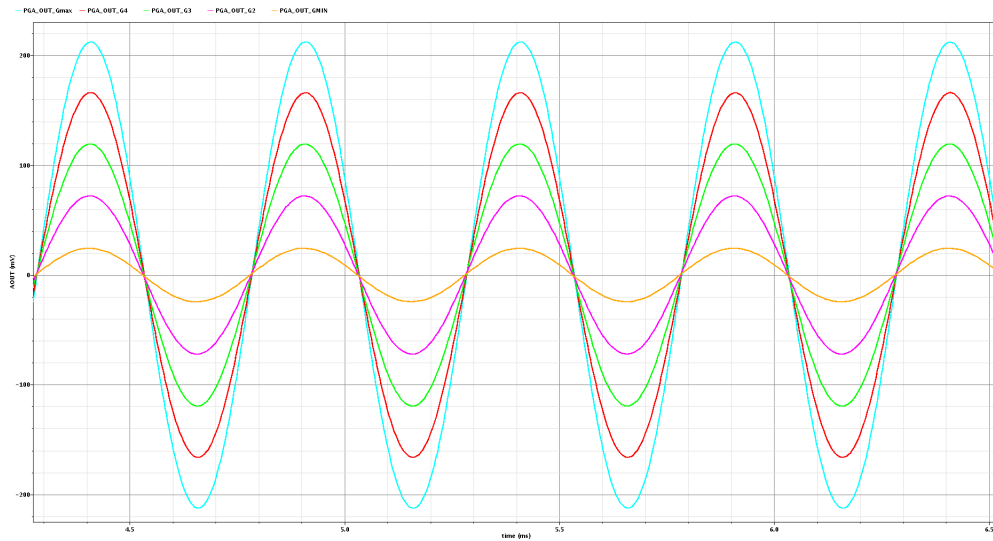


Figura 3.41: Transitorio dell'uscita del PGA in corrispondenza di un ingresso sinusoidale di  $250\mu V_{pp}$  e dei vari guadagni.

Per il buon interfacciamento dell'uscita della catena di amplificazione con l'ADC, descritto nel prossimo capitolo, è bene mantenere la tensione di uscita del PGA a 250mV e poterla eventualmente regolare. Come si può verificare dai risultati riportati nella tabella 3.3 e dalla figura 3.42, che mostra l'andamento del modo comune di uscita del PGA in una analisi DC (post-layout) con sweep della tensione di regolazione  $V_{REF}$ , anche questa esigenza è stata soddisfatta.

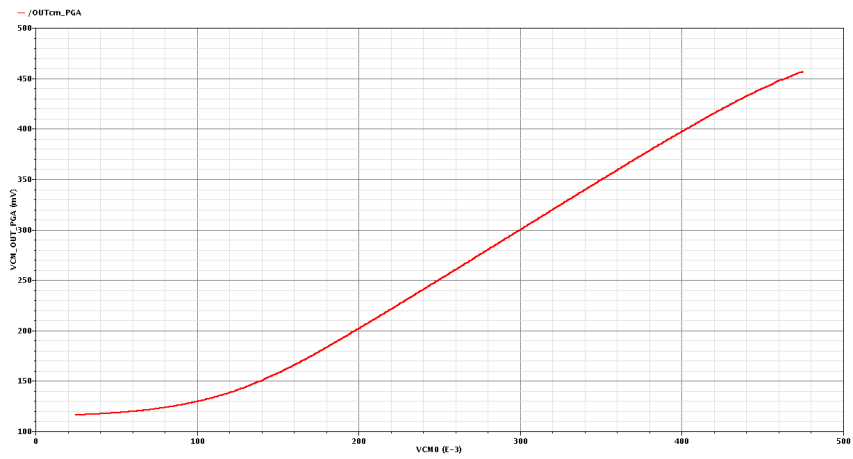


Figura 3.42: Regolazione della tensione di modo comune di uscita del PGA.

## Capitolo 4

# Valutazioni Preliminari al Progetto dell'ADC

### 4.1 Aspetti Generali

In questa sezione si riprendono brevemente le specifiche richieste all'ADC e le motivazioni sulle quali si fonda la scelta del tipo di convertitore: il SAR a redistribuzione di carica (SAR-CR). Si illustrano quindi la struttura e le caratteristiche che determinano le prestazioni in termini di linearità e consumo della principale classe di convertitori SAR-CR. La descrizione sarà accompagnata dall'illustrazione di alcuni risultati frutto di un lavoro di modellizzazione effettuato in MATLAB<sup>TM</sup> il cui scopo è duplice: da un lato permettere una migliore comprensione dei vantaggi, svantaggi e più in generale delle problematiche progettuali legate alle non idealità di questo genere di convertitori e dall'altro di presentare criteri di dimensionamento e fornire strumenti di analisi più fini e riutilizzabili anche per altre tipologie di SAR-CR cui è dedicata la sezione successiva.

#### 4.1.1 Specifiche e Panorama dei Convertitori ADC

Per essere integrato in ciascuno dei 64 canali che costituiscono il blocco di trattamento del segnale del sistema di monitoraggio neuronale, l'ADC deve essere progettato con una risoluzione nominale di 10bit e lavorare ad una frequenza di campionamento di 30kHz. I requisiti illustrati nel capitolo 3 imponevano al convertitore un consumo massimo di 500nW ed un numero effettivo di bit pari almeno a 8.2. Ciò corrisponde a una FoM inferiore a 60fJ/cstep. Dati i consumi estremamente ridotti previsti per gli stadi amplificanti progettati, gli obiettivi generali del progetto potrebbero essere raggiunti con un consumo di  $2.5\mu W$ , cioè con la realizzazione di un ADC la cui FoM è circa pari a 300fJ/cstep. Una delle principali intenzioni dell'attività di progetto è quella di investigare comunque il modo per ottenere la massima efficienza possibile e data la bassa velocità richiesta, si è ritenuto opportuno approfondire uno sforzo per la realizzazione di un ADC con una FoM ottimizzata.

Nel panorama dei convertitori analogico-digitali esiste una ampia varietà di topologie

secondo differenti principi di funzionamento; nel contesto dei sistemi a bassa velocità e bassissimi consumi solo alcune di queste topologie rappresentano valide opzioni progettuali. I convertitori a rampa fanno parte di questo insieme. La loro struttura è illustrata in figura 4.1 ed è costituita fondamentalmente da un comparatore, un contatore ed un generatore di segnale a rampa. Nonostante la semplicità, questa soluzione mostra alcuni aspetti critici per cui non è stata scelta per la realizzazione del convertitore. Per garantire buona risoluzione, infatti, è richiesto un generatore di rampa poco sensibile ai disturbi ed accurato anche al variare dei parametri di processo e di temperatura (ciò si traduce in maggiore consumo), ed inoltre necessita di un clock a frequenza sufficientemente alta per garantire la risoluzione desiderata [56].

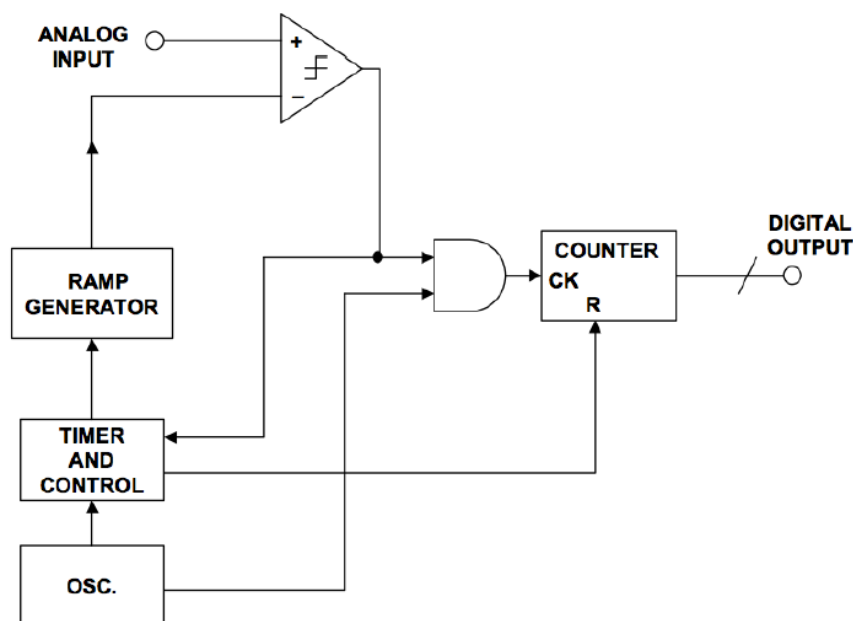


Figura 4.1: Single slope ADC [56].

Alla frequenza di campionamento prevista anche i convertitori sigma-delta rappresentano una opzione percorribile a seguito delle alte risoluzioni che sono in grado di garantire. I componenti analogici di questo tipo di circuiti sono relativamente semplici e sono rappresentati da comparatori, riferimenti di tensione, integratori e sommatore analogici, mentre i blocchi digitali di elaborazione del segnale richiesti, seppure scalabili con la tecnologia, sono caratterizzati da una elevata complessità. Grazie all'operazione di sovracampionamento del segnale e al noise shaping [32] che essi compiono, i  $\Sigma\Delta$  permettono di ottenere un ottimo rapporto segnale-rumore ed un elevato ENOB. Tali prestazioni potrebbero tuttavia risultare superflue in applicazioni come quella in questione, in cui un ENOB di 8-8.5 bit è già sufficiente a non degradare la qualità del segnale. Per questo motivo il bilancio tra i vantaggi e gli svantaggi di questi circuiti nell'ottica

di bassi consumi, risoluzioni medio basse ed area limitata e complessità progettuale, è sfavorevole. Il convertitore ad approssimazioni successive (fig. 4.3) è di fatto la migliore tra le soluzioni presentate per l'applicazione in questione (fig. 4.2).

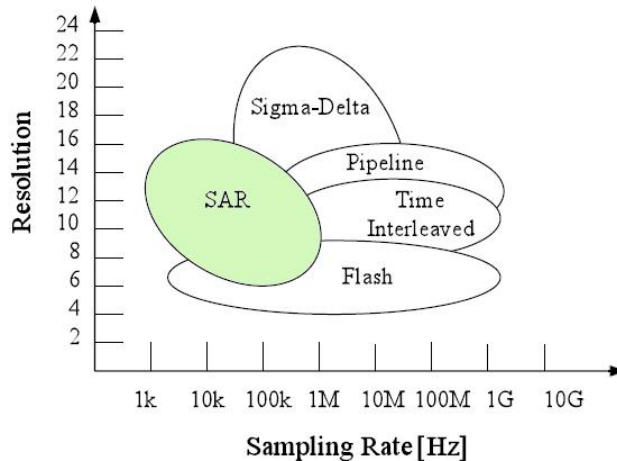


Figura 4.2: Utilizzo degli ADC in funzione di risoluzione e velocità [56].

La struttura è semplice e non richiede necessariamente circuiti analogici dai grandi consumi come amplificatori e riferimenti corrente. Questa classe di ADC sono noti per i bassi consumi [32] e la semplicità della struttura, costituita da un comparatore, un DAC ed una logica ad approssimazioni successive. La conversione è effettuata attraverso un processo ciclico. Subito dopo il termine della fase di campionamento, la tensione di segnale memorizzata è confrontata con la tensione di uscita del DAC (inizialmente equivalente al livello di metà dinamica) e a seconda dell'esito del confronto, il DAC viene aggiornato per generare una tensione maggiore o minore di quella già erogata. La procedura si ripete secondo un algoritmo di ricerca binaria in cui sono effettuate tante comparazioni quanti sono i bit di codifica fino ad arrivare a determinare il LSB. Al termine della codifica il DAC offre in uscita una tensione prossima al valore campionato, con un errore massimo di mezzo LSB. I risultati delle comparazioni sono memorizzati in opportuni registri e costituiscono la codifica binaria a  $n$  bit del valore campionato.

#### 4.1.2 Il SAR a Ridistribuzione di Carica

Esiste una particolare topologia di SAR, chiamata *charge redistribution* (CR, a ridistribuzione di carica), che dimostra ottima efficienza alle risoluzioni di interesse. Gli ADC con le migliori FoM presentate in letteratura, indicativamente comprese tra i  $4fJ/cstep$  e i  $50fJ/cstep$ , sono infatti ottenute con convertitori SAR CR la cui risoluzione varia tra i 6 e i 12 bit. La particolarità di questa classe di SAR è rappresentata dall'architettura del DAC: esso è costituito da una rete di condensatori ed interruttori pilotati da un opportuno circuito digitale, attraverso la quale è possibile riprodurre specifici valori di



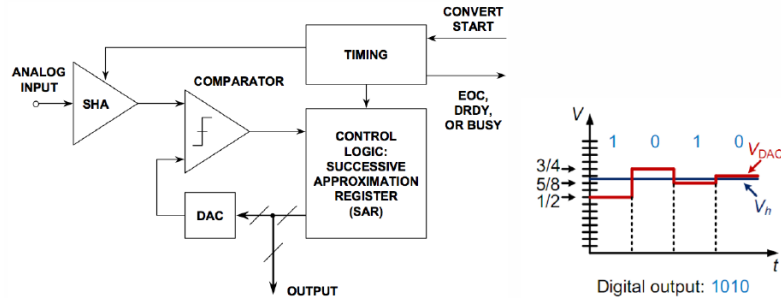


Figura 4.3: SAR ADC [56].

tensione al nodo connesso all'ingresso del comparatore. Il più importante e diffuso tipo di SAR CR utilizza un *array binario* (BW, Binary Weighted Array) come quello illustrato in figura 4.4, costituito dal parallelo di capacità dimensionate secondo uno schema binario [32] [65].

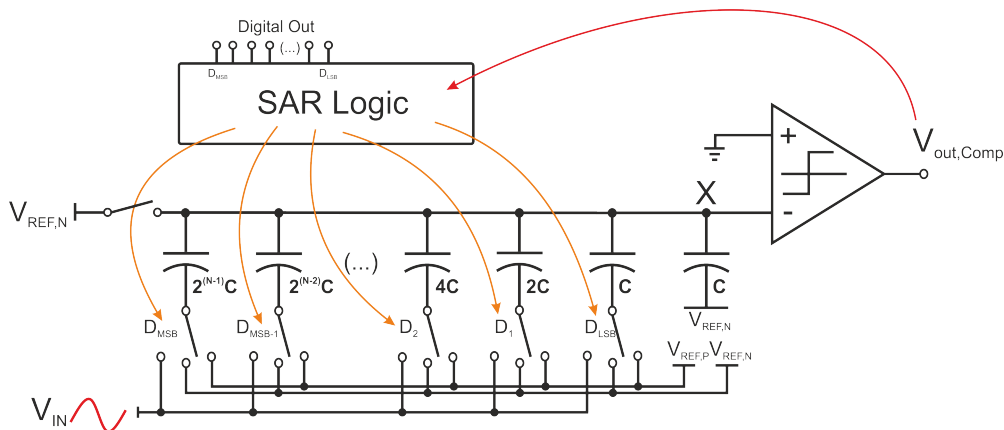


Figura 4.4: Binary Weighted Array Charge Redistribution SAR ADC.

In questa topologia il campionamento che precede la conversione è effettuato direttamente ai capi dell'array grazie alla preliminare chiusura e alla successiva apertura di due interruttori: uno che connette tutti i bottom plate dell'array al segnale di ingresso  $V_{IN}$  e un altro interruttore che connette al riferimento negativo il piatto superiore di tutte le unità (nodo X). Memorizzato il valore di tensione del segnale, la logica di comando, genericamente definita *Logica SAR*, è responsabile della corretta attuazione dell'algoritmo di ricerca binaria attraverso la sequenziale e selettiva attivazione degli interruttori che connettono il bottom plate della capacità dell'array alle due tensioni di riferimento superiore  $V_{REFP}$  ed inferiore  $V_{REFN}$ . Ad ogni configurazione di capacità commutate corrisponde un preciso valore di tensione al top plate (nodo X). La particolare organizzazione delle

capacità come blocchi costituiti da un multiplo binario di una certa capacità unitaria  $C_u$ , associata ad una adeguata configurazione della logica, garantiscono la possibilità di implementazione dell'algoritmo di ricerca.

Con riferimento allo schema di fig. 4.4, alla prima comparazione il bottom plate della capacità maggiore è portato ad alimentazione sintetizzando al nodo X una tensione corrispondente al valore dell'MSB. Il comparatore a questo punto viene attivato e la sua uscita è determinata dal segno dell'espressione:

$$-(-V_{IN} + \frac{2^{N-1}C_u}{2^N C_u}(V_{REFP} - V_{REFN})) = V_{IN} - \frac{\Delta V_{REF}}{2} \quad (4.1)$$

Se l'esito è positivo (uscita del comparatore alta), la comparazione del bit inferiore avviene dopo la semplice connessione della capacità successiva al riferimento superiore  $V_{REFP}$ . Se invece l'esito è negativo, prima di connettere a  $V_{REFP}$  la capacità successiva, il piatto inferiore (bottom) della capacità appena commutata è riportato al riferimento negativo. La procedura è la stessa per ogni comparazione fino al MSB e i valori di uscita del comparatore sono memorizzati in opportuni registri e costituiscono la codifica finale. I ritmi di campionamento e di comparazione sono in genere scanditi da un clock. Al di là del caso particolare appena descritto, le caratteristiche di ogni componente del convertitore e il valore delle tensioni di riferimento variano a seconda dell'applicazione, delle scelte progettuali e dei numerosi sottotipi di SAR capacitivi esistenti. Ciò che accomuna queste topologie è l'efficiente meccanismo di redistribuzione di carica tra le capacità utilizzato per riprodurre le tensioni di confronto.

Sebbene si sia finora esaminato solamente il caso di ADC single ended, la struttura del SAR progettato per il sistema sarà di tipo fully differential. In realtà ciò non rappresenta un problema in quanto le differenze tra i due tipi non risiedono nel principio di funzionamento ma solo in alcuni particolari di implementazione. Le strutture single ended, specialmente nel caso dell'array di capacità, sono più semplici da trattare e intuitive da comprendere. Per questo motivo in occasione della spiegazione di alcune importanti caratteristiche delle architetture analizzate si farà quasi sempre riferimento al caso single ended, chiarendo ove necessario l'impatto che può avere sulle prestazioni il passaggio ad una topologia fully differential.

La linearità ed il consumo del convertitore sono elementi da considerare con grande attenzione se si desidera ottenere FoM ridotte. Il consumo di un generico ADC SAR-CR binario realizzato con tipiche capacità di libreria offerte dalla maggior parte delle tecnologie submicrometriche (superiori a 20fF) si suddivide in percentuali simili tra array, logica e comparatore.

Tuttavia, curando la topologia dell'array e adottando algoritmi di commutazione opportuni è possibile ridurre molto i consumi rendendo il consumo dell'array minoritario rispetto a quello di una logica ed comparatore tipici. Anche in questi casi, comunque, l'array e la sua gestione determinano la qualità del segnale convertito e ciò, come vedremo, impone dei limiti all'adozione di topologie efficienti. Per questo motivo oltre che per motivi di sintesi, in questo capitolo sarà dedicato ampio spazio all'illustrazione delle

topologie di array finora più usate tra i lavori presenti in letteratura piuttosto che alle possibilità di realizzazione del comparatore e della logica, che possono essere resi più efficienti se realizzati rispettivamente in modo dinamico ed asincrono.

Nel prossimo paragrafo saranno analizzati in dettaglio i meccanismi alla base del consumo energetico associato alla carica e scarica delle capacità dell'array di un CR-SAR binario, evidenziando i fattori più importanti e confrontando tra loro differenti algoritmi di switching. Successivamente all'analisi energetica si affronta l'analisi delle caratteristiche di linearità di questa fondamentale topologia. Alcune considerazioni sono inoltre dedicate ad aspetti finora poco indagati che legano indici di qualità come parassiti e DNL, controllabili dal progettista, ad altri parametri come ad esempio l'ENOB, che stabiliscono nei fatti le prestazioni del chip ottenuto e che sono frutto di un procedimento di misura post-integrazione in cui il designer è soprattutto spettatore.

## 4.2 Topologie Alternative

In questa sezione troverà spazio una rapida analisi di alcune topologie di SAR a redistribuzione di carica adatte alla applicazione e all'implementazione in configurazione completamente differenziale e particolarmente interessanti per il principio di funzionamento e le prestazioni potenzialmente ottenibili. Dopo averne illustrato gli schemi di funzionamento saranno approfonditi i loro principali vantaggi e svantaggi in relazione all'applicazione per la quale sono candidate.

### 4.2.1 Generalità

Nel panorama dei convertitori analogico-digitali SAR a redistribuzione di carica la varietà di topologie di DAC, intese come combinazioni di struttura fisica dell'array capacitivo e algoritmo di commutazione sono numerose e una loro descrizione esaustiva richiede troppo spazio per essere affrontata in questo contesto. Per valutare l'opportunità di adottare una topologia piuttosto che altre si è scelto di esaminare unicamente le opzioni compatibili con una configurazione completamente differenziale e che per rilevanza (utilizzo in letteratura), o prestazioni potessero essere ritenute le migliori candidate. Topologie come la charge sharing, presentata per la prima ed unica volta nel 2010 da Chandrakasan et Al. [59] o la junction splitting [60], costituiscono esempi di convertitori CR SAR single ended innovativi ma solitari. Essi sono caratterizzati da procedure di commutazione molto più complesse di quelle dei SAR tradizionali e presentano problematiche come la marcata sensibilità al mismatch e ai parassiti (nel charge sharing) o l'utilizzo di capacità flottanti (nel junction-splitting), non accompagnate da linee guida progettuali sufficientemente consolidate. Per questo motivo si è preferito evitare lo sforzo nella direzione di un loro adattamento all'utilizzo differenziale richiesto dall'applicazione per esaminare con più attenzione altre topologie interessanti e che offrono comunque un certo spazio per introdurre elementi di novità. Sono due le principali topologie di DAC utilizzate nel maggior numero di lavori dedicati ai convertitori SAR CR: la topologia binaria e la topologia con capacità di attenuazione, detta anche *Bridge Capacitor Array Topology*. Ciascuna di esse può essere utilizzata con una opportuna procedura

di commutazione che può implicare modifiche secondarie alla struttura fisica dell'array e eventualmente migliorarne le prestazioni. Con poche eccezioni, adeguatamente sottolineate nel corso del capitolo, queste due strutture sono utilizzabili sia in configurazione single ended che fully differential. Le due strutture alternative sono semplicemente schematizzate in figura 4.5.

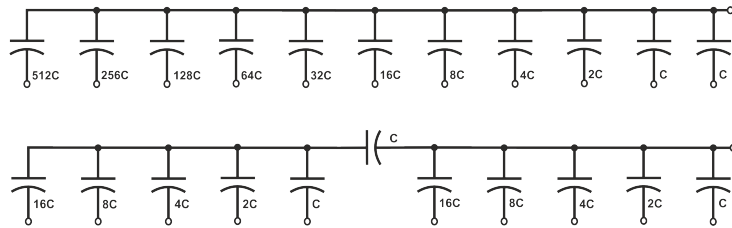


Figura 4.5: Binary Weighted Array (Binario), in alto. Bridge Capacitor Array, in basso. Esempi per DAC a 10bit

In questa sezione sono analizzate e confrontate le caratteristiche di consumo e di linearità delle differenti topologie ad array binario e bridge. Sia l'analisi di consumo che quella di linearità sono state svolte con l'ausilio di modelli *MATLab<sup>TM</sup>* opportunamente realizzati per ciascuna topologia.

## 4.2.2 Consumo delle Topologie ad Array Binario

Il consumo riconducibile all'attività di commutazione dell'array binario di CR-SAR ADC è associato alla carica che ad ogni confronto viene trasferita, persa oppure richiesta all'alimentazione per generare l'opportuna tensione di confronto al nodo di ingresso del comparatore. Essa dipende dalla tensione da convertire perchè ad ogni codifica si associa una sequenza di switching più o meno dispendiosa in termini di carica. Nondimeno, i differenti algoritmi di attivazione degli interruttori secondo i quali è possibile attuare la procedura di ricerca binaria così come alcune varianti dell'array binario sono in grado di influenzare il consumo. In questo paragrafo è presentata una breve analisi del consumo energetico relativo al solo array di capacità e al modo in cui è utilizzato per la conversione. L'analisi è effettuata per alcune importanti topologie di SAR ADC ad array binario a pari capacità unitaria e risoluzione. Per semplicità e dato che si tratta di un caso frequente in letteratura, si confronteranno architetture in cui i riferimenti di tensione superiore ed inferiore dell'array sono rappresentati da alimentazione  $V_{DD}$  o  $V_{REF}$  e massa.

L'energia consumata dall'array non è unicamente dovuta alla procedura di commutazione. Anche in fase di campionamento è necessaria carica per caricare le capacità e quindi una certa corrente richiesta all'alimentazione. Tuttavia questa dipende in prima approssimazione dalla capacità d'ingresso dell'array e non è particolarmente influenzata dall'algoritmo di commutazione.

Una dettagliata analisi del dispendio energetico associato ad alcuni algoritmi è presentato in [57] con riferimento ad una struttura elementare a 3 bit riferita tra alimentazione e

massa. Nello stesso lavoro sono confrontate l'efficienza della procedura classica con quella di una diversa procedura proposta ed associata ad un array binario in cui la capacità dell'MSB è organizzata in blocchi speculari, come in figura 4.6 e definita *split capacitor topology*.

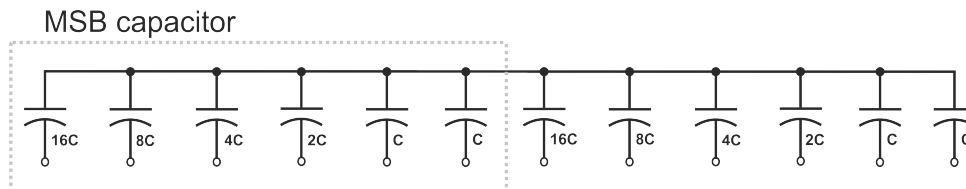


Figura 4.6: Split Capacitor Binary Weighted Array.

In questa particolare procedura, come in quella classica, al primo confronto si porta al riferimento positivo (alimentazione) la capacità del MSB. Ma mentre nella procedura tradizionale in caso di codifica negativa del MSB è necessario riportare l'intera capacità a massa per poi testare il bit successivo, nel caso dello splitting, si può semplicemente riportare a massa solo la subunità della capacità del MSB che corrisponde al bit che si sta confrontando. In questo modo è possibile ridurre il numero di volte in cui si commuta una capacità ad alimentazione. L'efficienza è calcolata rispetto all'energia media di commutazione tra tutte le codifiche ed è normalizzata rispetto all'energia  $E_{unit}$  associata alla carica e scarica di una capacità unitaria:

$$E_{unit} = C_u V_{REF}^2 = C_u V_{DD}^2 \quad (4.2)$$

Si dimostra che essa è migliore del 37% [57]. In realtà, sebbene sia difficile da quantificare, è lecito aspettarsi anche un certo incremento del consumo della logica SAR, più complessa rispetto a quella delle topologie tradizionali.

Per verificare la validità di questa soluzione e poterla comparare con altre topologie di ADC a redistribuzione di carica sono stati realizzati dei modelli *MATLab<sup>TM</sup>* per un numero di bit pari a 10, come quello dell'ADC da realizzare, sia nelle loro configurazioni single ended che in quelle fully differential. Con riferimento al caso single ended, i risultati delle simulazioni effettuate sono coerenti con i dati della letteratura. In figura 4.7 sono illustrati i consumi delle topologie ad array binario esaminate.

Mentre la procedura classica mostra un consumo energetico che peggiora in corrispondenza della codifica minima (ogni capacità va portata a alimentazione e poi riportata a massa) e offre un consumo medio di  $1365.3 E_{unit}$  (come confermato in [58]), i risultati delle simulazioni per il modello di split capacitor array ADC realizzato confermano un consumo medio inferiore di quasi il 40% rispetto al precedente, pari a circa  $852 E_{unit}$ . La curva inferiore si riferisce ad una terza possibile procedura di commutazione, particolarmente interessante e presentata in [58] per un ADC completamente differenziale ed il cui schema è evidenziato in figura 4.8.

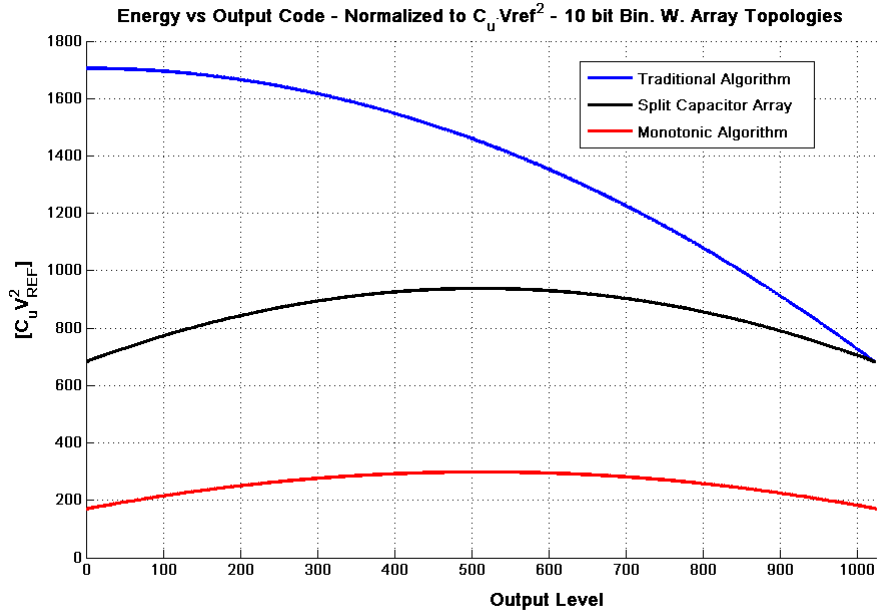


Figura 4.7: Consumo energetico per codifica normalizzato rispetto a  $E_{unit}$  delle topologie ad array binario esaminate.

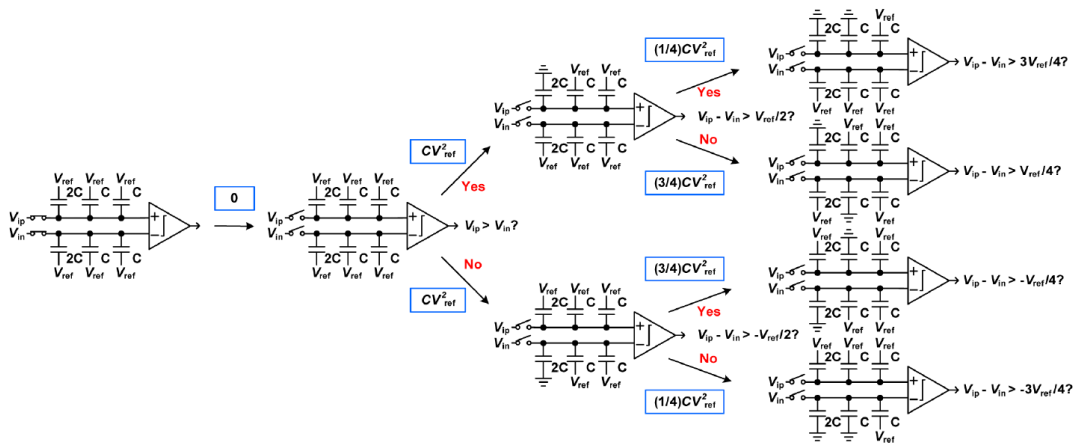


Figura 4.8: Schematizzazione della procedura di commutazione monotona per un ADC SAR CR fully differential binario a 3 bit [58].

Questa architettura è possibile solo in modalità completamente differenziale e prevede il campionamento del segnale attraverso interruttori afferenti al top plate dell'array e con i piatti inferiori di tutte le capacità connesse ad alimentazione. In questo modo, ai due morsetti di ingresso è imposta la corrispondente tensione di segnale e una volta aperti gli interruttori di campionamento, senza ulteriori commutazioni, è immediatamente possibile valutare il MSB. Da notare che la possibilità di effettuare la valutazione del MSB subito dopo il campionamento permette l'eliminazione della capacità maggiore dell'array.

A seconda dell'esito della comparazione, è semplicemente commutata a massa la corrispondente capacità di uno dei due array. Nessuna capacità è mai commutata verso l'alto durante il ciclo. In questo modo è possibile ottenere un consumo energetico pari a circa  $255.5E_{unit}$  [58], confermato dalle simulazioni e pari al 20% di quello ottenibile con l'algoritmo tradizionale.

Il SAR ADC a redistribuzione di carica più efficiente mai presentato in letteratura utilizza un array binario ed è stato realizzato nel 2010 dal team dell'Università di Twente guidato dal Prof. Nauta [61]. Lo schema di conversione era quello di una topologia binaria classica ma le capacità unitarie full-custom di 0.5fF ciascuna permettevano un consumo estremamente ridotto. Inoltre, anche se per stessa ammissione degli autori ciò non è risultato un fattore determinante, la carica delle capacità corrispondenti al MSB avveniva secondo una tecnica *stepwise* adiabatica [64]. In questo modo il consumo relativo alla carica ed alla scarica delle capacità era ridotto di un fattore pari al numero  $K$  di passi effettuati per la carica (e scarica). L'inconveniente associato alla tecnica è rappresentato dalla necessità di utilizzare  $K$  capacità in più per la memorizzazione delle tensioni intermedie, ciascuna pari o superiore a quella del MSB. La FoM ottenuta dal convertitore è di  $4.4fJ/cstep$ , grazie anche all'utilizzo di un comparatore dinamico ad alta efficienza. Sebbene rappresenti un punto di riferimento per il progetto di convertitori, questo lavoro non presenta una struttura di array innovativa ed utilizza capacità appositamente progettate.

In definitiva, alla luce di quanto visto finora, la topologia più interessante dal punto di vista dell'efficienza è quella che utilizza un algoritmo monotono, ma questa non è esente da potenziali problemi. La particolarità di questa procedura realizzata in convertitori fully differential è la diminuzione della tensione di modo comune  $V_{CM}$  ai nodi di ingresso del comparatore che avviene ad ogni passo, pari ogni volta alla metà dell'equivalente tensione analogica della caratteristica corrispondente al bit codificato. Tale fenomeno, mostrato in figura 4.9 non si ha nel caso di una topologia fully differential con procedura tradizionale (in cui  $V_{CM}$  rimane costante) ed è dovuto al fatto che ad ogni passo nessuna capacità è mai commutata verso l'alto ma solo verso il basso, con la conseguente diminuzione della tensione al top plate di ciascun array. L'algoritmo monotono quindi impone uno specifico progetto del comparatore, che deve poter lavorare anche quando il modo comune dei suoi ingressi è vicino al rail inferiore. Più avanti verrà approfondita questa problematica. Il prossimo paragrafo è invece dedicato all'analisi delle prestazioni di linearità associate alle architetture binarie. Nella prossima sezione, si passerà ad esaminare topologie più avanzate e interessanti ai fini progettuali, riproponendo lo stesso ordine e approccio analitico utilizzato per le topologie appena esaminate.

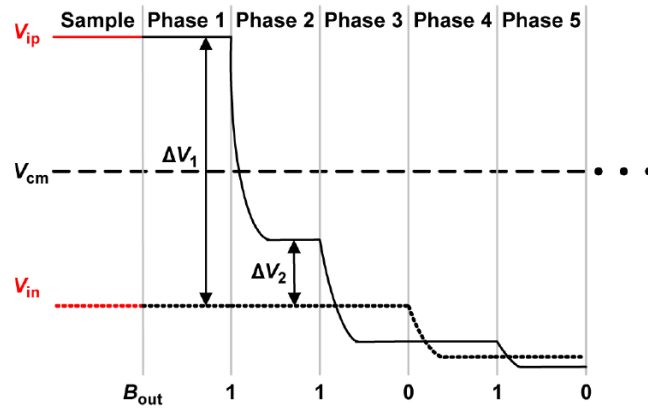


Figura 4.9: Riduzione della tensione di modo comune all'ingresso del comparatore associato alla procedura di switching monotonica a 5 bit [58].

### 4.2.3 Linearità delle Topologie ad Array Binario

Insieme al consumo, l'elemento che concorre maggiormente a determinare la figura di merito di un convertitore è la linearità. A differenza del consumo, cui l'array contribuisce in una percentuale simile a quelle di comparatore e logica, la linearità dipende per la maggior parte dalla qualità di implementazione dell'array ed entra nel calcolo della figura di merito attraverso il numero effettivo di bit, cioè l'ù. Quest'ultimo è un parametro che in genere è possibile valutare durante il processo di misura oppure anche in simulazione ma comunque successivamente alla realizzazione dell'array ed è legato ad altri parametri descrittivi della linearità dell'ADC, cioè DNL ed INL, già introdotti nel capitolo dedicato alle specifiche. Questi sono parametri controllabili dal progettista perchè legati da relazioni deterministiche ai parassiti oppure dovuti a fenomeni stocastici con caratteristiche note, come il mismatch della tecnologia. Tuttavia il legame tra DNL, INL ed ENOB non è quantitativamente immediato nonostante sia importante per il progettista. Sollecitati da questa esigenza, nel prossimo paragrafo illustriamo un approccio basato su simulazioni con il quale si è cercato di affrontare il problema e più avanti, quando si tratterà di illustrare le caratteristiche e le scelte di progetto per il convertitore realizzato, questo tema verrà ripreso. Come primo passo, in questo paragrafo sono invece illustrati gli elementi che contribuiscono a determinare DNL e INL in un CR SAR ADC binario. Il mismatch tecnologico è l'unica sorgente di non idealità che introduce in modo significativo non linearità nel convertitore. Ogni capacità dell'array binario nel layout è di fatto assemblata mettendo in parallelo il numero di capacità unitarie necessarie a sintetizzarne il valore, pari ad uno per la capacità del LSB oppure pari ad un intero potenza di due per gli altri bit. Questa soluzione è adottata tipicamente per poter disporre le capacità dell'array secondo distribuzioni meno sensibili ai gradienti di ossido e in generale alle non idealità di processo che sono all'origine dei mismatch. Nonostante questi accorgimenti il mismatch non può essere del tutto eliminato. Esso determina una



modifica percentuale del valore di ciascuna capacità unitaria, che introduce una variazione assoluta  $\Delta C$  modellizzabile come un processo stocastico gaussiano a media nulla e deviazione standard ricavabile dalle relazioni di Pelgrom e pari a:

$$\sigma_C = C_u \cdot \frac{k_c}{\sqrt{Area}} = k_c \sqrt{C_u c_{spec}} \quad (4.3)$$

dove  $c_{spec}$  è la capacità per unità di area associata alle capacità di tecnologia utilizzate,  $C_u$  è il valore di capacità unitaria scelto e  $k_c$  è parametro di mismatch tecnologico indicato in  $\% \cdot \mu m$ . Nella tecnologia UMC  $0.13 \mu m$  utilizzata si ha  $k_c$  pari a 0.0095 mentre il massimo valore di  $c_{spec}$  disponibile è quello delle capacità MIM (Metal-Insulator-Metal) ed è pari a  $1.11 fF / \mu m^2$ . In un SAR CR binario ad ogni configurazione di capacità connesse a massa oppure a alimentazione corrisponde un livello di tensione al nodo di ingresso del comparatore con il quale si confronta il segnale campionato. Questi valori discriminano tra due livelli successivi di codifica di uscita e rappresentano di fatto le ascisse della caratteristica tra ingresso  $V_{IN}$  e uscita  $D_{out}$  dell'ADC in corrispondenza del quale si ha il cambio di livello. Questa relazione è rappresentata in figura 4.10, nella quale è anche evidenziato come scostamenti delle tensioni sintetizzate rispetto ai loro valori ideali siano causa di DNL.

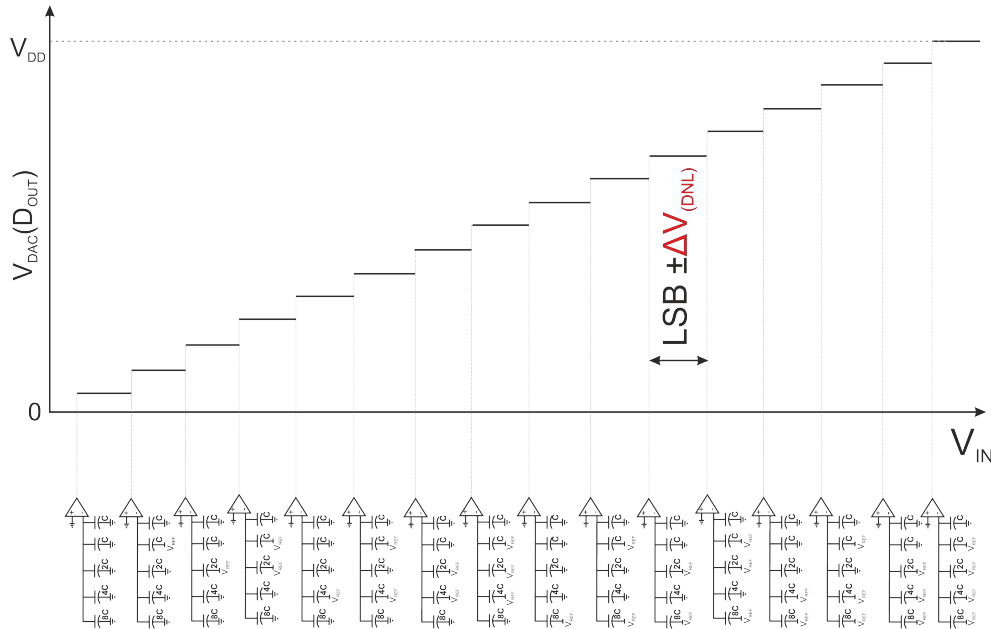


Figura 4.10: Legame tra configurazioni capacitive e caratteristica dell'ADC CR SAR binario.

Ad una variazione statistica della capacità unitaria corrisponde una variazione statistica di ogni blocco capacitivo dell'array on conseguente alterazione dei livelli di tensione sintetizzabili durante il ciclo di conversione e la generazione di una non linearità differenziale. Uno scostamento dei livelli di tensione effettivi superiore al LSB dai corrispettivi ideali comporterebbe la perdita netta di un livello di conversione e la non monotonicità della caratteristica. Per questo motivo è in genere richiesto di realizzare convertitori con una caratteristica di DNL in funzione della codifica che non presenti mai picchi con valore assoluto superiore a 0.5LSB. Il legame tra la DNL massima ed il mismatch in una topologia ad array binario di N bit con procedura di switching tradizionale è stato analizzato da Chandrakasan et al. [73]. Nel lavoro in questione è sottolineato come in genere il massimo contributo alla DNL sia dovuto all'errore della capacità equivalente dell'MSB, che essendo la più grande presenta anche una variazione in valore assoluto superiore a quella di tutte le altre capacità, determinando uno scostamento del livello di tensione dell'MSB che a pari percentuale risulta il maggiore. Tale errore entra in gioco durante la fase di valutazione del LSB, quando il bottom della relativa capacità è connesso prima al riferimento superiore, e poi eventualmente è riconnesso al riferimento inferiore. L'entità di questo errore, relativo alla transizione tra l'MSB e la codifica appena inferiore, è ricavata in [73] e risulta essere circa pari a:

$$\Delta V_{err,MSB} \approx \frac{\Delta C_{total}}{2^{\frac{N}{2}} C_u} V_{REF} \quad (4.4)$$

pertanto la deviazione standard della DNL dovuta all'MSB si ottiene calcolando la deviazione standard del rapporto tra la tensione errore tra i due livelli adiacenti e l'LSB, cioè:

$$\sigma_{DNL_{max}} \approx \sqrt{Var\left[\frac{\Delta C_{total} 2^{\frac{N}{2}}}{C_u}\right]} \quad (4.5)$$

dove la variazione  $\Delta C_{total}$  è la somma degli errori sulle capacità unitarie portate al riferimento positivo nelle configurazioni dei due livelli adiacenti. Si noti che tali errori sono identicamente distribuiti e indipendenti tra loro per cui non conta quali capacità ad ogni codifica hanno il bottom connesso al riferimento positivo ma solo il numero totale di queste. La deviazione standard della DNL può essere approssimata con quella corrispondente al livello di transizione appena precedente all'MSB. Questo in genere rappresenta il caso peggiore perchè somma il contributo al mismatch del maggior numero di capacità unitarie; si ha quindi:

$$\sigma_{DNL_{max}} \approx \sigma_{DNL_{MSB}} \approx 2^{\frac{N}{2}} \frac{\sigma_C}{C_u} \quad (4.6)$$

ed in generale, come ogni mismatch capacitivo, è un processo gaussiano e con media nulla. Lo stesso calcolo è proposto sempre in [73] per la topologia con split capacitor array ed evidenzia una situazione migliore di un fattore  $\sqrt{2}$ :

$$\sigma_{DNL_{max}} \approx 2^{\frac{N}{2}} \frac{\sigma_C}{\sqrt{2} C_u} \quad (4.7)$$

Nella topologia con split capacitor infatti la capacità dell'MSB (con relativo mismatch) è connessa a  $V_{REF}$  all'inizio, come nella procedura classica, ma con una differenza: in caso di esito negativo della comparazione (bit 0), nella fase appena successiva è ricollegata a massa solo una parte di essa, quella corrispondente alla metà della capacità totale. Ciò comporta un fattore due di riduzione della varianza dell'errore di tensione prodotta all'ingresso del comparatore. Una buona regola di dimensionamento della capacità minima di un array binario classico per evitare la non monotonicità della caratteristica dovrebbe fare in modo di generare una DNL la cui probabilità di essere 0.5 sia remota, ad esempio imponendo che tale valore cada 3 deviazioni standard distante dalla media:

$$3 \cdot \sigma_{DNL_{max}} \leq 0.5 \quad (4.8)$$

Questa relazione si tradurrebbe per un array binario classico in:

$$C_u \geq \frac{9 \cdot 2^N k_c^2 C_{spec}}{(0.5)^2} \quad (4.9)$$

Per dare l'idea dei valori in gioco, con i parametri della tecnologia disponibile dovremmo avere capacità minime di almeno 3.7fF. Nell'ottica di un progetto con componenti standard tale valore non rappresenta un problema dato che è ben al di sotto della capacità minima di libreria fornita nel designkit, ma indica che con le capacità disponibili l'ottimizzazione non è possibile.

Se il binario, dati i parametri di tecnologia in gioco, si rivela una topologia piuttosto robusta e poco sensibile ai mismatch, altre strutture che vedremo più avanti non si rivelano altrettanto robuste ma possono rappresentare una scelta appetibile per altri vantaggi. In previsione di affrontare l'analisi di altre topologie per le quali stime analitiche delle non linearità risultano più complesse e con minori riferimenti in letteratura, è stato deciso di elaborare i modelli MATLAB<sup>TM</sup> dei vari convertitori, tra cui anche il binario e lo split. Tali modelli sono stati realizzati pensando di riprodurre l'esatta implementazione fisica dell'array con le sue capacità affette da mismatch e il preciso algoritmo di switching. Il risultato di questo lavoro ha permesso di avere una stima più realistica dell'impatto effettivo che il mismatch ha sulla caratteristica ed in particolare sulle grandezze. Lo scopo delle simulazioni è stato quello di estrarre per ogni topologia realizzata con dati parametri tecnologici e capacità unitaria, alcune grandezze di interesse relative alla linearità. In particolare, allo scopo di verificare la coerenza dei modelli con le formule 4.6 e 4.7 è stata valutata la distribuzione del valore assoluto della DNL massima tra tutte le codifiche ed in particolare della codifica del MSB. In figura 4.11 è illustrato l'andamento della deviazione standard della DNL di ogni codifica per un CR SAR binario classico a 10 bit e capacità unitaria di 20fF, valutato su 2000 realizzazioni. Come si osserva dalla figura, il valore della deviazione standard in corrispondenza della codifica di metà dinamica (MSB) è pari a 0.0712, cioè precisamente quanto si ottiene analiticamente dalla 4.6 sostituendo i parametri di tecnologia e la capacità unitaria.

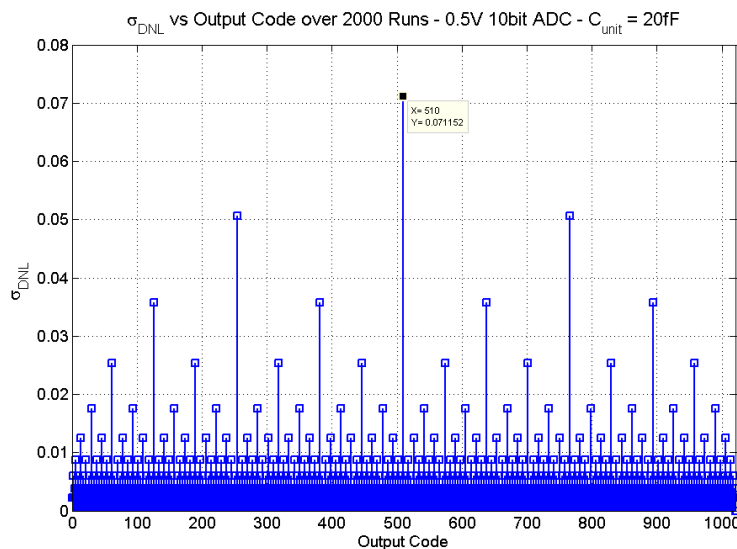


Figura 4.11: Distribuzione della deviazione standard della DNL di un convertitore ad array binario classico, per tutte le codifiche.

Nel caso della topologia a split capacitor a pari capacità unitaria e numero di bit, la deviazione standard della DNL in funzione della codifica è mostrata in figura 4.12. Anch'essa mostra un picco in corrispondenza della codifica di transizione del MSB ed esso è pari a , cioè esattamente di un fattore  $\sqrt{2}$  inferiore a quello ottenuto con la procedura classica, come previsto dalle stime teoriche.

E' da notare come i valori di deviazione standard su tutte le codifiche si configurino secondo una precisa distribuzione. Questa dipende sia dall'architettura dell'array che dall'algoritmo di commutazione utilizzato. A questo punto è opportuna una precisazione. Al contrario di applicazioni nel campo della spettroscopia in cui occorrono convertitori che codifichino perfettamente ogni livello della caratteristica, l'elaborazione e la digitalizzazione di segnali analogici a tempo continuo come quelli neuronali richiede semplicemente la conservazione della qualità del segnale a fronte di distorsioni e rumore. In questo caso l'ENOB, piuttosto che la DNL, rappresenta il parametro di riferimento. Assumendo ragionevolmente che la linearità (quindi l'ENOB) del convertitore dipenda sia dalla particolare distribuzione delle deviazioni standard della DNL che dal loro valore, si può comprendere come il dimensionamento della capacità unitaria del convertitore sulla base di relazioni come la 4.6 possa risultare troppo rigido. Queste relazioni sono infatti riferite al caso peggiore, che è caratteristico di una singola codifica, mentre tutte le altre sono caratterizzate da indici di dispersione della DNL inferiori o molto inferiori.

I modelli implementati, al contrario, sono in grado di fornire cifre più realistiche che tengono conto del contributo di tutte le codifiche. Questi modelli sono stati utilizzati con due finalità. La prima finalità è poter verificare per dati parametri di mismatch tecnologico e parassiti e con un definito dimensionamento, come si distribuisce (su mol-

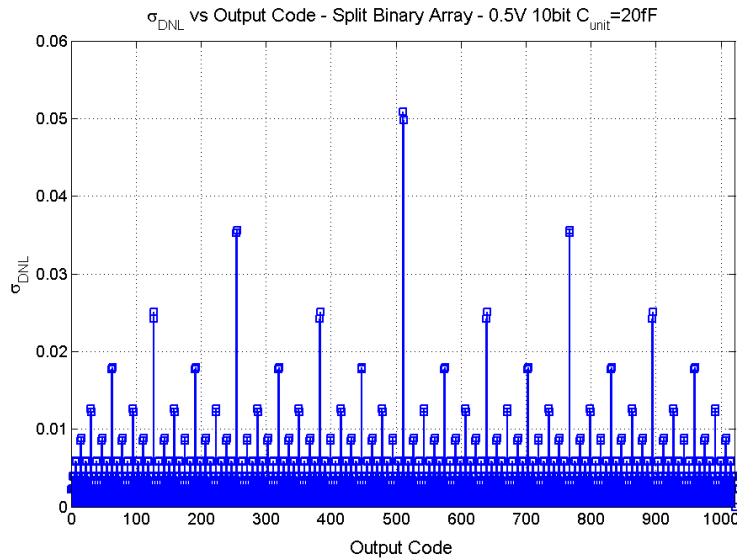


Figura 4.12: Distribuzione della deviazione standard della DNL di un convertitore ad array binario con split capacitor, per tutte le codifiche.

teplici realizzazioni) il massimo valore assoluto di DNL. Il secondo obiettivo è quello di valutare l'impatto sull'ENOB delle DNL tipiche di ogni topologia ed al variare della capacità unitaria. Quest'ultima questione sarà ripresa alla fine del capitolo per effettuare un confronto tra le varie topologie.

La schematizzazione del tipo di simulazioni effettuate per il calcolo del massimo modulo di DNL è mostrato in figura 4.13 e un esempio dei risultati da esse fornito è in figura 4.14. Si tratta di istogrammi della distribuzione del valore assoluto massimo della DNL di ogni caratteristica, come quello illustrato in figura 4.14.

Come si può osservare, la distribuzione degli istogrammi non è una gaussiana a media nulla. Infatti, sebbene la DNL rappresenti un processo gaussiano a media nulla il cui contributo principale è riconducibile al mismatch della capacità del MSB, il valore assoluto massimo della DNL è il risultato di operazioni matematiche e logiche evidenziate in figura 4.15 che ne modificano le caratteristiche di distribuzione.

Da una distribuzione gaussiana con media nulla e deviazione standard simile alla 4.6, attraverso l'operazione di valore assoluto si passa a una distribuzione di tipo *gaussiano ripiegato* e successivamente all'individuazione dei massimi di ogni caratteristica ad una distribuzione analoga a quella mostrata negli istogrammi. La trattazione matematica completa del problema delle distribuzioni di queste grandezze non è stata svolta per motivi di complessità e perchè non è stata ritenuta necessaria ai fini progettuali. Inoltre, simulazioni per confrontare le prestazioni dell'array binario rispetto a quello con split capacitor sono state effettuate al variare della capacità unitaria e per 1000 realizzazioni in corrispondenza di ogni valore. Questo è servito a verificare la coerenza dei modelli rispetto al confronto tra varie topologie. Il risultato di queste simulazioni è illustrato in

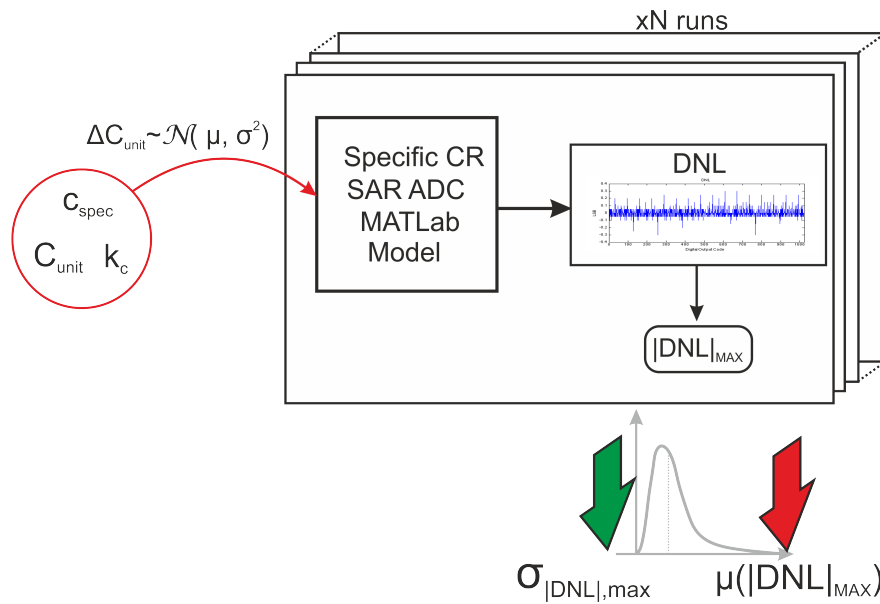


Figura 4.13: Simulazioni alternative effettuate per estrarre informazioni sulle prestazioni di linearità delle tipologie di array.

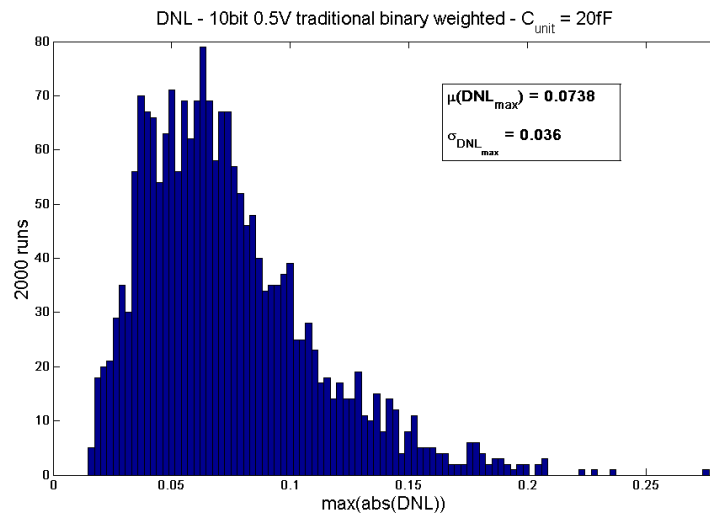


Figura 4.14: Istogrammi della distribuzione del massimo modulo di DNL in un binary weighted array con algoritmo di commutazione tradizionale.

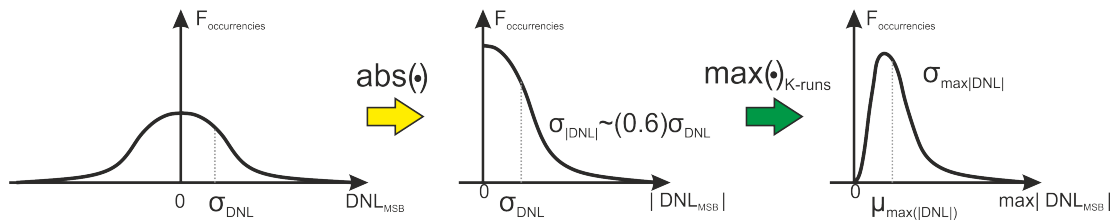


Figura 4.15: Cambiamento delle distribuzioni delle variabili di interesse a seconda delle elaborazioni effettuate in simulazione.

figura 4.16 per la media della DNL massima (valore assoluto) e mostra un rapporto pari a  $\sqrt{2}$  a favore dello split, come era ragionevole aspettarsi dalle stime teoriche.

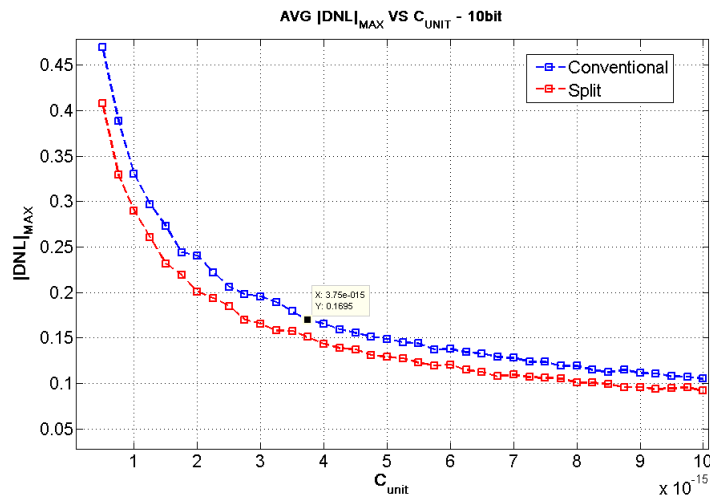
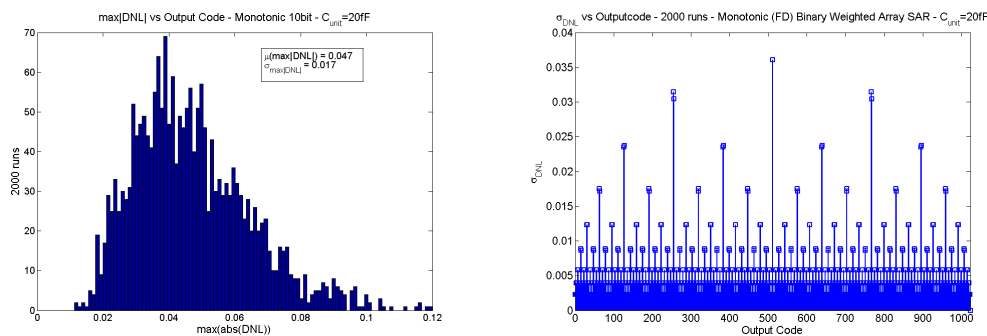


Figura 4.16: Confronto tra massimo modulo della DNL di ADC CR SAR binario classico e con split capacitor.

Dalla figura 4.16 risulta evidente come già con pochi fF di capacità unitaria le topologie binarie esaminate mostrano avere buone prestazioni in termini di DNL. Anche per la topologia binaria monotona (fully differential) sono state realizzate queste simulazioni e sono mostrate in figura 4.17. Si evidenzia una distribuzione delle deviazioni standard delle DNL per codifica analogica a quella del binario ma migliore di un fattore  $\sqrt{2}$ . A causa della configurazione fully differential infatti il LSB raddoppia ma il contributo dei mismatch alla varianza peggiora lievemente a causa della diminuzione di capacità totale dell'array.

Anche per la INL sono state svolte simulazioni analoghe ma si è posta su di essa una minore attenzione rispetto a quella rivolta alla DNL per una serie di motivi. Di fatto, una volta depurata dal contributo di puri errori di guadagno e di offset (che non sono delle non linearità), la DNL contiene di per sé tutte le informazioni relative alla non linearità



(a) Istogrammi del massimo modulo di DNL. (b) Deviazione standard della DNL in funzione della codifica.

Figura 4.17: Risultati delle simulazioni effettuate con i modelli di ADC binario monotono (2000 realizzazioni).

della caratteristica. In base alle caratteristiche della sua distribuzione si possono trarre conclusioni circa la qualità del convertitore e, come vedremo nel prossimo paragrafo, fondare ragionamenti sui possibili legami tra la caratteristica dell'ADC e l'ENOB. In realtà, come vedremo, la INL ha un ruolo nel determinare il degrado dell'ENOB ma è una grandezza più scomoda da trattare e per la quale attualmente non ci sono formule chiuse che ne descrivano i legami con la capacità unitaria e i parametri di mismatch della tecnologia.

Un altro elemento da considerare è rappresentato dalle capacità parassite. Esse dipendono da diversi fattori: la particolare geometria delle capacità, gli accoppiamenti tra le capacità e altri elementi fisici dell'integrato e le capacità equivalenti verso massa delle connessioni che afferiscono al top plate dell'array. Mentre le capacità parassite ai bottom delle capacità non rappresentano un problema di alcun tipo perchè a quei nodi è applicata direttamente una delle tensioni di riferimento oppure di ingresso, quelle poste tra top plate dell'array e i vari bottom delle capacità provocano effetti di tipo identico al mismatch ed è necessario controllarli con opportune scelte di layout. Infine, le capacità tra piatto superiore e massa (o riferimento fisso) hanno influenza sulla caratteristica ma non introducono non linearità, bensì un errore di guadagno. Si consideri l'array di figura 4.18 in cui è evidenziato il parassita al top plate.

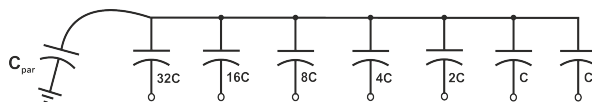


Figura 4.18: Parassita al top plate di un binary weighted array a 6 bit.

La tensione generata al nodo di top è funzione del numero  $m$  di capacità unitarie in tutto l'array che a seconda della codifica hanno il bottom connesso ad alimentazione ed è definita dalla relazione:



$$V_{comp} = \frac{mC_u}{C_{tot}} = \frac{mC_u}{2^N C_u + C_{par}} \quad (4.10)$$

Come si può notare il parassita ha il solo effetto di aumentare la capacità totale dell'array e quindi di avvicinare in uguale misura tutti i livelli di tensione che discriminano tra una codifica e la successiva. Da questo fenomeno deriva un errore di guadagno che non ha ripercussioni sulla linearità del convertitore.

#### 4.2.4 Bridge Capacitor Array

L'architettura cosiddetta *Bridge* deve la sua notorietà ad alcuni lavori del gruppo di ricerca del Prof. Maloberti dell'Università di Pavia [62] [63] e rappresenta un candidato molto interessante per il progetto del convertitore. Il nome dell'architettura deriva dalla sua struttura fisica, illustrata in figura 4.19, in cui due array binari identici sono collegati con una capacità di ponte (bridge) o anche *di attenuazione*.

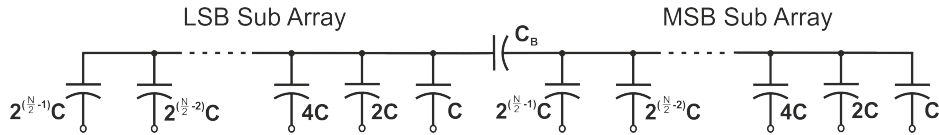


Figura 4.19: Topologia con Bridge Capacitor Array.

Il vantaggio più evidente di questa topologia è la drastica riduzione della capacità totale di un fattore  $2^{\frac{N}{2}}$ , dove  $N$  è il numero di bit del convertitore. L'array appare diviso in due metà identiche (esistono varianti asimmetriche ma faremo riferimento al caso più tradizionale, simmetrico), ciascuna costituita da un array binario equivalente a quello che si utilizzerebbe per un convertitore con la metà dei bit, ma senza una capacità unitaria. La capacità unitaria mancante di fatto è costituita dalla capacità vista dal top plate di ciascun *sub-array* verso la capacità di ponte  $C_B$ . Con una logica SAR di pilotaggio ed una procedura di switching praticamente identiche a quelle di un CR-SAR binario tradizionale si possono attivare gli interruttori di bottom delle capacità per riprodurre i livelli di tensione di transizione della caratteristica di conversione. Il sub-array direttamente connesso al comparatore è dedicato alla codifica della prima metà dei bit a partire dal MSB, l'altro codifica i restanti. La corretta riproduzione delle tensioni è resa possibile dalla stessa capacità di attenuazione che nel caso dei primi 5 bit offre al top-plate del primo sub-array una capacità equivalente all'unitaria (è in serie all'altro sub-array), mentre per i passi successivi riporta al top plate del MSB array (il primo) le tensioni prodotte al top plate del secondo sub-array, che chiameremo *LSB sub-array*, attraverso un rapporto di attenuazione  $AR$  dato dalla partizione capacitiva:

$$AR = \frac{C_B}{C_B + C_{TOT_{sub}}} \quad (4.11)$$

che idealmente dovrebbe essere pari a  $2^{-\frac{N}{2}}$ . In realtà, a differenza della topologia binaria classica, questo array è affetto da un intrinseco errore di guadagno associato alla presenza della capacità di bridge. Per fare in modo che la capacità equivalente vista dal nodo di top di ciascun sub-array verso il bridge capacitor sia pari a una capacità unitaria, servirebbe dimensionare la capacità  $C_B$  secondo la relazione:

$$C_B = C_u \cdot \frac{2^{\frac{N}{2}-1}}{\frac{N}{2} - 2} \quad (4.12)$$

ma così facendo il rapporto di attenuazione AR non risulterebbe più pari a  $2^{-\frac{N}{2}}$ . Allo stesso modo, facendo  $C_B$  pari a una capacità unitaria si otterrebbe il corretto valore del rapporto di attenuazione ma la capacità equivalente vista da ciascun array verso l'altro non sarebbe più pari all'unitaria. Entrambe le situazioni comportano un errore di guadagno; data la difficoltà di realizzare valori precisi di capacità secondo la relazione 4.12 per un numero di bit medio o elevato, si sceglie tipicamente di imporre  $C_B$  pari ad una capacità unitaria. Il conseguente errore di guadagno è equidistribuito su tutti i livelli e si traduce in un LSB equivalente pari a:

$$LSB = \frac{FSR}{2^N - 1} \quad (4.13)$$

non causa INL ma solo la perdita di un LSB alla fine della caratteristica. Tale errore è poco problematico per l'applicazione in esame e facilmente correggibile. La presenza della capacità di attenuazione rende più sensibile ai parassiti ed al mismatch la struttura, per la quale è richiesta una progettazione attenta specialmente nella fase di layout. I lavori presentati in letteratura [62] [63] [31] in cui sono proposte topologie di CR SAR ADC con questo tipo di array ottengono figure di merito comprese tra i 35fJ/cstep e i 50 fJ/cstep. Una caratteristica che è opportuno sottolineare è che come per il binario, anche per l'array con capacità di ponte è possibile l'implementazione di differenti algoritmi di switching per ottimizzare ulteriormente il consumo energetico.

#### 4.2.5 Consumo di topologie SAR con Bridge Capacitor

Il bridge capacitor array può essere approssimativamente considerato come l'insieme di due array binari classici corrispondenti a dei convertitori con la metà dei bit. Nel caso di una procedura di switching tradizionale, come la principale già analizzata per l'array binario, è possibile associare a ciascun sub array un consumo in funzione della codifica che mostra un andamento analogo a quello evidenziato in figura 4.7 per l'array binario classico. In un convertitore a N bit, ciascun sub-array ha  $2^{\frac{N}{2}}$  configurazioni possibili associate ad un numero di consumi energetici equivalenti pari a  $2^{\frac{N}{2}-1}$  (il consumo cambia ogni due variazioni di bit meno significativo). Dato che per ogni salto di codifica del

sub-array principale si susseguono tutte le  $2^{\frac{N}{2}}$  configurazioni dell'array del LSB, la caratteristica dell'energia spesa in funzione della codifica avrà un andamento che rifletterà questa proprietà. Ci si aspetta pertanto che il consumo di ogni codifica del MSB array si sommi alternativamente per i corrispondenti  $2^{\frac{N}{2}}$  livelli consecutivi ai corrispondenti consumi dell'array del LSB (la cui varietà è minore di un fattore 2). L'andamento globale della caratteristica riproduce l'andamento del consumo dell'array binario tradizionale sia a livello generale (definito dall'energia del sub-array principale) che a livello locale, cioè ogni  $2^{\frac{N}{2}}$  codifiche i corrispondenti sottolivelli si distribuiscono secondo lo stesso pattern. Quest'ultimo naturalmente dipende dal tipo di algoritmo di conversione ed un discorso analogo si può fare per altre procedure. Allo scopo di valutazione dell'efficienza sono stati realizzati modelli di convertitore di tipo bridge fully differential con diverse procedure di switching corrispondenti alle varietà presenti in letteratura. In particolare sono state valutate due differenti topologie: la tradizionale ed un array bridge con capacità del MSB analoga a quella della topologia con split capacitor (chiameremo questa topologia split-bridge) [74]. Le curve di consumo energetico in funzione della codifica di uscita ottenuta come risultato delle simulazioni con i modelli implementati sono mostrate in figura 4.20 per un convertitore a 10bit.

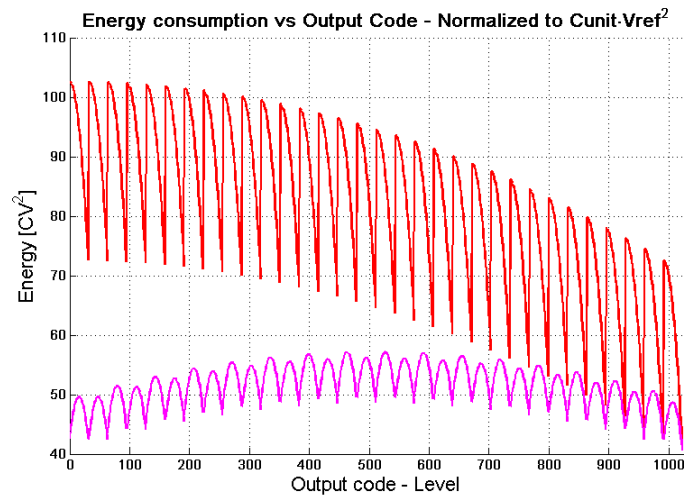


Figura 4.20: Consumo energetico per codifica delle topologie bridge analizzate.

Il consumo medio della topologia bridge tradizionale è pari circa a  $81.5C_uV_{REF}^2$ , cioè pari a poco meno del 6% rispetto al corrispondente binario, scalando circa di un fattore  $2^{\frac{N}{2}-1}$ . Per la topologia bridge-split tale cifra è ancora inferiore, di circa  $51.5C_uV_{REF}^2$ . Si noti che rispetto al consumo di un array binario con procedura tradizionale, a pari capacità unitaria l'utilizzo di una architettura split permette un consumo ridotto fino al 3.8%. Quanto appena evidenziato suggerisce come la scelta di topologie di array così efficienti e il contenimento della capacità totale renda potenzialmente l'array un componente il cui consumo non risulta più un contributo determinante per l'efficienza

dell'intero convertitore. Questa considerazione vale a meno di una riduzione equivalente dei consumi di logica e comparatore e se non si considerano le non linearità di queste topologie di DAC. In realtà, come vedremo nei prossimi paragrafi, per questioni di linearità delle architetture bridge la capacità minima dell'array dovrà essere aumentata. Una valutazione del compromesso tra linearità, area, consumo e capacità unitaria per le varie topologie binarie e bridge sarà affrontata a fine capitolo dopo aver illustrato il legame tra la DNL e l'ENOB.

#### 4.2.6 Linearità di topologie SAR con Bridge Capacitor

Rispetto alla tradizionale struttura binaria, l'architettura con capacità di attenuazione mostra una superiore sensibilità alle non idealità di processo rappresentate dal mismatch ed anche a quelle di progetto rappresentate dai parassiti. Al fine di progettare un convertitore adottando un array di questa classe è necessario poter stimare l'impatto di entrambe le sorgenti di non linearità per effettuare un dimensionamento adeguato della capacità unitaria e eventualmente valutare, infine, se il risultato è compatibile con i requisiti del sistema. In letteratura non si trova una analisi completa dell'effetto di mismatch e parassiti in topologie con capacità di attenuazione. L'architettura scelta rappresenta un esempio tipico di array con capacità di attenuazione e si è cercato di ottenere per essa delle formule chiuse che con una certa approssimazione od in modo rigoroso potessero rappresentare strumenti rapidi di stima e dimensionamento. Con l'ausilio dei modelli MATLAB implementati è stato possibile sia la verifica delle relazioni trovate che una stima il più possibile realistica e con la minima approssimazione possibile, dei risultati ottenibili al variare dei parametri in gioco. L'analisi dell'impatto del mismatch è stata impostata con lo stesso criterio adottato da Chandrakasan et al. [73] per l'analisi delle topologie binarie. I calcoli sono riferiti al caso di un array con capacità di attenuazione single ended, ma si possono facilmente estendere al caso differenziale con un miglioramento di un fattore 2 della deviazione standard massima. La tensione generata all'ingresso del comparatore dipende dalla codifica secondo la relazione:

$$V_{comp} = \left( \frac{MSD_i \cdot (\sum_{i=0}^{\frac{N}{2}} (2^{i-1} C_u + \Delta C_i))}{C_{TOT_{sub}} + C_X} + AR \cdot \frac{LSD_i \cdot (\sum_{i=0}^{\frac{N}{2}} (2^{i-1} C_u + \Delta C_i))}{C_{TOT_{sub}} + C_Y} \right) \cdot V_{REF} \quad (4.14)$$

Dove con MSD e LSD sono indicate le possibili codifiche per i due sub-array (del MSB e LSB rispettivamente).  $\Delta C_i$  indica la variazione di capacità dovuta al mismatch associata al blocco capacitivo del bit equivalente, mentre  $\Delta C_{TOT_{sub}}$  indica il mismatch della capacità totale di ciascun sub-array (esclusa la capacità di ponte):

$$C_{TOT_{sub}} = \Delta C_{TOT_{sub}} + (2^{\frac{N}{2}} - 1) C_u \quad (4.15)$$

$C_X$  e  $C_Y$  indicano invece la capacità equivalente effettivamente vista da ciascun nodo di top verso l'altro attraverso la capacità di ponte. Esse sono rappresentate dalla serie di quest'ultima con la capacità totale di ciascun subarray, cioè:

$$C_{X,Y} = \frac{C_B C_{TOT_{sub}}}{C_B + C_{TOT_{sub}}} \quad (4.16)$$

con i relativi mismatch. Tutte le variazioni di capacità possono essere espresse in funzione del mismatch relativo alla singola capacità unitaria  $\Delta C_k^i$  nel seguente modo:

$$\Delta C_i = \sum_{k=1}^{2^{i-1}} \Delta C_k^i \quad (4.17)$$

$$\Delta C_{TOT_{sub}} = \sum_{k=1}^{2^{\frac{N}{2}-1}} \Delta C_k^i \quad (4.18)$$

Il pedice  $k$  a differenza di  $i$  non indica l'unità funzionale capacitiva dell'array cui appartiene la capacità unitaria ma serve solo a sottolineare precisamente di quale capacità unitaria si tratta. Si tenga infatti presente che i blocchi capacitivi sono realizzati come parallelo di capacità unitarie ciascuna affetta da un mismatch assunto come fenomeno stocastico dagli identici momenti statistici ma ciascuna indipendente dalle altre. Sempre nella 4.14  $AR$  indica il rapporto di attenuazione tra nodo di top dell'array secondario (LSB array) e nodo di top connesso al comparatore (MSB array). Anche questo rapporto è affetto da mismatch ed è pari a:

$$AR = \frac{C_B}{C_B + C_{TOT_{sub}}} = \frac{C_u + \Delta C}{C_u + \Delta C + (2^{\frac{N}{2}} - 1)(C_u + \Delta C)} \quad (4.19)$$

cioè:

$$AR = \frac{C_u + \Delta C}{2^{\frac{N}{2}} (C_u + \Delta C)} \quad (4.20)$$

Data la natura statistica dei processi di variazione nella la semplificazione tra numeratore e denominatore non è, a rigore, appropriata. Si osservi tuttavia che il mismatch in questo rapporto non è determinante perchè riguarda un effetto il cui peso equivalente all'ingresso del comparatore è quello di una capacità unitaria, dunque può essere trascurato. Analogamente a quanto fatto in [73] si trascurino i mismatch relativi alla

capacità totale dei subarray e delle capacità  $C_{X,Y}$  che compaiono a denominatore in entrambi gli addendi della 4.14. Questa approssimazione impedisce di trovare una descrizione completa della INL ma non compromette la stima della DNL. Tenendo conto dell'indipendenza del mismatch delle singole capacità, si ha che per ogni configurazione non conta che capacità è connessa ad alimentazione ma il numero totale di capacità unitarie  $M$  nel primo sub-array e  $L$  nel secondo sub-array che hanno il bottom connesso ad alimentazione. La 4.14 può essere riscritta nel seguente modo:

$$V_{comp} = \left( \frac{\sum_{k=1}^M \cdot (C_u + \Delta C_k)}{C_{TOT_{sub}} + C_X} + AR \frac{\sum_{k=1}^L \cdot (C_u + \Delta C_k)}{C_{TOT_{sub}} + C_Y} \right) \cdot V_{REF} \quad (4.21)$$

Si considerino ora i livelli di tensione corrispondenti alle transizioni tra due codifiche di uscita, la differenza tra due livelli adiacenti dovrebbe essere idealmente pari a circa un LSB. In realtà a causa del mismatch esiste una tensione di errore tra questa differenza e il valore nominale (ottenuta senza tenere conto dei mismatch) data da:

$$\Delta V_{err} = \left( \frac{\sum_{k=1}^{\Delta M} \cdot \Delta C_k}{C_{TOT_{sub}} + C_X} + AR \frac{\sum_{k=1}^{\Delta L} \cdot \Delta C_k}{C_{TOT_{sub}} + C_Y} \right) \cdot V_{REF} \quad (4.22)$$

I differenziali  $\Delta M$  e  $\Delta L$  indicano il numero di mismatch di capacità unitarie indipendenti coinvolti nella data transizione di codifica. Il caso peggiore si ha quando la differenza nel numero di capacità unitarie indipendenti connesse ad alimentazione tra due codifiche adiacenti è massima. Nelle architetture bridge, esattamente come nelle binarie, questa eventualità si verifica in corrispondenza della codifica appena inferiore al MSB, dove sono coinvolti sia i mismatch della capacità del MSB, sia quelli legati a tutte le altre capacità dell'array. Per maggiore chiarezza, la situazione è illustrata in figura 4.21. Nell'immagine è sottolineato come sia la transizione tra la codifica del MSB e il livello inferiore, e non quello superiore, a determinare il caso peggiore perchè coinvolge il maggior numero di processi stocastici indipendenti.

Si ha pertanto :

$$\Delta M = 2 \cdot 2^{\left(\frac{N}{2} - 1\right)} \quad (4.23)$$

$$\Delta L = 2^{\frac{N}{2}} - 1 \quad (4.24)$$

Trattandosi di processi stocastici e volendo valutare la loro varianza per risalire alla deviazione standard della DNL, il segno delle variazioni non è preso in considerazione e si considerano tutte concordi. In ogni caso il mismatch delle capacità appartenenti all'array secondario viene attenuato di un fattore  $AR$ , che affetto o non affetto da mismatch, rende il ruolo del mismatch appartenente all'array del LSB equiparabile o inferiore a quello di una capacità unitaria dell'array principale. La varianza dell'errore  $\Delta V_{err}$  può essere

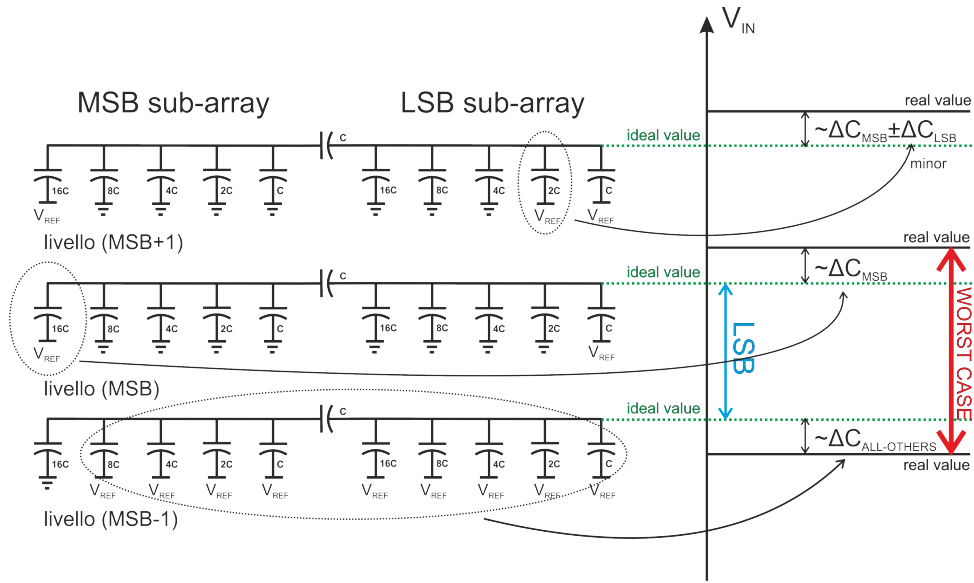


Figura 4.21: Schematizzazione della situazione potenzialmente peggiore per la DNL: la transizione a cavallo della codifica MSB.

calcolata tenendo conto dell'indipendenza di tutte le diverse capacità unitarie, si ha quindi:

$$\sigma_{err}^2 \approx 2 \cdot \frac{2^{(\frac{N}{2}-1)} \sigma_C^2}{2^N C_u} V_{DD}^2 = \frac{\sigma_C^2}{2^{\frac{N}{2}} C_u^2} V_{DD}^2 \quad (4.25)$$

la sua deviazione standard è pertanto pari a:

$$\sigma_{err} \approx \frac{\sigma_C}{2^{\frac{N}{4}} C_u} V_{DD} \quad (4.26)$$

Dividendo il risultato per l'LSB nominale si ottiene la corrispondente deviazione standard di DNL:

$$\sigma_{DNL_{max}} \approx 2^{\frac{3}{4}N} \cdot \frac{\sigma_C}{C_u} \quad (4.27)$$

Confrontandola la grandezza appena definita con quella equivalente della topologia binaria classica (BWC, Binary Weighted Classico) si ottiene:

$$\sigma_{DNL_{max}} \approx 2^{\frac{N}{4}} \cdot \sigma_{DNL_{max}BWC} \quad (4.28)$$

La capacità di ponte infatti, dividendo l'array in due metà con una capacità totale equivalente ridotta di un fattore  $2^{\frac{N}{2}}$ , peggiora dello stesso valore la varianza della DNL, e dunque la sua deviazione standard risulta superiore di un fattore  $2^{\frac{N}{4}}$ . Le simulazioni confermano questo risultato; in figura 4.22 è mostrata la distribuzione delle deviazioni standard per ogni codifica di uscita per un array a 10 bit di classe bridge realizzato con i parametri della tecnologia a disposizione e una capacità unitaria di 20fF. Si noti come per la codifica centrale (appena inferiore al MSB) si ha una deviazione standard di circa 0.4, esattamente come ricavabile sostituendo alla 4.28 il numero di bit, i parametri di tecnologia e la capacità unitaria.

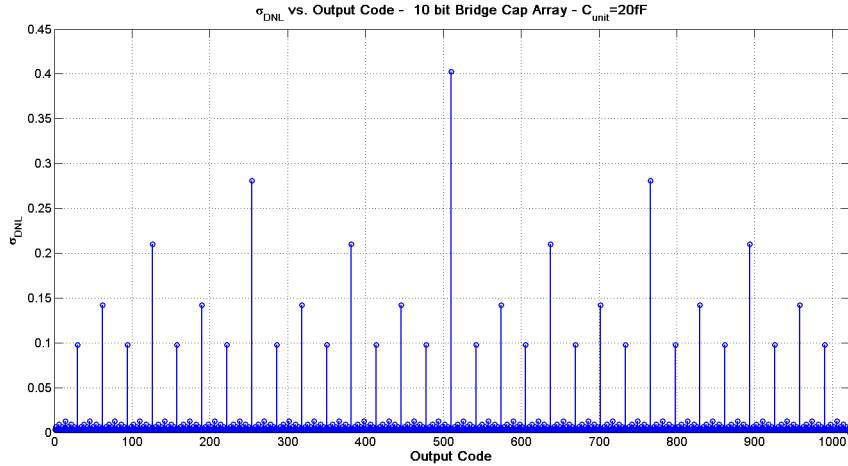


Figura 4.22: Deviazione standard della DNL in funzione della codifica per un Bridge Array a 10bit e capacità unitaria di 20fF.

Dalla 4.28, imponendo che il valore di tre deviazioni standard di DNL sia inferiore a 0.5 si otterrebbe una capacità minima di almeno 118fF nel caso single ended (29fF nel caso fully differential) circa  $2^{\frac{N}{2}}$  superiore rispetto ai 3.8 fF calcolati nel single ended binario classico. Inoltre, il confronto tra topologia bridge con algoritmo di commutazione classico e la topologia split-bridge mostrano le stesse relazioni che si hanno nell'insieme dei binari. Come ricavabile con un approccio analitico e verificabile da simulazioni, la topologia split bridge presenta una deviazione standard nel caso della codifica dell'MSB che si dimostra migliore di un fattore  $\sqrt{2}$  rispetto alla tradizionale :

$$\sigma_{DNL_{max}} \approx \frac{2^{\frac{3}{4}N} \sigma_C}{\sqrt{2} C_u} = 2^{\frac{N}{4}} \cdot \sigma_{DNL_{max,bin-split}} \quad (4.29)$$



Il passaggio ad una equivalente configurazione fully differential del convertitore porta a un dimezzamento del valore della deviazione standard calcolata nella 4.28. Un altro elemento che non è stato ancora preso in considerazione e che risulta determinante rispetto alle prestazioni del convertitore è l'effetto delle capacità parassite. Sempre con riferimento ad un array singolo, esse sono perlopiù schematizzabili come in figura 4.23, dove sono rappresentate una capacità  $C_{pA}$  vista verso massa in parallelo all'array del MSB, e una capacità  $C_{pB}$  equivalente sull'array secondario.

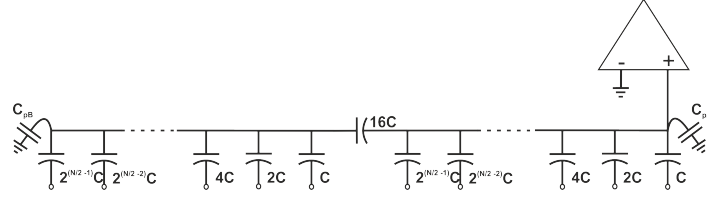


Figura 4.23: Principali parassiti di interesse in un array con capacità di ponte.

In generale esiste la possibilità di avere altri tipi di parassiti oltre a quelli raffigurati, ma non sono stati evidenziati per dei motivi precisi. Le capacità indesiderate tra ciascun bottom e il top delle capacità all'interno di un subarray, infatti, hanno un effetto qualitativamente equivalente ad un mismatch, devono essere minimizzati in fase di layout ma non introducono novità in quanto al tipo di effetto prodotto. Al contrario, i parassiti che da un bottom plate si accoppiano con il nodo di top dell'array alternativo, passando a cavallo della capacità di ponte, possono avere un effetto devastante sulla DNL ma non è difficile realizzare dei layout praticamente esenti da questa categoria di parassita, pertanto non verranno trattati.

In letteratura non è finora stata presentata una analisi dettagliata dell'effetto dei parassiti di figura 4.23 sulla DNL di un Bridge Array. In questo lavoro è presentata una analisi rigorosa che porta all'ottenimento di una formula chiusa che lega la DNL ai parassiti e che è supportata da risultati di simulazione con modelli appositamente costituiti. A differenza del mismatch, l'effetto dei parassiti può essere trattato più facilmente come un effetto deterministico. La tensione generata al nodo di ingresso del comparatore in una architettura come quella raffigurata in figura 4.23 è:

$$V_{comp} = \left( \frac{\sum_{i=1}^{\frac{N}{2}} C_i \cdot MSD_i}{C_{TOT_{sub}} + C_X + C_{pA}} + AR \cdot \frac{(\sum_{i=1}^{\frac{N}{2}} C_i \cdot LSD_i)}{C_{TOT_{sub}} + C_Y + C_{pB}} \right) \cdot V_{REF} \quad (4.30)$$

con il rapporto AR che stavolta è:

$$AR = \frac{C_B}{C_B + C_{TOT_{sub}} + C_{pA}} = \frac{C}{2^{\frac{N}{2}} C + C_{pA}} \quad (4.31)$$

mentre le capacità  $C_X$  e  $C_Y$  sono:

$$C_X = \frac{C_B(C_{TOT_{sub}} + C_{pB})}{C_B + C_{TOT_{sub}} + C_{pB}} = \frac{C(2^{\frac{N}{2}} - 1) + C_{pB}}{2^{\frac{N}{2}}C + C_{pB}} \quad (4.32)$$

$$C_Y = \frac{C_B(C_{TOT_{sub}} + C_{pA})}{C_B + C_{TOT_{sub}} + C_{pA}} = \frac{C(2^{\frac{N}{2}} - 1) + C_{pA}}{2^{\frac{N}{2}}C + C_{pA}} \quad (4.33)$$

Le capacità totale equivalenti  $C_{TSE_{1,2}}$  viste verso massa dai nodi di top di ciascun subarray (rispettivamente di MSB e LSB) sono:

$$C_{TSE_1} = C_X + C_{TOT_{sub}} + C_{pB} \quad (4.34)$$

$$C_{TSE_2} = C_Y + C_{TOT_{sub}} + C_{pA} \quad (4.35)$$

e sostituendo ai vari termini le loro espressioni precedentemente ricavate si ottiene:

$$C_{TSE_1} = \frac{C^2(2^N - 1) + 2^{\frac{N}{2}}C(C_{pA} + C_{pB}) + C_{pA}C_{pB}}{2^{\frac{N}{2}}C + C_{pB}} \quad (4.36)$$

$$C_{TSE_2} = \frac{C^2(2^N - 1) + 2^{\frac{N}{2}}C(C_{pA} + C_{pB}) + C_{pA}C_{pB}}{2^{\frac{N}{2}}C + C_{pA}} \quad (4.37)$$

Si osservi l'identità del numeratore tra le due espressioni; d'ora in poi sarà indicato con  $\Gamma$ . In ciascuna codifica D si avranno un numero  $M_{(D)}$  ed un numero  $L_{(D)}$  di capacità unitarie totali il cui bottom plate è connesso ad alimentazione (riferimento positivo). Alla luce di quanto esaminato finora la 4.30 può essere scritta nella forma:

$$V_{comp} = \left( \frac{M_{(D)} \cdot C_i}{C_{TSE_1}} + AR \cdot \frac{L_{(D)} \cdot C_i}{C_{TSE_2}} \right) \cdot V_{REF} \quad (4.38)$$

che, sostituendo le espressioni delle varie capacità in funzione di parassiti e capacità unitaria, diventa:

$$V_{comp} = \left( \frac{M_{(D)} \cdot C(2^{\frac{N}{2}}C + C_{pB})}{\Gamma} + \frac{L_{(D)} \cdot C^2}{\Gamma} \right) \cdot V_{REF} \quad (4.39)$$

La 4.39 mostra come solo il parassita dell'array del LSB  $C_{pB}$  sia responsabile di non linearità, mentre  $C_{pA}$  introduce un semplice errore di guadagno. La differenza tra il livello associato a due codifiche adiacenti può essere scritto in funzione delle variazioni dei numeri di capacità unitaria connessi al riferimento superiore. Tenendo conto di ciò e sostituendo  $\Gamma$  la sua espressione si ha:

$$\Delta V_{comp_{D,D-1}} = \left( \frac{\Delta M_{D,D-1} \cdot C(2^{\frac{N}{2}} C + C_{pB}) + \Delta L_{D,D-1} \cdot C^2}{C^2(2^N - 1) + 2^{\frac{N}{2}} C(C_{pA} + C_{pB}) + C_{pA} C_{pB}} \right) \cdot V_{REF} \quad (4.40)$$

Il caso pessimo di DNL si può a questo punto valutare individuando il massimo scostamento tra la differenza di tensione  $\Delta V_{comp_{D,D-1}}$  e il valore ideale del LSB:

$$DNL_{max} = \frac{\Delta V_{comp_{D,D-1}}^{max}}{LSB} - 1 \quad (4.41)$$

e tale circostanza si verifica, in un array di tipo bridge, ogni  $2^{\frac{N}{2}}$  codifiche, quando:

$$\Delta M_{D,D-1} = 1; \quad (4.42)$$

$$\Delta L_{D,D-1} = -(2^{\frac{N}{2}} - 1) \quad (4.43)$$

Se si sostituiscono questi valori nella 4.40 e si applica la 4.41 considerando il LSB classico di un ADC a 10bit si ottiene:

$$\frac{\Delta V_{err}}{LSB} = \frac{(2^N - 2^{\frac{N}{2}})C_{pB} - C_{pA}2^{\frac{N}{2}} + C - \frac{C_{pA}C_{pB}}{C}}{C(2^N - 1) + 2^{\frac{N}{2}}(C_{pA} + C_{pB}) + \frac{C_{pA}C_{pB}}{C}} \quad (4.44)$$

In generale la 4.44 mostra che anche il parassita  $C_{pA}$  contribuisce all'errore di ampiezza di ogni gradino della caratteristica rispetto al LSB, ma rende conto di un contributo equidistribuito associato all'errore di guadagno che essa introduce. L'effetto propriamente non lineare si deve unicamente a  $C_{pB}$ , si dovrebbe calcolare utilizzando un LSB che tenga conto degli effetti lineari (errori di guadagno) ed è pari a:

$$DNL_{max} = \frac{(2^N - 2^{\frac{N}{2}})C_{pB} + C - \frac{C_{pA}C_{pB}}{C}}{C(2^N - 1) + 2^{\frac{N}{2}}(C_{pB} + C_{pA}) + \frac{C_{pA}C_{pB}}{C}} \quad (4.45)$$

Si noti come il parassita all'ingresso del comparatore non solo non introduce errori di linearità, ma attenua quelli provocati da  $C_{pB}$ . Per questo è conveniente, in fase di progetto, rivolgere il bottom plate della capacità di attenuazione (più affetto da parassiti) verso l'array principale piuttosto che verso il secondario e preoccuparsi soprattutto di ridurre il parassita al nodo di top dell'array secondario. Il risultato delle simulazioni confermano quanto ricavato analiticamente. In figura 4.24 è mostrata a titolo di esempio la DNL di un convertitore a 10 bit con capacità unitaria 30fF e entrambi i parassiti di 20fF.

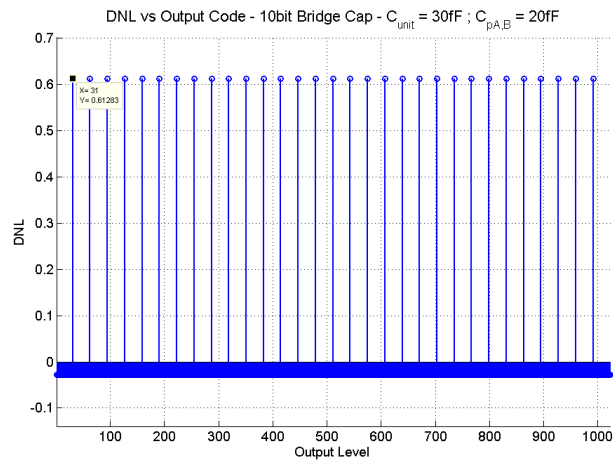


Figura 4.24: Caratteristica DNL di un Bridge Capacitor Array con parassiti.

L'immagine mostra una caratteristica DNL ottenuta dagli errori di guadagno dovuti a  $C_{pA}$ . Essa mostra picchi regolari di DNL di circa 0.6, esattamente pari al calcolo analitico della 4.45. In un fully differential a pari capacità minima l'LSB raddoppia mentre i parassiti entrano in gioco nello stesso modo e l'errore si riduce di un fattore due indipendentemente dalla procedura di commutazione.

#### 4.2.7 Confronto tra Topologie

Nella tabella 4.1 sono riportate le espressioni delle deviazioni standard massime per le topologie bridge e binarie analizzate nei precedenti paragrafi ricavate analiticamente con lo stesso approccio proposto in [73].

Un breve confronto tra le tipologie classica binaria e bridge mostra che la capacità unitaria necessaria a fare in modo che un valore pari a tre volte la deviazione standard della DNL massima (transizioni prossime al MSB) sia inferiore a metà del LSB è pari a 3.8fF (circa 4fF) nel primo e 118fF nel secondo, un fattore circa  $2^{\frac{N}{2}}$  di differenza. Una simile differenza renderebbe l'impiego di topologie bridge evidentemente sconveniente rispetto alle binarie.

Nella seconda colonna della tabella 4.2 sono riportati, per ciascuna topologia, i valori di capacità unitaria necessaria a mantenere il triplo della deviazione standard di DNL

Tabella 4.1: Tabella riassuntiva dei legami tra capacità unitaria, parametri di tecnologia e deviazione standard massima della DNL per ciascuna topologia esaminata.

$max \sigma_{DNL} $		
Topologia	Single Ended	Fully Differential
Binario Classico	$2^{\frac{N}{2}} \cdot \frac{\sigma_C}{C_u}$	$2^{\frac{N}{2}} \cdot \frac{\sigma_C}{\sqrt{2}C_u}$
Binario Split	$2^{\frac{N}{2}} \cdot \frac{\sigma_C}{\sqrt{2}C_u}$	$2^{\frac{N}{2}} \cdot \frac{\sigma_C}{2C_u}$
Binario Monotono	NA	$2^{\frac{N}{2}} \cdot \frac{\sigma_C}{\sqrt{2}C_u}$
Bridge Classico	$2^{\frac{3N}{4}} \cdot \frac{\sigma_C}{C_u}$	$2^{\frac{3N}{4}} \cdot \frac{\sqrt{2}\sigma_C}{C_u}$
Bridge Split	$2^{\frac{3N}{4}} \cdot \frac{\sigma_C}{\sqrt{2}C_u}$	$2^{\frac{3N}{4}} \cdot \frac{\sigma_C}{2C_u}$

massima al di sotto della metà di un LSB. Nella seconda e terza colonna sono rispettivamente riportati, in tutti i casi, l'equivalente capacità totale ed il consumo medio per topologie a 10bit.

Tabella 4.2: Tabella riassuntiva delle capacità unitarie, totali e consumi medi per avere il triplo della deviazione standard massima della DNL inferiore a 0.5.

Topologia	$C_{unit,min}$	$C_{TOT}$	$E_{avg,conversion}$ [fJ]
Binario Classico S.E.	4fF	4.096pF	1365fJ
Binario Classico F.D.	2fF	4.096pF	683fJ
Binario Split S.E.	2f	2.048pF	426fJ
Binario Split F.D.	1fF	2.024pF	213fJ
Binario Monotono	2fF	2.048pF	127fJ
Bridge Classico S.E.	120fF	7.56pF	2445fJ
Bridge Classico F.D.	60fF	7.56pF	1225fJ
Bridge Split S.E.	60fF	3.78pF	772fJ
Bridge Split F.D.	29.5fF	3.78pF	386fJ

I numeri evidenziano come accettando di dimensionare l'array secondo le relazioni riportate nella tabella 4.1 gli array bridge fully differential ottengono capacità totali superiori o prossimi al limite della compatibilità con l'applicazione e consumi paragonabili a quelli degli array binari. Questi ultimi invece ottengono capacità totali anche inferiori, ma richiedono necessariamente la realizzazione di capacità full-custom i cui parametri di mismatch non corrisponderanno a quelli forniti dalla casa produttrice per le capacità di libreria e pertanto potrebbero portare a risultati peggiori delle attese. La topologia più interessante tra quelle riportate è la Bridge-Split, che nel caso fully differential permetterebbe di utilizzare le capacità di libreria (29.5fF) e ottenere capacità totale e consumi compatibili con l'applicazione, ma aumenterebbe la complessità della logica perchè richiederebbe il controllo di due array aggiuntivi associate alle capacità del MSB. Quello della tabella 4.2 rappresenta un sovradimensionamento fondato sul caso pessimo della

DNL tra tutte le codifiche ed atto a ad avere caratteristiche con il 99% di probabilità di non avere picchi di DNL superiori a 0.5. Se da un lato questo dimensionamento permette di limitare molto l'impatto dei mismatch, dall'altro non contiene quello dei parassiti. L'effetto di questi ultimi, come vedremo più avanti, è più significativo nelle architetture bridge. Anche nelle architetture binarie realizzate con capacità full-custom molto piccole, tuttavia, pochi fF di parassita tra il bottom e il top plate delle capacità potrebbero risultare devastanti. Il problema dei parassiti suggerisce al progettista di aumentare ulteriormente la capacità unitaria in base a previsioni sulla possibile entità degli accoppiamenti capacitivi indesiderati tra i vari nodi del circuito, difficile da prevedere. A questo punto le dimensioni ed il consumo dell'array potrebbero diventare un problema.

Scegliendo di occuparsi, almeno per il momento, unicamente dell'impatto del mismatch, non è ancora chiaro quale sia il vero rapporto, a pari prestazioni di linearità tra la capacità unitaria delle topologie binarie e di quelle bridge. Una condizione così restrittiva come quella imposta da una qualsiasi delle relazioni riportate nella tabella 4.1 potrebbe non essere strettamente necessaria. Nel contesto dell'elaborazione di segnali analogici, infatti, il principale parametro di riferimento per la qualità della conversione è l'ENOB e non è detto che non sia possibile ottenerne uno compatibile con l'applicazione anche facendo ricorso a dimensionamenti più rilassati della capacità.

Inoltre, si è visto in precedenza che la deviazione standard della DNL in funzione delle codifiche presenta delle distribuzioni caratteristiche per le diverse topologie di array ed è possibile che l'ENOB sia influenzato anche da questi pattern di distribuzione. Allo scopo di avere stime più realistiche dell'effetto del mismatch e di verificare la possibilità di ridurre ulteriormente la capacità unitaria sono state svolte simulazioni *MATLab<sup>TM</sup>* con i modelli implementati per tutte le topologie allo scopo di stimare la media e la deviazione standard del massimo valore assoluto della DNL.

Gli istogrammi per la prima delle grandezze elencate in un bridge array single ended con procedura tradizionale sono illustrati come esempio in figura 4.25. Come già illustrato in precedenza per l'omologa topologia binaria, si è considerato un convertitore a 10bit ed una capacità minima di 20fF.

Per ottenere lo stesso risultato di  $|DNL|_{max}$  della topologia binaria realizzata a pari capacità unitaria servirebbe aumentare quest'ultima di un fattore di poco superiore a  $2^{(\frac{N}{2}-2)}$  cioè nel caso dei un convertitore 10 bit, poco superiore a 8. Si tratta di un aumento significativo ma molto inferiore a quello ricavabile utilizzando la 4.28 ed evidente nella tabella 4.2.

Si potrebbe procedere nel generare istogrammi per ogni topologia al variare della capacità unitaria e poi scegliere la più adatta a seconda della distribuzione ottenuta e della capacità totale. Sarebbe tuttavia preferibile avere uno strumento più rigoroso per ricavare numeri utili al dimensionamento, cioè uno strumento in grado di collegare la capacità unitaria alla linearità richiesta dall'applicazione. L'osservazione degli istogrammi ottenuti dai modelli realistici dei vari convertitori offre, infatti, un'informazione in più ma non fornisce un vero strumento per capire fino a quali capacità minime (da ridurre per ottimizzare area e consumi) ci si può spingere pur tendendo sotto controllo la de-

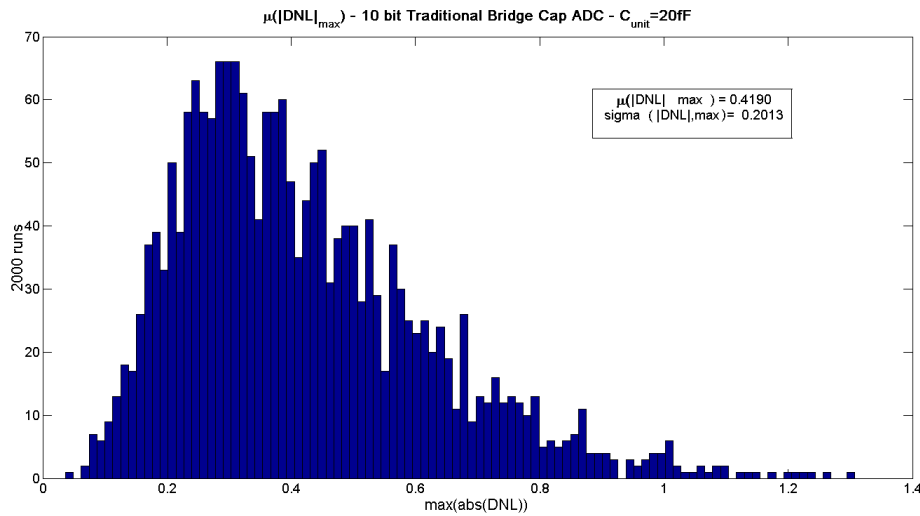


Figura 4.25: Istogrammi per un array con capacità di attenuazione e capacità minima 20fF.

gradazione della linearità al di sotto del limite imposto dall'applicazione. Stabilire quale sia il valore limite accettabile con gli strumenti finora a disposizione non è possibile.

Un criterio più adeguato nel contesto del progetto dovrebbe essere fondato sul legame tra capacità ed ENOB, caratteristico di ogni singola architettura. In questo contesto è proposto un criterio progettuale nuovo fondato su un tipo di simulazioni differente rispetto a quelle utilizzate finora. Come si è visto, ogni topologia è caratterizzata da struttura fisica dell'array e algoritmo di commutazione e da una distribuzione delle  $\sigma_{DNL}$  tra tutte le codifiche, dipendente da numerosi fattori: struttura, algoritmo di commutazione, capacità unitaria, numero di bit e mismatch tecnologici. In ogni topologia, qualsiasi siano questi parametri, la distribuzione presenta un picco massimo in corrispondenza del MSB, calcolabile con le relazioni della tabella 4.1 e rispetto al quale tutti gli altri possono essere normalizzati per ottenere un vero e proprio pattern, come quello di figura 4.26.

Dati i parametri tecnologici e una capacità unitaria è possibile ottenere la  $\sigma_{DNL,max}$  per ogni topologia. Moltiplicando quest'ultima per un pattern è possibile ottenere una distribuzione realistica delle deviazioni standard di ogni topologia che è strettamente legata alla capacità unitaria scelta. Con l'ausilio di *MATlab*<sup>TM</sup>, per ciascuna topologia, al variare della capacità unitaria e per ogni suo valore sono state simulate 100 realizzazioni di caratteristiche e sono state utilizzate per il calcolo dell'ENOB secondo lo schema di simulazione di figura 4.27. Ogni caratteristica è stata utilizzata per convertire una sinusoide ad ampiezza crescente ed attraverso il calcolo delle FFT in uscita è stato possibile ottenere l'ENOB massimo di ciascuna e costruirne una statistica per ogni topologia, definirne cioè con valori medi e deviazione standard al variare della capacità unitaria.

Data la specifica del sistema di almeno 8.2 bit di codifica, sono state svolte simulazioni

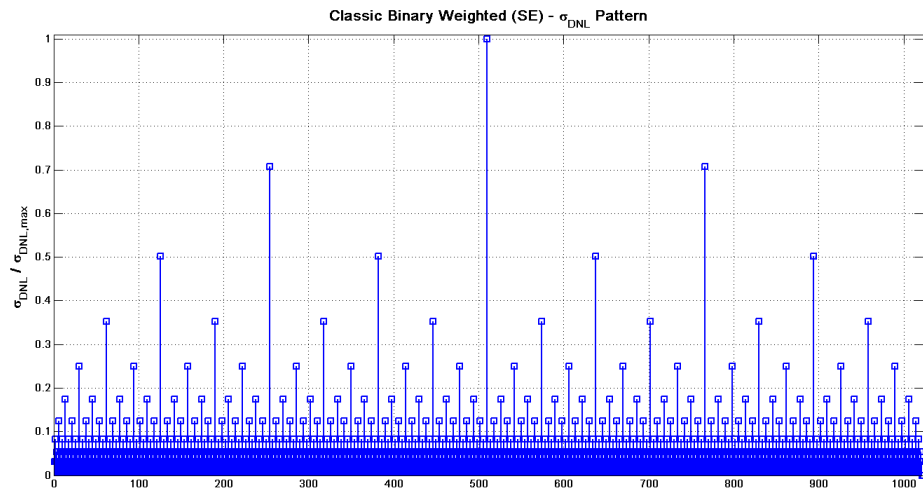


Figura 4.26: Esempio di pattern di distribuzione delle deviazioni standard della DNL.

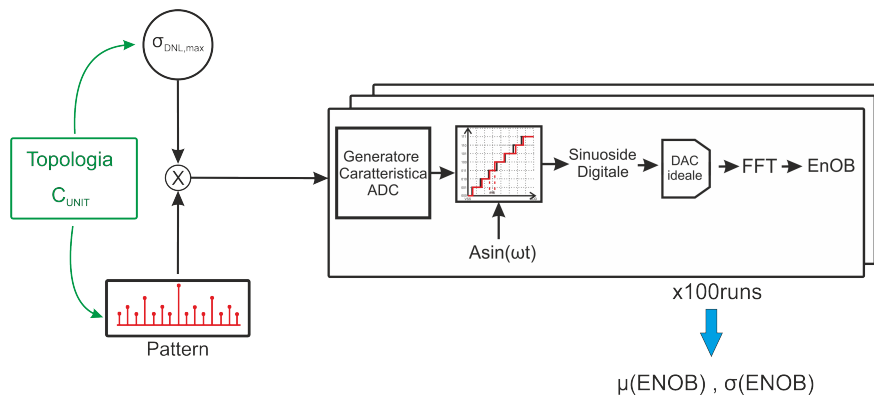


Figura 4.27: Schema delle simulazioni utilizzate per ottenere l'enob.



per verificare quale capacità unitaria per ogni topologia (a 10 bit) garantisca che il 99% delle realizzazioni corrispondessero ad un ENOB di almeno 9. La tabella 4.3 è l'omologa della 4.2 ma basata su questo nuovo criterio di calcolo.

Tabella 4.3: Tabella riassuntiva delle capacità unitarie, totali e consumo medio per ogni topologia e per avere un ENOB superiore a 9.

Topologia	$C_{unit,min}$	$C_{TOT}$	$E_{avg,conversion}$
Binario Classico S.E.	1fF	1.024pF	341fJ
Binario Classico F.D.	0.5fF	1.024pF	170fJ
Binario Split S.E.	0.5fF	512fF	170fJ
Binario Split F.D.	0.25fF	512fF	85.3fJ
Binario Monotono	1.6fF	1.638pF	102.5fJ
Bridge Classico S.E.	18fF	1.134pF	366.75fJ
Bridge Classico F.D.	9fF	1.134fF	183.25fJ
Bridge Split S.E.	16fF	1.008pF	206fJ
Bridge Split F.D.	8fF	1.008pF	103fJ

Si osserva un significativo rilassamento delle specifiche sulla capacità unitaria per tutte le topologie ed un conseguente miglioramento dei consumi. Tuttavia mentre per le binarie questo miglioramento probabilmente non è realistico a causa delle variazioni dei parametri di mismatch dovuto alla necessità di realizzare unità full-custom, nelle architetture bridge, almeno in quelle single-ended, i valori di capacità si riducono ma si allontanano meno dal range di capacità standard della tecnologia e permettono di ottenere aree e consumi molto interessanti. Questo significa che il tipo di distribuzione della deviazione standard della DNL caratteristico di ogni topologia è fortemente influente sulle non linearità. Dunque anche l'INL, che dipende esattamente da come le DNL sono distribuite, ha un ruolo determinante nel progetto di convertitori per segnali analogici. Come già detto in precedenza, si è preferito trattare il problema utilizzando la DNL piuttosto che la INL perchè è risultato più semplice legarne i parametri statistici alla capacità unitaria con formule chiuse. In ogni caso, la trattazione è equivalente, perchè a meno di errori di offset, la DNL contiene sempre anche l'informazione sulla morfologia e la ampiezza della INL ad ogni codifica.

### 4.3 L'Architettura Scelta

Dalla tabella 4.3 illustrata nel precedente paragrafo risulta molto interessante la scelta di topologie Bridge perchè permettono di ottenere strutture compatte ed efficienti senza dover ricorrere ad un progetto full-custom delle capacità. La più efficiente tra queste risulta essere la Split-Bridge. Per il progetto tuttavia è stata scelta una architettura ancora differente, finora mai presentata in letteratura ed ottenuta dalla combinazione dell'architettura bridge con la procedura di commutazione monotona già vista per array binari. Possiamo definire questa nuova topologia con la denominazione

Monotonic-Bridge. L'array utilizzato è illustrato insieme ad un comparatore ideale in figura 4.28.

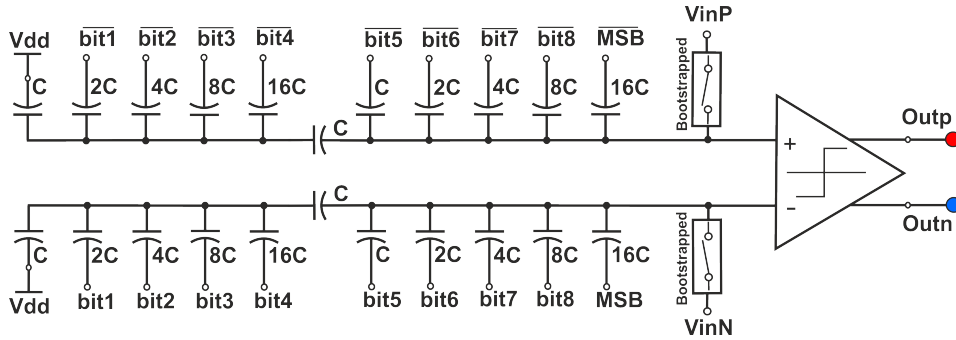


Figura 4.28: Topologia con Bridge Capacitor array scelta.

A differenza del caso binario, in presenza di capacità di attenuazione non è possibile eliminare la capacità del MSB. Per la linearità, infatti, l'array deve essere simmetrico e tale per cui la capacità vista da ciascun sub-array verso l'altro, attraverso il condensatore di ponte, sia identica.

Questo appare come uno svantaggio in termini di area, ma in realtà è più che bilanciato dalla riduzione di capacità tipica della architettura bridge. Al contrario, dal punto di vista del mismatch l'aumento della capacità totale al sub-array del MSB costituisce un vantaggio perchè attenua con effetto quadratico l'effetto dei mismatch. La deviazione standard massima della DNL, come nelle altre topologie, la si ha sempre in corrispondenza del MSB. Essa è stata stimata analiticamente con calcoli tediosi che non riportiamo e risulta essere circa pari a:

$$\sigma_{DNL,max} \approx 2^{\frac{3N}{4}} \cdot \frac{\sigma_C}{2C_u} \quad (4.46)$$

Questo valore è identico a quello della topologia Split-Bridge completamente differenziale; la monotona però presenta un algoritmo di commutazione molto più semplice e permette l'impiego di una logica più compatta. Inoltre, come nel caso degli array binari, la monotona risulta la procedura di commutazione più efficiente. Il suo consumo medio è pari a  $30.5C_u V_{REF}^2$  e l'energia per ogni codifica è rappresentata e confrontata con quella della altre architetture bridge in figura 4.29.

Lo schema della procedura di commutazione utilizzata nella struttura di array scelta è illustrato in figura 4.30 per un convertitore a 4 bit.

Ad ogni passo successivo al campionamento, come in ogni SAR ADC, viene codificato un bit procedendo dal più significativo al meno significativo e di volta in volta la particolare configurazione dell'array riproduce un livello di tensione corrispondente alla transizione tra due codifiche o insiemi di codifiche successive. Nella procedura

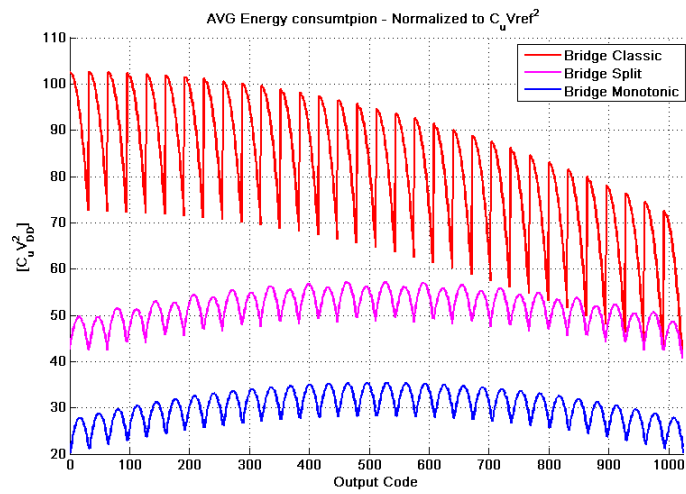


Figura 4.29: Confronto tra le energie spese durante il ciclo di commutazione del Monotono e delle altre topologie Bridge.

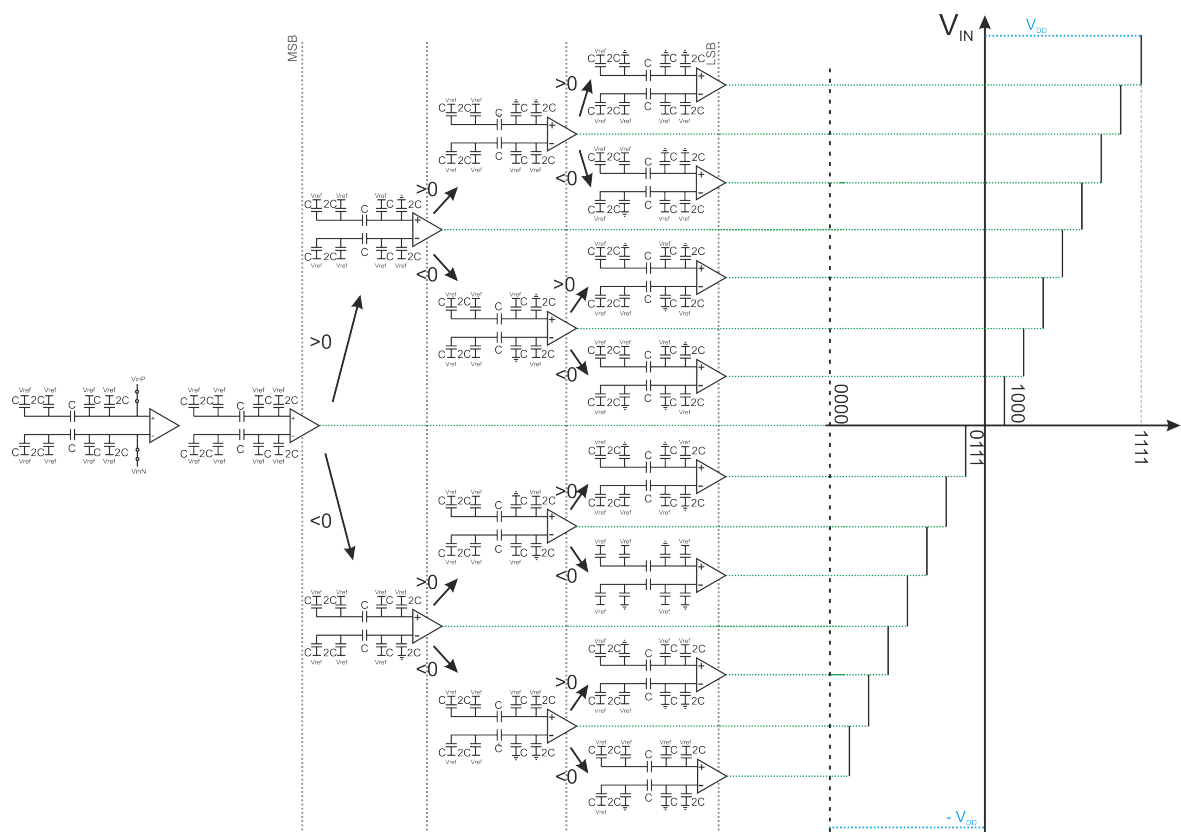
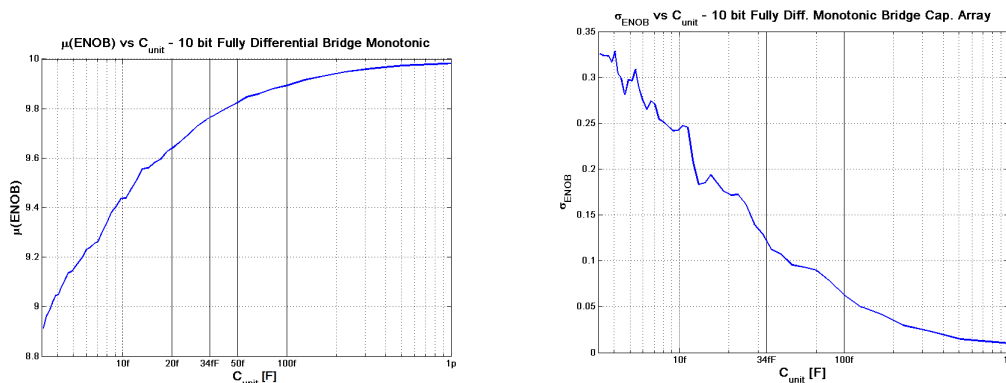


Figura 4.30: Schema di funzionamento della topologia scelta.

monotonica durante la fase di sampling tutti i bottom delle capacità sono connessi all'alimentazione del convertitore e la corrente in fase di sampling è assorbita proprio da quest'ultima. Passato un tempo di campionamento  $T_S$  si aprono gli interruttori di ingresso, il segnale è memorizzato sulla capacità equivalente vista dal top dell'array del MSB e può iniziare la fase di comparazione. Ad ogni passo, a seconda dell'esito della comparazione viene connessa a massa la capacità del bit corrispondente di uno solo dei due array. Non vi è nessuna differenza algoritmica rispetto alla procedura monotonica illustrata in figura 4.8 per un array classico, varia solo la struttura dell'array. La tensione di modo comune dei nodi di ingresso del comparatore scende ad ogni confronto della metà della tensione equivalente al bit codificato, fino ad avvicinarsi gradualmente allo zero. Questa caratteristica, come vedremo, rende necessario un attento progetto del comparatore e lo studio della variabilità della sua tensione di offset con il modo comune dell'ingresso, che rappresenta un possibile fonte di non linearità.

Il risultato delle simulazioni effettuate in MATLAB per controllare l'effetto del mismatch in modo da garantire un ENOB superiore a 9 nel 99% dei casi indicano la necessità di una capacità unitaria di almeno 16fF, ben inferiore ai 30fF che otterremmo dal dimensionamento solo basato sulla 4.46. La topologia split bridge nel caso fully differential sembrerebbe essere più vantaggiosa perchè dovrebbe permettere l'utilizzo di 8fF, in realtà tale valore cade nell'insieme delle capacità realizzabili solo in modo full-custom e risulta meno affidabile. Inoltre, l'efficienza quasi doppia e la maggiore semplicità della topologia monotona realizzabile utilizzando capacità di libreria giustificano la scelta adottata.

In figura 4.31 è mostrato l'andamento di media e deviazione standard dell'ENOB in funzione della capacità unitaria come risultato di simulazioni comprensive del solo effetto del mismatch.



(a) Media dell'ENOB in funzione della capacità unitaria. (b) Deviazione standard dell'ENOB in funzione della capacità unitaria.

Figura 4.31: Risultati delle simulazioni effettuate con i modelli di ADC Bridge Monotono per la valutazione dell'ENOB in funzione della capacità unitaria e del mismatch tecnologico.

Per conservare l'affidabilità dei parametri di mismatch tecnologico e ridurre lo sforzo

progettuale si può pensare di realizzare l'array con capacità minima di libreria, cioè 17fF. Con questo valore otterremmo una capacità totale di circa 2pF, di cui solo 1pF visto a ciascun nodo d'uscita del PGA, e un consumo durante il ciclo di comparazione di 129fJ. A questo punto è opportuno sottolineare che in realtà occorre un ulteriore dimensionamento per cercare di prevedere e ridurre l'effetto dei parassiti ed in particolare quello caratteristico presente nelle architetture bridge per le quali sono state ricavate le formule chiuse 4.44 e 4.45 nei precedenti paragrafi. Questo è uno dei temi cui è dedicato il prossimo capitolo.

## Capitolo 5

# Progetto e Dimensionamento del Convertitore

In questo capitolo sono illustrati i dettagli di progetto del convertitore effettivamente progettato: il dimensionamento dell'array, la scelta ed il progetto del comparatore e la realizzazione della logica SAR.

### 5.1 Dimensionamento dell'array e Risultati Pre-Layout

Il dimensionamento della capacità unitaria dell'array sulla base delle specifiche di ENOB, considerando unicamente le non idealità rappresentate dal mismatch, è stato affrontato nel precedente capitolo. In particolare si è osservato che per la topologia scelta, 16fF di capacità unitaria potrebbero garantire 9 bit effettivi. Scegliendo di utilizzare solo componenti di libreria si avrebbe a disposizione una capacità minima di 17fF ed il mismatch non rappresenterebbe dunque un problema. In realtà il dimensionamento deve essere deciso anche in funzione dei potenziali parassiti dell'array. Gli accoppiamenti capacitivi tra piatto inferiore e superiore di ogni capacità agiscono in modo equivalente a dei mismatch. Ciò suggerirebbe un aumento della capacità unitaria, ma quantificare a priori l'entità dell'aumento è complesso perchè è difficile prevedere i parassiti. Ciò che però è maggiormente controllabile è l'effetto dei parassiti connessi tra il top plate di ciascun sub-array e la massa, i cui effetti sono stati derivati analiticamente nel precedente capitolo in cui è stato anche mostrato che l'unica fonte di non linearità è il parassita al top plate del LSB:  $C_{pB}$ . Il legame tra  $DNL_{max}$ , parassita  $C_{pB}$  e capacità unitaria per la tipologia di ADC fully differential bridge con procedura monotona scelta per il progetto, è rappresentato in figura 5.1. Le curve sono la fedele rappresentazione delle relazioni analitiche espresse nel precedente paragrafo; una di queste, evidenziata in rosso, corrisponde alla capacità unitaria scelta per il progetto, pari a 34fF, il doppio della capacità minima offerta dalla tecnologia.

Per tale valore di capacità la DNL massima resta inferiore a 0.5 anche con parassiti prossimi ai 40fF. L'impatto della DNL distribuita deterministicamente per effetto del parassita  $C_{pB}$  è stato stimato attraverso l'ausilio dei modelli *MATlab*<sup>TM</sup> realizzati.

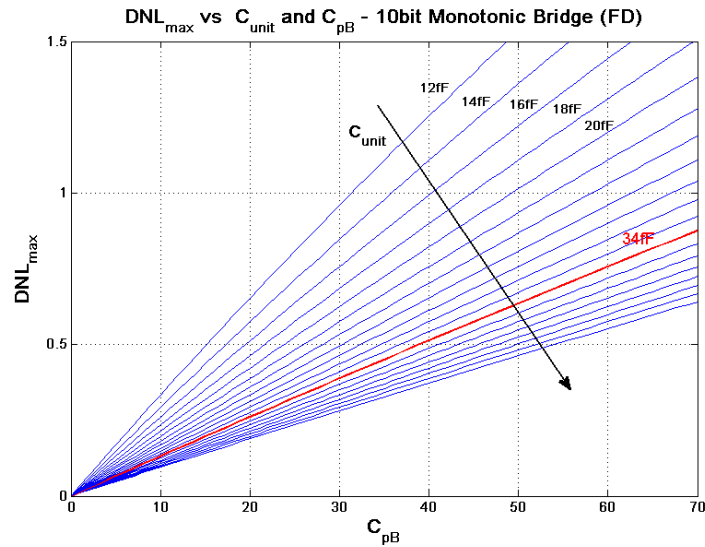


Figura 5.1: Legame tra parasita, deviazione standard della DNL e Capacità unitaria, per l'ADC SAR CR Bridge Monotonico scelto.

In figura 5.2 è mostrato l'andamento del numero di bit effettivo in funzione del picco massimo di DNL per effetto del parassita sul top plate dell'LSB.

Risulta evidente da quanto visto finora che anche non rispettando in modo rigoroso la specifica di monotonicità della caratteristica è possibile ottenere una linearità compatibile con l'applicazione. Il contributo stimato del mismatch della tecnologia per l'array progettato ( $k_c = 0.0095$ ) è riassunto negli istogrammi di figura 5.3.

La distribuzione degli istogrammi mostra come oltre il 95% delle realizzazioni abbiano un picco massimo di DNL inferiore a 0.4. Questo risultato era prevedibile dato che il valore di capacità unitaria scelto è superiore a quello ottenibile effettuando il sovradimensionamento rappresentato dalla 4.46. Gli stessi modelli, perfezionati per tenere conto contemporaneamente di mismatch e di parassiti, sono stati utilizzati per verificare a priori la distribuzione del picco massimo del valore assoluto della DNL e per valutare l'ENOB atteso (e confrontarlo successivamente con quanto ottenuto a seguito del layout). Gli istogrammi relativi all'array scelto tenendo conto contemporaneamente di mismatch e parassiti (30fF) sul top plate del LSB sono illustrati in figura 5.4. Con la capacità scelta il 60% delle realizzazioni ha un valore assoluto di DNL massimo inferiore alla media, cioè 0.58, mentre nel 99% dei casi la DNL massima non supera 1LSB.

I risultati di simulazione mostrano come i due fenomeni, parassita e mismatch, si sommino con buona approssimazione in modo lineare. Nonostante esista una probabilità del 50% che il picco del massimo valore assoluto di DNL cada oltre 0.5, l'ENOB corrispondente alle caratteristiche affette da mismatch e parassiti simulate per l'ADC scelto infatti è in media 9.57, come si evince dai risultati di simulazione di figura 5.5

Il valore medio è inferiore ai 9.8 bit predicibili in base agli effetti dei soli parassiti (fig.

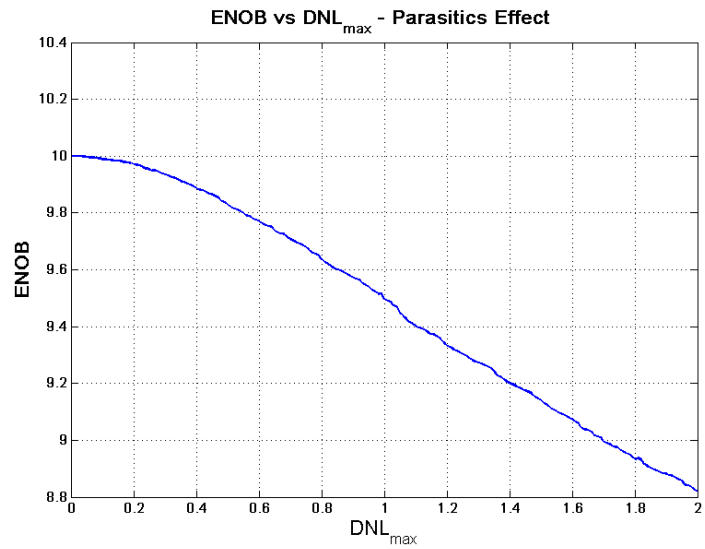


Figura 5.2: Andamento dell'ENOB in funzione del picco massimo di DNL dovuto al parassita sul top plate del LSB.

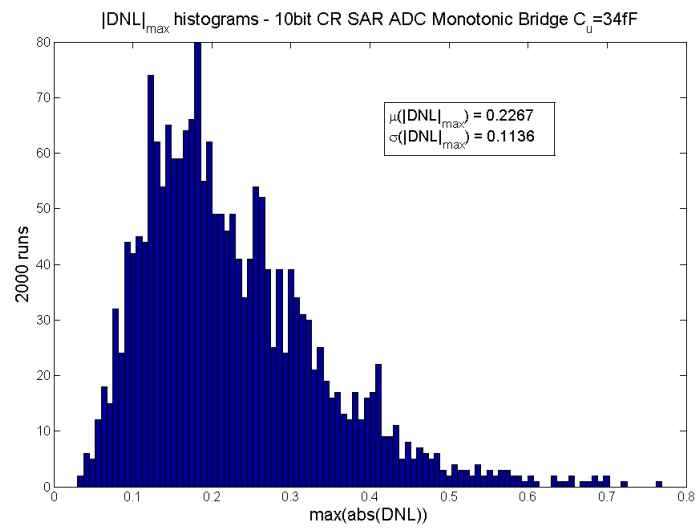


Figura 5.3: Istogrammi e parametri descrittivi dell'impatto del mismatch nell'array progettato.



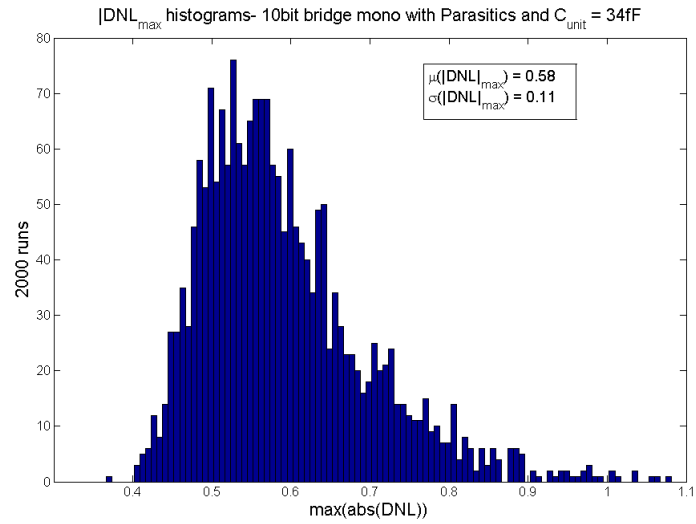


Figura 5.4: Istogrammi e parametri descrittivi dell'impatto di parassiti (ai piatti superiori) e mismatch nell'array progettato.

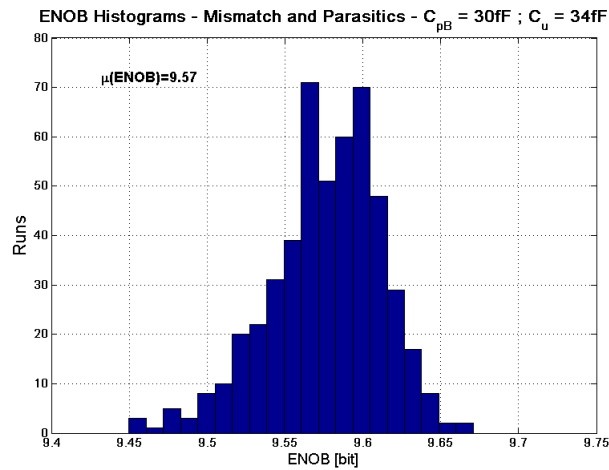


Figura 5.5: Istogrammi relativi all'ENOB stimato per l'ADC progettato con i mismatch tecnologici e parassiti di 30fF previsti ad ogni top plate dei sub-array.

5.2) ed è inferiore anche al valore di ENOB atteso per effetto del solo mismatch, pari a circa 9.75 bit con una deviazione standard di 0.11 bit. Se però si sommano gli effetti il valore appare molto più vicino alle stime ed in ogni caso è comunque accettabile perchè molto superiore agli 8 bit minimi richiesti dalle specifiche e perchè, come prevedibile dato il dimensionamento, il 100% delle realizzazioni è caratterizzata da un ENOB superiore a 9 bit.

In realtà l'associazione dell'array al comparatore e soprattutto l'effetto dei parassiti (non modellizzati in *MATLab<sup>TM</sup>*) tra top e bottom di ciascuna capacità determinerà una ulteriore diminuzione di questa cifra, Non si è pertanto ritenuto opportuno correggere al ribasso il valore di capacità unitaria, il cui dimensionamento a 34fF è stato confermato. L'ENOB atteso a seguito dell'estrazione di tutti i parassiti sarà esaminato nell'ultima sezione del capitolo, dopo aver illustrato anche le caratteristiche di logica e comparatore. La capacità totale dell'array fully differential è di 4.28pF e quella equivalente all'ingresso, cioè quella offerta al PGA, è di circa 1pF su ciascun morsetto. Ciascun blocco capacitivo dell'array è pilotato al bottom da un inverter a sua volta comandato dalla logica SAR, che vedremo nei prossimi paragrafi.

La potenza consumata dall'array con la capacità unitaria scelta si divide in tre contributi:

- potenza consumata nel ciclo di conversione;
- potenza consumata durante la fase di campionamento;
- potenza consumata per attivare gli interruttori e gli inverter che pilotano le capacità;

Il primo contributo, come visto nel precedente capitolo, si può stimare attraverso la relazione:

$$P_{conv} \approx f_s 30.5 C_u V_{DD}^2 \approx 8nW \quad (5.1)$$

In media il secondo contributo vale invece:

$$P_{samp} \approx f_s 2 C_{in} \left(\frac{V_{DD}}{2}\right)^2 \approx 4nW \quad (5.2)$$

dove  $C_{in}$  è la capacità di ingresso effettivamente caricata durante il campionamento. Attribuendo al terzo un contributo stimato di 1-2nW, il consumo atteso per l'array è di circa 13-14nW.

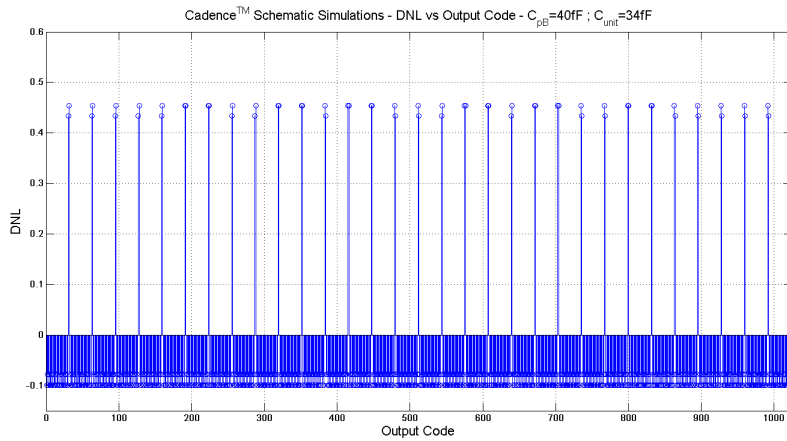
Le caratteristiche di DNL e INL ottenute dalle simulazioni in ambiente Cadence *Virtuoso<sup>TM</sup>*

Tabella 5.1: Tabella riassuntiva delle prestazioni attese per l'array ottenute da simulazioni in fase di schematico.

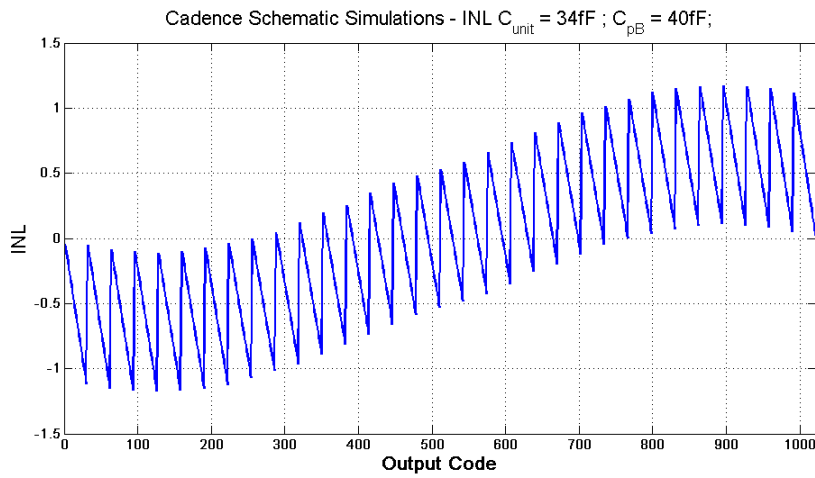
Parametri	Stime Pre-Layout
$V_{DD}$	0.5V
N bit	10
$C_u$	34fF
$C_{tot}$	4.28pF
$E_{avg}$	$\approx 260$ fJ
$P_{cons}$	$\approx 13$ nW
$ DNL _{max}$	0.4
ENOB	9.57

per l'array realizzato simulato insieme al comparatore (illustrato in seguito) e con blocchi logici modellizzati in VerilogAMS sono rappresentate in figura 5.6. La modellizzazione dei parassiti dei piatti superiori dei sub-array è stata effettuata collegando a questi ultimi delle capacità ideali, in questo caso pari a 40fF (sovrastima).

Queste simulazioni sono risultate molto lunghe e per motivi di tempo non si sono potute realizzare in modalità montecarlo allo scopo di verificare l'effetto dei mismatch. I risultati delle simulazioni su schematico in ambiente Cadence *Virtuoso<sup>TM</sup>* hanno dimostrato una ottima coerenza con i risultati delle simulazioni *Matlab<sup>TM</sup>*. In realtà, come vedremo nella sezione dedicata ai risultati, la mancata modellizzazione dei parassiti tra top plate e bottom plate delle varie capacità si tradurrà in un certo scostamento. A meno di questa incongruenza, che sarà oggetto di futuri perfezionamenti, il confronto tra le simulazioni MATLAB e le simulazioni Cadence in fase di schematico suggerisce una buona accuratezza dei modelli realizzati. Per questo motivo sono stati utilizzati anche per stimare l'effetto del mismatch. Una delle realizzazioni di caratteristiche ottenute con i modelli implementati, comprensive di mismatch e parassiti dai top plate verso massa, è illustrata in figura 5.7. I risultati delle simulazioni pre-layout e dei calcoli effettuati per la stima delle prestazioni relative all'array di capacità sono riassunti nella tabella 5.1.

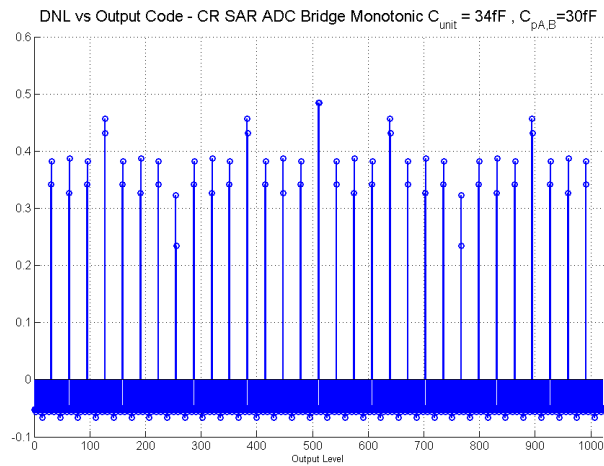


(a) Singola realizzazione di DNL dell'ADC realizzato simulata in *Cadence™*.

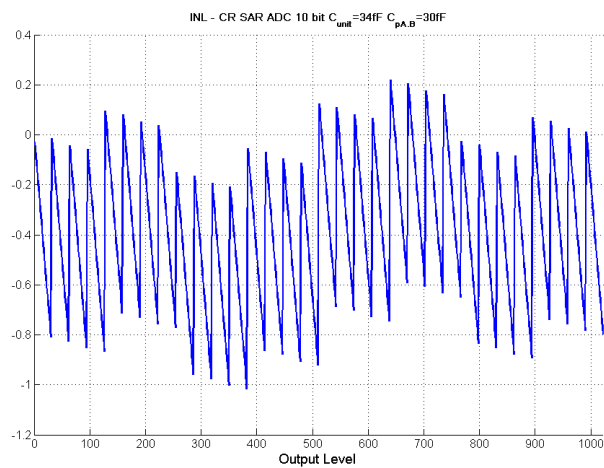


(b) Caratteristica INL dell'ADC realizzato simulata in *CadenceVirtuoso™*.

Figura 5.6: INL e DNL simulate in *CadenceVirtuoso™* per l'ADC SAR realizzato. Array e comparatore sono stati simulati in modalità schematico tenendo conto di parassiti modellizzati con capacità ideali di 40fF ai top plate dei subarray. La logica è stata implementata in linguaggio Verilog-A.



(a) Singola realizzazione di DNL dell'ADC realizzato simulata in MATLAB.



(b) Caratteristica INL dell'ADC realizzato simulata in MATLAB.

Figura 5.7: Singola realizzazione di INL simulata in MATLAB per l'ADC SAR realizzato.

### 5.1.1 Interruttori di ingresso

Il campionamento del segnale all'ingresso del convertitore avviene direttamente sul top-plate del sub-array relativo al MSB che offre al circuito a monte, cioè il PGA, una capacità totale pari a  $2^{\frac{N}{2}}$  capacità unitarie per ogni nodo d'uscita. Per un'accurata memorizzazione della tensione di segnale all'ingresso sono stati utilizzati due switch con circuito di bootstrap. Durante questa fase, infatti, la capacità totale del subarray è vista praticamente in serie alla resistenza dell'interruttore di top e la rapidità di completamento del campionamento è determinata dalla costante di tempo associata. Contemporaneamente, in fase di off, cioè durante il ciclo di conversione, lo stesso interruttore deve offrire una resistenza molto elevata, tale che nessuna corrente di leakage sia sufficiente a ad alterare di uno o più LSB la tensione memorizzata sull'array. Questo significa imporre almeno che:

$$\frac{I_{leak}T_{hold}}{C_{TOT_{sub}}} \leq \frac{LSB}{2} \quad (5.3)$$

cioè, considerando il caso peggiore in cui la differenza ai capi dell'interruttore aperto è di  $V_{DD}$  e con un tempo di hold di circa  $15\mu s$ , si ha:

$$R_{OFF} \geq \frac{2^{N+1}T_{hold}}{C_{TOT_{sub}}} \approx 15T\Omega \quad (5.4)$$

L'unico modo per ottenere resistenze così elevate è utilizzare transistori lunghi facendo in modo che essi siano polarizzati, in fase di off, sotto soglia, analogamente a pseudoresistori. I transistori della tecnologia a disposizione dimensionati con questo criterio e pilotati con tensioni non superiori a quelle di alimentazione ( $0.5V$ ) non sarebbero in grado di garantire resistenze di accensione sufficientemente piccole per agganciare il segnale nei tempi limite e memorizzarne il valore con accuratezza sufficiente. Si tenga conto che con il tempo di campionamento disponibile per l'applicazione il tempo da dedicare alla fase di sampling dovrebbe essere al massimo di  $10 - 15\mu s$  (la metà del totale) e la capacità totale del subarray non dovrebbe superare i  $2pF$ . Perché il segnale sia campionato con un errore inferiore alla metà del LSB dev'essere rispettata la relazione [32]:

$$\tau \ln\left(\frac{2V_{DD}}{LSB}\right) = \tau \ln(2^{N+1}) \leq 10\mu s \quad (5.5)$$

ed essendo la costante di tempo  $\tau$  pari a:

$$\tau = C_{TOT_{sub}}R_{ON} \quad (5.6)$$

si ha:

$$R_{ON} \leq \frac{10\mu s}{C_{TOTsub} \ln(2^{N+1})} \approx 1.3M\Omega \quad (5.7)$$

Il rapporto tra  $R_{OFF}$  e  $R_{ON}$  richiesto è di oltre  $10^7$  ed è ottenibile solo utilizzando una tecnica di bootstrap. L'architettura scelta per gli interruttori bootstrappati al top plate è simile a quella presentata in [67] ed illustrata in figura 5.8.

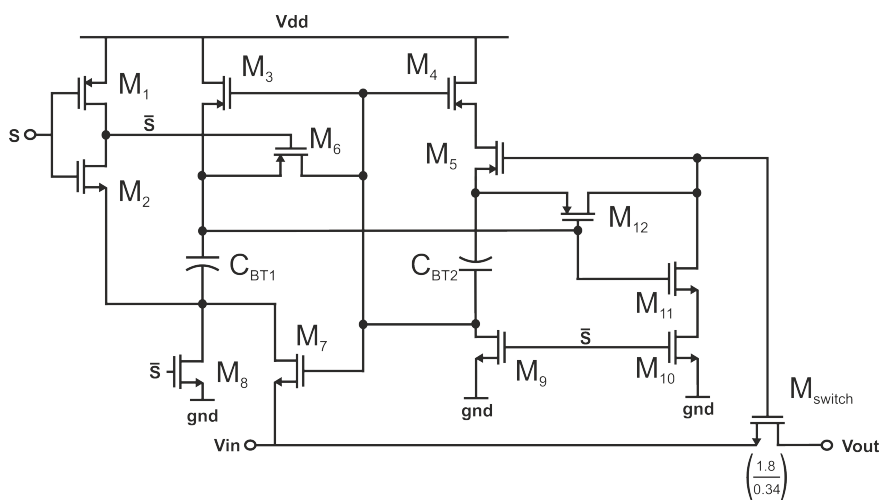


Figura 5.8: Architettura dello switch con circuito di bootstrap.

Rispetto al circuito in [67] questo interruttore non utilizza tecniche di *rise time acceleration* perchè la rapidità di accensione di questo tipo di circuito non è un particolare problema dell'applicazione. Il funzionamento del circuito si divide in due fasi corrispondenti a due diversi valori, alto o basso, del segnale di SAMPLE (S) e del suo negato. Quando il segnale di SAMPLE è basso, i MOS M3, M4, M5, M9 e M8 sono accesi e permettono la carica di entrambe le capacità di bootstrap. Nella fase successiva questi si spengono mentre M2, M7, M12 e M11 si accendono e pongono in serie tra  $V_{in}$  e il gate del MOS  $M_{switch}$  le due capacità. In tal modo è virtualmente possibile imporre  $2V_{DD}$  tra il gate e il source del MOS che funziona da vero e proprio interruttore di campionamento. In realtà a causa dei parassiti capacitivi si verifica una certa redistribuzione di carica quando il segnale SAMPLE è alto per cui la tensione  $V_{GS}$  effettivamente imposta è lievemente inferiore, ma comunque sufficiente a garantire l'adeguata accensione. Scegliendo infatti le capacità  $C_{BT1}$  e  $C_{BT2}$  di circa 300fF ciascuna si ottiene una accensione soddisfacente con una  $R_{ON}$  di circa 10k $\Omega$ . Ciò significa, per un array la cui capacità di vista all'ingresso è di 1.088pF, un tempo di settling (necessario a agganciare il segnale entro la metà di un LSB) durante il campionamento pari a circa:

$$t_{sample_{sett}} = R_{ON}C_{TSE1} \ln(2^{(N+1)}) \approx 80ns \quad (5.8)$$

In realtà questo tempo sarà superiore perchè l'interruttore di bootstrap impiega un certo tempo per l'accensione (circa 120ns) durante il quale la resistenza si riduce progressivamente fino al valore minimo. Le simulazioni mostrano in ogni caso un tempo di settling dell'array durante il campionamento di circa 250ns e con il caso peggiore osservato in simulazioni di corner con dispositivi meno conduttivi di circa 400ns ( $R_{ON}$  di circa  $50k\Omega$ ). Questi tempi sono compatibili con la velocità richiesta dall'applicazione e anche con velocità molto superiori. In figura 5.9 è illustrato l'andamento del segnale SAMPLE (S), del suo negato e della tensione  $V_{GS}$  del MOS  $M_{switch}$  utilizzato come interruttore durante il funzionamento del convertitore alla frequenza di 30kSps.

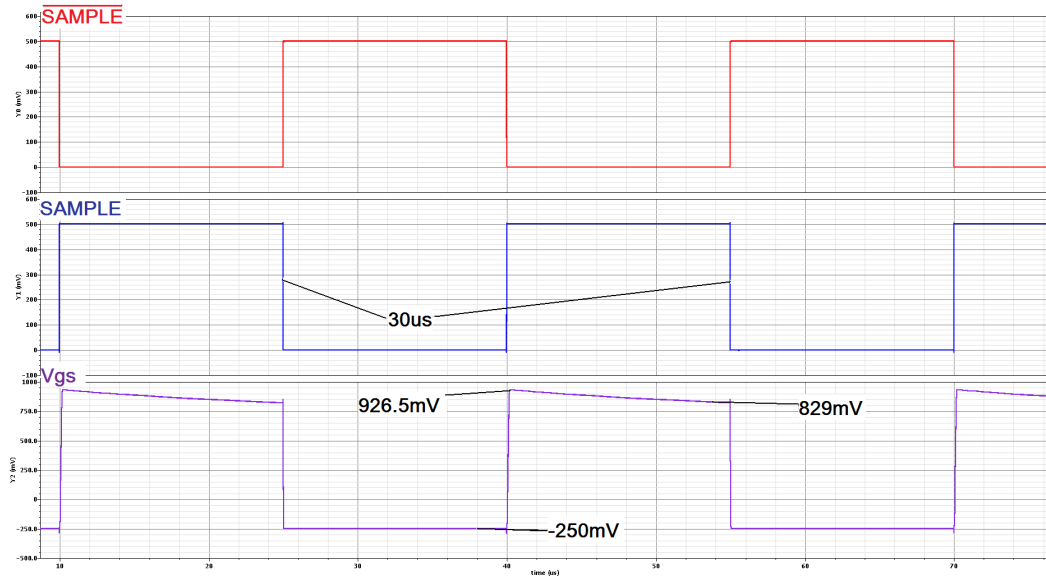


Figura 5.9: Forme d'onda di interesse per il funzionamento dello switch bootstrappato.

Il settling time dell'array tra una configurazione e la successiva durante il ciclo di conversione costituisce un'altra specifica. Esso è stato simulato secondo parametri tipici ed è di circa 15ns mentre attraverso simulazioni di corner è stato possibile osservare una sua possibile variazione fino a 50ns. Anche nel caso pessimo i numeri sono compatibili con l'intervallo di tempo disponibile per ciascuna comparazione alla frequenza di campionamento dell'applicazione, cioè circa  $1.6\mu s$ .

Non è stato necessario scegliere particolari circuiti di bootstrap per gli interruttori posti ai bottom delle capacità perchè a quei nodi afferisce sempre una bassa impedenza verso un riferimento di tensione (c'è un inverter) e pertanto ciascuno di essi può essere liberamente dimensionato per ottenere resistenze di accensione molto inferiori a quelle



dell'interruttore di ingresso e senza preoccuparsi del valore della resistenza in fase di spegnimento. Come vedremo anche nella sezione dedicata alla logica di pilotaggio, al posto della complessa rete di interruttori che si avrebbe se si utilizzasse una procedura di conversione classica, sono stati utilizzati dei semplici inverter. Questa semplicità è permessa dal campionamento sul nodo di top (e non di bottom) tipica della procedura monotonica. Successivamente all'introduzione dei dettagli progettuali di logica e comparatore, saranno illustrate le caratteristiche del convertitore definitivo ottenute da simulazioni successive alla fase di layout ed all'estrazione dei parassiti.

## 5.2 Il Comparatore

Questa sezione è dedicata alla scelta al progetto ed al dimensionamento del comparatore utilizzato per il convertitore SAR.

### 5.2.1 Funzionamento

Il comparatore è un circuito non lineare che in funzione del segno della differenza tra due segnali ai due nodi positivo e negativo di ingresso produce un gradino d'uscita che va al minimo o al massimo della dinamica. Questi due livelli di tensione corrispondono a equivalenti valori logici binari e nel contesto di un convertitore SAR costituiscono i possibili bit di codifica del segnale campionato. In generale un comparatore può essere di tipo statico oppure di tipo dinamico. La prima categoria è costituita tipicamente da comparatori a doppio stadio in cui il primo stadio funziona da preamplificatore, mentre il secondo stadio è un latch in grado di scattare ad un rail o all'altro a seconda del segnale in uscita al preamplificatore. La funzione del preamplificatore in questo tipo di comparatori è di ridurre al minimo le probabilità di metastabilità anche in corrispondenza di ingressi molto vicini tra di loro. In genere esso necessita di un certo tempo di settling per cui il latch viene attivato con un certo ritardo rispetto all'istante in cui gli ingressi sono disponibili all'ingresso del comparatore [70] [72]. Tralasciando la complessità dell'architettura e del coordinamento degli interruttori di set e reset, il principale inconveniente della struttura è la corrente continua che il preamplificatore richiede in polarizzazione, tanto più elevata quanto più alti sono i requisiti di velocità e di rumore. Una alternativa più efficiente è rappresentata dai comparatori dinamici, generalmente costituiti da un latch o una cascata di latch. La possibilità di utilizzare meccanismi di power-on per accendere il circuito solo in corrispondenza degli istanti in cui il segnale all'ingresso è disponibile rende questa categoria di comparatori molto più attraenti dal punto di vista energetico rispetto a quelli statici. Al contrario della controparte statica, le architetture dinamiche sono piuttosto soggette a disturbi come ad esempio quelli dovuti all'accensione e allo spegnimento e il dimensionamento dei componenti per la minimizzazione del rumore è più complesso. L'algoritmo monotonic scelto per la conversione del segnale nel convertitore progettato prevede la riduzione del modo comune dei segnali di ingresso ad ogni passo del ciclo di conversione fino a valori vicini allo zero. Il comparatore da progettare deve avere dinamica di modo comune che si estende fino al rail inferiore o comunque deve essere in grado di funzionare anche per differenze minime tra segnali entrambi molto vicini allo zero. I requisiti di velocità sono poco stringenti. Per lavorare alla frequenza di circa 30kHz con 10bit significa che la conversione di ogni bit ha a disposizione un intervallo temporale di circa  $1.6\mu\text{s}$ . La commutazione del comparatore dovrebbe impiegare il minimo tempo possibile. Un buon risultato sarebbe ottenerlo inferiore al 10% di quello disponibile, cioè circa 160ns. Coerentemente all'obiettivo di realizzare un convertitore efficiente abbiamo scelto di adottare una topologia dinamica dello stesso genere di quelle presentate in [69] e in [61] ma opportunamente modificata per soddisfare i requisiti richiesti ed illustrata in figura 5.10 con i relativi dimensionamenti dei dispositivi.

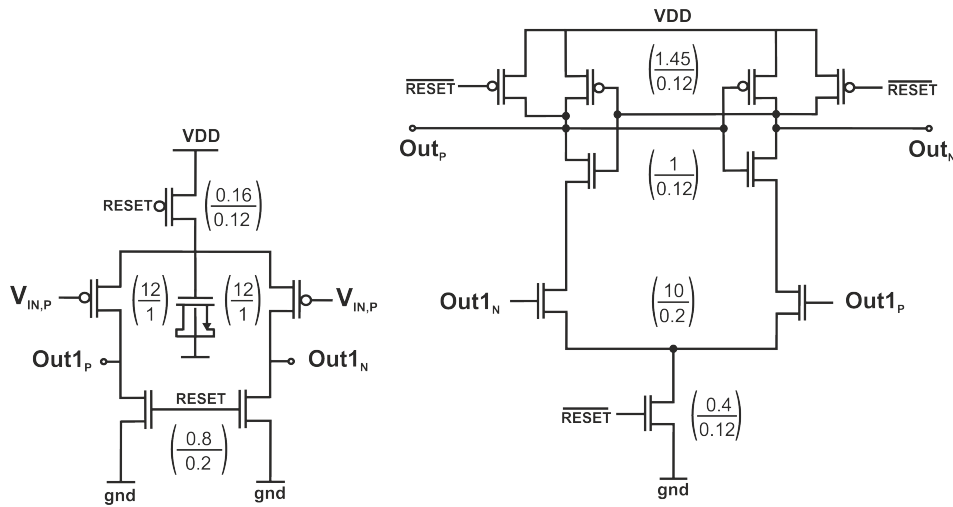
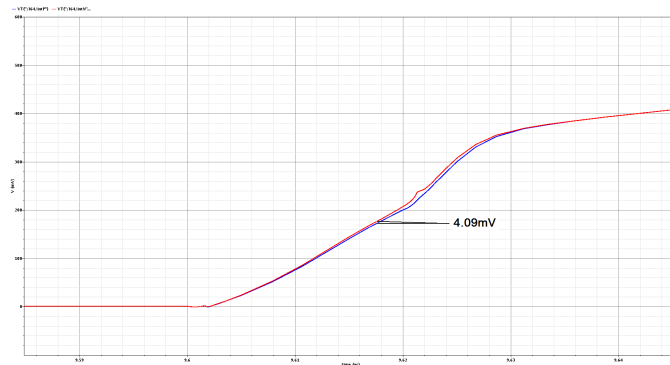


Figura 5.10: Comparatore progettato.

Quando il valore della tensione ai nodi di top dei sub array è stabilizzata e il segnale differenza tra i due array è pronto per essere valutato (fine fase di sample), il segnale di reset transisce da alto a basso attivando simultaneamente primo e secondo stadio del comparatore. Analogamente a quanto spiegato in [61], il primo stadio funziona come un amplificatore a basso guadagno che in base al segnale presente al gate di ciascuno dei due MOS di ingresso smista in proporzioni differenti la corrente di coda caricando con velocità differenti le capacità parassite ai nodi di uscita, cui contribuiscono i gate d'ingresso del latch. Ciò che si osserva in corrispondenza di ogni comparazione sono due differenti rampe di salita, come illustrato in figura 5.11

Anche il secondo stadio, cioè il latch, è potenzialmente pronto a comparare fin dall'istante successivo alla transizione del RESET ma in realtà ciò non avviene finché il modo comune della rampe di tensione ai due ingressi non raggiunge una tensione pari a qualche tensione termica al di sotto della soglia dei MOS di ingresso del secondo stadio. Prima del raggiungimento di questo valore di soglia, infatti, il gate dei MOS di ingresso è a tensione troppo bassa e i transistori sono praticamente spenti. All'istante di raggiungimento della soglia le rampe sono distanziate di un valore da 3 a 6 volte superiore rispetto a quello d'ingresso anche quando questo è dell'ordine di un LSB. Il latch commuta a seguito della differenza di tensione tra le uscite del primo stadio e le sue uscite partono da una situazione di reset in cui sono vincolate all'alimentazione. All'accensione, esse si portano a una tensione di circa 420mV dalla quale si allontanano successivamente in seguito alla commutazione, come illustrato in figura 5.12.

In serie alle uscite sono posti due inverter con soglia 200mV (50mV sotto la metà della dinamica) per evitare commutazioni spurie dovute al transitorio decrescente delle uscite che precede la commutazione. La commutazione avviene in circa 15ns (fig. 5.12), tempo più che adatto all'applicazione e che può garantire, logica permettendo, il funzionamento a frequenze superiori rispetto a quelle richieste dall'applicazione. Considerando



(a) Caratteristiche rampe all'uscita del primo stadio del comparatore in corrispondenza di una tensione differenziale di ingresso pari a un LSB.



(b) Dettaglio del valore delle rampe in corrispondenza della soglia di innesco del latch.

Figura 5.11: Caratteristiche del segnale di uscita al primo stadio del comparatore durante la commutazione.

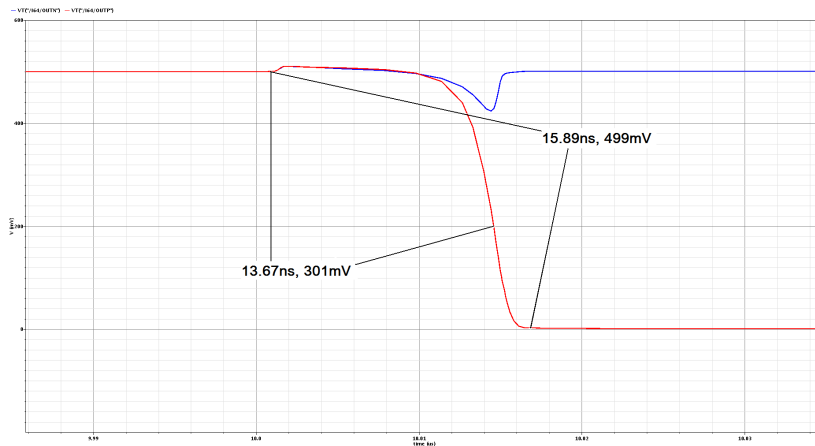


Figura 5.12: Commutazione del comparatore.

il settling time dell'array pari a circa 25ns e altri 100ns di margine, queste potrebbero ragionevolmente essere superiori a 1Msps. Il consumo del comparatore alla frequenza di 30kSps è di circa 12.5nW.

### 5.2.2 Sensibilità ai disturbi e al rumore

Per non degradare troppo le prestazioni di linearità del convertitore oltre quanto già inevitabilmente imposto dalle non idealità dell'array, è bene che il comparatore abbia adeguate prestazioni di rumore e sia poco sensibile ai disturbi, che in questo caso sono rappresentati dalle escursioni di tensione dovute al set-reset. Inoltre, in un comparatore di questo tipo inserito in un SAR CR ADC monotono la variazione del modo comune di ingresso potrebbe indurre una variazione dell'offset potenzialmente devastante per la linearità del convertitore. In generale un offset costante durante il ciclo di conversione non rappresenta un problema perchè può essere corretto con opportune configurazioni circuitali, con procedure di correzione, oppure, come in questo caso, può essere corretto durante il post-processing off-chip del segnale. Se tuttavia l'offset varia allora la codifica può risultare erronea e la qualità del segnale degrada. La discussione circa l'idoneità del comparatore progettato rispetto a tutte queste problematiche è affrontata in questo paragrafo. L'analisi del rumore in latch dinamici con un primo stadio di questo tipo è ampiamente affrontata in [61], pertanto è qui presentata una analisi impostata in modo analogo ma con gli opportuni adeguamenti al circuito in esame. In base a quanto spiegato circa il funzionamento del primo stadio del comparatore il suo guadagno equivalente è di circa 3-6, per cui si ha almeno un fattore pari a circa 10 di riduzione del contributo alla densità spettrale di rumore del secondo stadio, che può essere trascurato. E' dunque il primo stadio a fornire un contributo dominante al rumore e ciò che avviene ad ogni comparazione è l'integrazione della densità spettrale di rumore in corrente all'uscita sulle capacità parassite ai relativi nodi di uscita. Il rumore termico dei MOS di ingresso,

assumendo che essi lavorino in debole inversione come praticamente ogni altro MOS del front-end, si può scrivere nel seguente modo:

$$\bar{I}_{n,out}^2 \approx 2qI_D \quad (5.9)$$

dove con  $I_D$  si indica la corrente che scorre in ciascun ramo dello stadio di ingresso. Questo rumore, proveniente dai quattro MOS di ingresso, è integrato sulla capacità parassita  $C_L$  al nodo, pertanto la densità spettrale di rumore in uscita al primo stadio è:

$$\bar{E}_{n,out}^2 \approx 8qI_D \frac{1}{(2\pi\omega)^2} \quad (5.10)$$

Tale rumore è integrato fino alle più alte frequenze a partire da una banda equivalente BW pari all'inverso del tempo di integrazione. Infatti, tra ogni comparazione e la successiva, le capacità parassite vengono scaricate e non esiste alcuna memoria del segnale; si ha solo l'integrazione del rumore nell'intervallo di tempo in cui il comparatore è acceso per la singola comparazione. Ciò equivale a filtrare le componenti a più bassa frequenza. Si ha quindi:

$$V_{n,rms}^2 = \sqrt{\int_{BW}^{\infty} 8qI_D \frac{1}{(2\pi\omega C_L)^2} d\omega} = \frac{8qI_{MOS}T_{int}}{4\pi^2 C_L^2} \quad (5.11)$$

Il tempo di integrazione dipende dalla soglia al quale il secondo stadio inizia la commutazione e coincide con il ritardo con il quale è raggiunta dalla tensione di modo comune delle uscite del primo stadio. Dalle simulazioni si osserva che questa è prossima ai 180mV ma per il calcolo in corso si può adottare il valore della tensione di soglia  $V_{TH}$  dei MOS, pari 274mV. Ciò rappresenta una stima peggiorativa ed in ogni caso si ha:

$$T_{int} = \frac{C_L V_{TH}}{I_D} \quad (5.12)$$

Sostituendo l'ultima relazione nella 5.11 si osserva che questo non dipende dalla corrente che scorre in ciascun MOS. Invertendo la relazione in funzione di  $C_L$  per avere un rumore equivalente in ingresso inferiore alla metà del LSB ( $476\mu V$ ) si ottiene:

$$C_L \geq \frac{8qV_{TH}}{G_{eq}4\pi^2(476\mu V)^2} \approx 9.75 fF. \quad (5.13)$$

Questo valore è inferiore alla capacità parassita simulata nello schematico (circa 25fF) con il dimensionamento dei transistori già mostrato in figura 5.10, pertanto la specifica si ritiene soddisfatta.

Un'altra categoria di disturbi rispetto ai quali lo stadio deve dimostrarsi insensibile è legata al set e al reset del comparatore. Questa procedura sottopone il nodo di source dei MOS di ingresso ad escursioni di tensione di diverse centinaia di mV. A causa dell'accoppiamento capacitivo con i nodi di ingresso rappresentato dalla capacità  $C_{gs}$  dei MOS, queste variazioni possono riportarsi all'ingresso alterando il segnale prima di ogni comparazione. Una variazione di tensione al nodo di source può ripartirsi diversamente tra i due nodi di ingresso solo in ragione di differenze significative della capacità  $C_{gs}$  dei due rami oppure della capacità vista verso massa dal top plate dell'array in ingresso. Queste ultime dipendono dal mismatch e dai parassiti, elementi che si mantengono costanti nell'integrato e pertanto non introducono effetti non lineari bensì solo un offset correggibile nel post-processing. La variazione di capacità  $C_{gs}$  che si può verificare quando uno dei due transistori si trova in debole inversione mentre l'altro si trova in forte inversione, al contrario, può dare origine a effetti non lineari. Il rapporto H di attenuazione con il quale le escursioni di power-on e power-off al source alterano il segnale differenziale presente all'ingresso è:

$$H = \frac{\Delta C_{gs}}{C_{TSE1} + C_{gs}} \quad (5.14)$$

La capacità totale equivalente vista dal top plate del sub-array del MSB  $C_{TSE1}$  è stata definita nel precedente capitolo. Si tenga presente che le variazioni di capacità in questione dipendono dal regime di funzionamento dei MOS (sottosoglia, debole o forte inversione) e quindi dalla tensione  $V_{GS}$ . Se l'alterazione della tensione di ingresso è grande, allora significa che anche il segnale differenziale all'ingresso è grande, più grande del disturbo (che ritorna attenuato di H) e quindi non c'è rischio di errore nella singola comparazione. Può invece sussistere il rischio di errore a cavallo di più cicli perchè di volta le tensioni all'ingresso si avvicinano e le differenze di capacità tra gate e source tendono allo zero, aumentando il peso di eventuali errori compiuti durante la comparazione del MSB. Le simulazioni in realtà mostrano che le variazioni di  $C_{gs}$  nella configurazione in questione sono ridotte e non evidenziano un impatto realmente pericoloso sul segnale.

In prima approssimazione quindi si ha che la  $\Delta V_{GS}$  coincide con la differenza di segnale in ingresso, per cui tra questa e la variazione del segnale indotta dal power-on c'è sempre un rapporto inferiore a uno e non vi sono ripercussioni sulla qualità del segnale. In ogni caso al nodo di source è stata posta una capacità MIM di circa 50fF per cercare di ridurre l'escursione di tensione al nodo. Essa non incrementa il consumo del circuito perchè successivamente alla fase di accensione, quando il circuito va a regime, la capacità conserva tra un ciclo e il successivo la maggior parte della carica accumulata. Le simulazioni in *Cadence<sup>TM</sup>* durante l'accensione confermano che le variazioni della capacità  $C_{GS}$  non sono mai sufficienti a generare errori superiori al LSB.

L'offset del comparatore risulta dalle simulazioni MonteCarlo circa pari a 10mV e nonostante la riduzione della tensione di modo comune dei morsetti di ingresso la sua variazione non è mai superiore al valore del LSB.



## 5.3 La Logica SAR

In questa sezione è illustrato lo schema generale della logica utilizzata per la gestione del convertitore, i dettagli progettuali dei suoi principali componenti e le sequenze con i quali questi si attivano in modo coordinato per implementare l'algoritmo di switching monotonicamente desiderato.

### 5.3.1 Caratteristiche Generali e Funzionamento

La logica di ADC SAR CR è tipicamente costituita da elementi dinamici (che commutano in corrispondenza dei fronti di un clock esterno) come dei registri o dei flip-flop, disposti su due file, secondo lo schema generale di figura 5.13 (esempio a 9 bit). La prima fila di clock serve a scandire ad ogni fronte di clock la sequenza di switching delle capacità del convertitore, attuata dalla seconda fila di flip-flop utilizzando l'informazione fornita di volta in volta dall'uscita del comparatore.

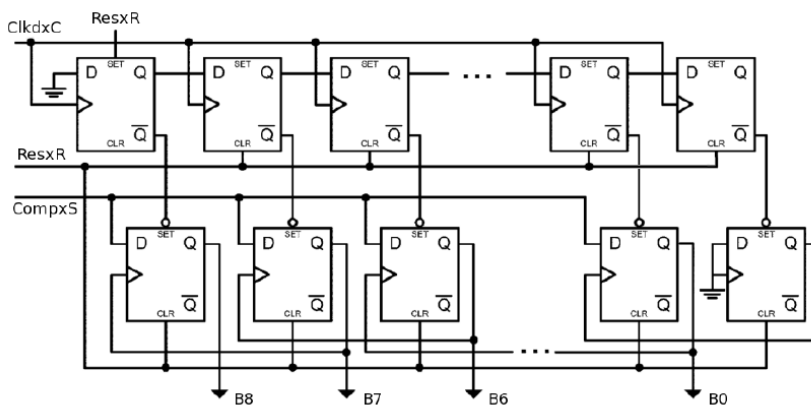


Figura 5.13: Logica SAR classica (esempio a 9 bit).

A seconda del bit che quest'ultimo genera la logica connette a massa oppure ad alimentazione il bottom plate della capacità corrispondente al bit appena codificato (caso single ended). La necessità di utilizzare un clock esterno ad una frequenza relativamente alta, pari almeno a  $N$  (bit) volte la frequenza di campionamento è penalizzante dal punto di vista del consumo. Il segnale di sample e i fronti necessari ad attivare le singole comparazioni utili all'applicazione sono a frequenza relativamente bassa e potrebbero essere generato da una divisione di un clock più veloce. Quest'ultimo lavorerebbe continuamente, anche durante la fase di sampling, ed ha un consumo costante causato sia dalle commutazioni utili al funzionamento del comparatore, sia a molte altre commutazioni delle sue uscite (e di eventuali altri componenti) che non hanno alcuna utilità logica. Per questo motivo il convertitore è stato realizzato con una logica asincrona in grado di attuare in modo automatizzato la sequenza di commutazioni necessarie alla conversione in seguito ad ogni fronte del segnale di sampling (questo deve necessariamente provenire

dall'esterno). Data la relativa semplicità della logica, la ristretta varietà di componenti e la volontà di ridurre al minimo possibile i consumi, è stato abbandonato il progetto con celle standard, di cui peraltro la tecnologia forniva gli schematici ma non i dettagli di dimensionamento nè il layout, a favore di un approccio full-custom. Questo, come vedremo, ha permesso di ridurre al minimo il numero di gate soggetti a commutazioni nel ciclo di conversione e di ridurre l'area occupata e i consumi complessivi. Ogni componente o gruppo di componenti progettato sarà illustrato in un paragrafo dedicato. Lo schema generale della logica implementata per il convertitore è illustrato in figura 5.14 insieme alla legenda dei segnali di interesse che verrà convenzionalmente utilizzata in questa sezione per descrivere il funzionamento dei circuiti.

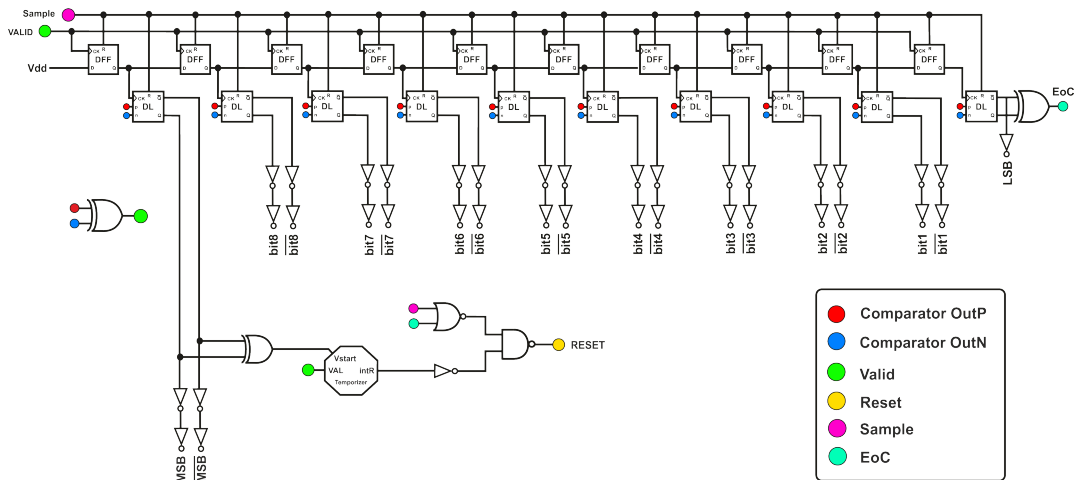


Figura 5.14: Logica SAR monotonica realizzata.

Il segnale VALID (in verde) è pari al livello logico 1 (0.5V) ogniqualvolta il comparatore termina la valutazione degli ingressi, altrimenti è pari a 0. Ad ogni fronte positivo del VALID un blocco logico, che chiameremo temporizzatore, progettato specificatamente ed illustrato nel prossimo paragrafo, genera un impulso rettangolare che agisce come clock locale con un ritardo ed una durata regolabili dal progettista. L'uscita del temporizzatore (intR) entra in una semplice rete logica per la generazione del segnale di SET/RESET del comparatore. L'utilità di questa configurazione è garantire il completamento degli eventi necessari ad ogni passo della conversione. Quando il VALID transisce da 0 a 1, infatti, con un certo ritardo il temporizzatore induce il suo RESET per un tempo prestabilito e sufficiente al completamento di una serie di eventi prima della successiva comparazione. In corrispondenza della salita del VALID, del corrispondente RESET del comparatore e negli istanti appena successivi, anche la fila di flip-flop dinamici è sollecitata. Inizialmente le uscite di tutti in flip-flop della fila sono a 0 e solo l'ingresso del primo è alto. Ad ogni VALID, ciascun flip-flop trasferisce il segnale d'ingresso fino all'uscita. In questo modo si ha l'avanzamento di unità in unità lungo la fila del valore logico 1.

La fila inferiore è costituita da latch dinamici con il pin di clock connesso alle uscite dei corrispondenti flip-flop della fila superiore e gli ingressi direttamente connessi alle uscite del comparatore. Con l'avanzare del valore logico 1 lungo la fila di flip-flop, è di volta in volta attivato il latch che riceve al pin di clock il fronte positivo e sulla base della posizione delle uscite del comparatore commuta e impone la corretta configurazione ai bottom delle capacità dell'array connesse in uscita. Naturalmente le uscite del comparatore devono ancora essere ai valori logici conseguenti alla commutazione pertanto il RESET del comparatore deve avvenire successivamente alla commutazione del latch. Inoltre, una volta resettato, occorre che questo resti inattivo per un tempo sufficiente al settling dell'array (almeno 15ns, ma è opportuno offrire del margine). Il meccanismo si attiva ogni volta che il segnale esterno SAMPLE transisce verso lo 0, altrimenti il temporizzatore è fermo, ed il comparatore e i flip-flop sono resettati; quando a seguito del completamento della conversione il segnale di EoC posto in fondo alla fila di flip-flop transisce da 0 a 1, il ciclo è terminato e il comparatore è resettato in attesa del campionamento successivo. L'unica eccezione è rappresentata dalla prima comparazione. In quel momento il temporizzatore non è ancora entrato a regime e una opportuna rete logica costituita da una XOR permette l'attivazione del temporizzatore e il reset del comparatore solo dopo che la prima coppia di capacità (MSB) è stata commutata.

### 5.3.2 Il Temporizzatore

In figura 5.15 è mostrata l'architettura del circuito responsabile della corretta sequenzialità delle commutazioni logiche nel convertitore.

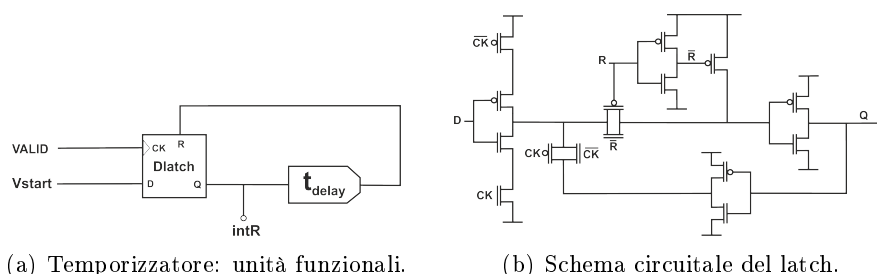


Figura 5.15: Struttura del temporizzatore.

La struttura è semplice ed è costituita da un latch dinamico retroazionato attraverso un anello in cui è posto un elemento di ritardo selezionabile dal progettista. Il funzionamento del circuito è schematizzato e spiegato in figura 5.16

Il temporizzatore garantisce sia il tempo di commutazione ai latch della seconda fila per memorizzare i bit delle comparazioni prima che il comparatore sia spento, sia il tempo di settling dell'array prima della comparazione conseguente. Dati i tempi di commutazione del comparatore e il tempo di settling dell'array di circa 15ns è stato scelto di utilizzare una unità di ritardo (catena di inverter) di circa 20ns. In questo modo il tempo disponibile alla memorizzazione del dato per i latch dinamici differenziali

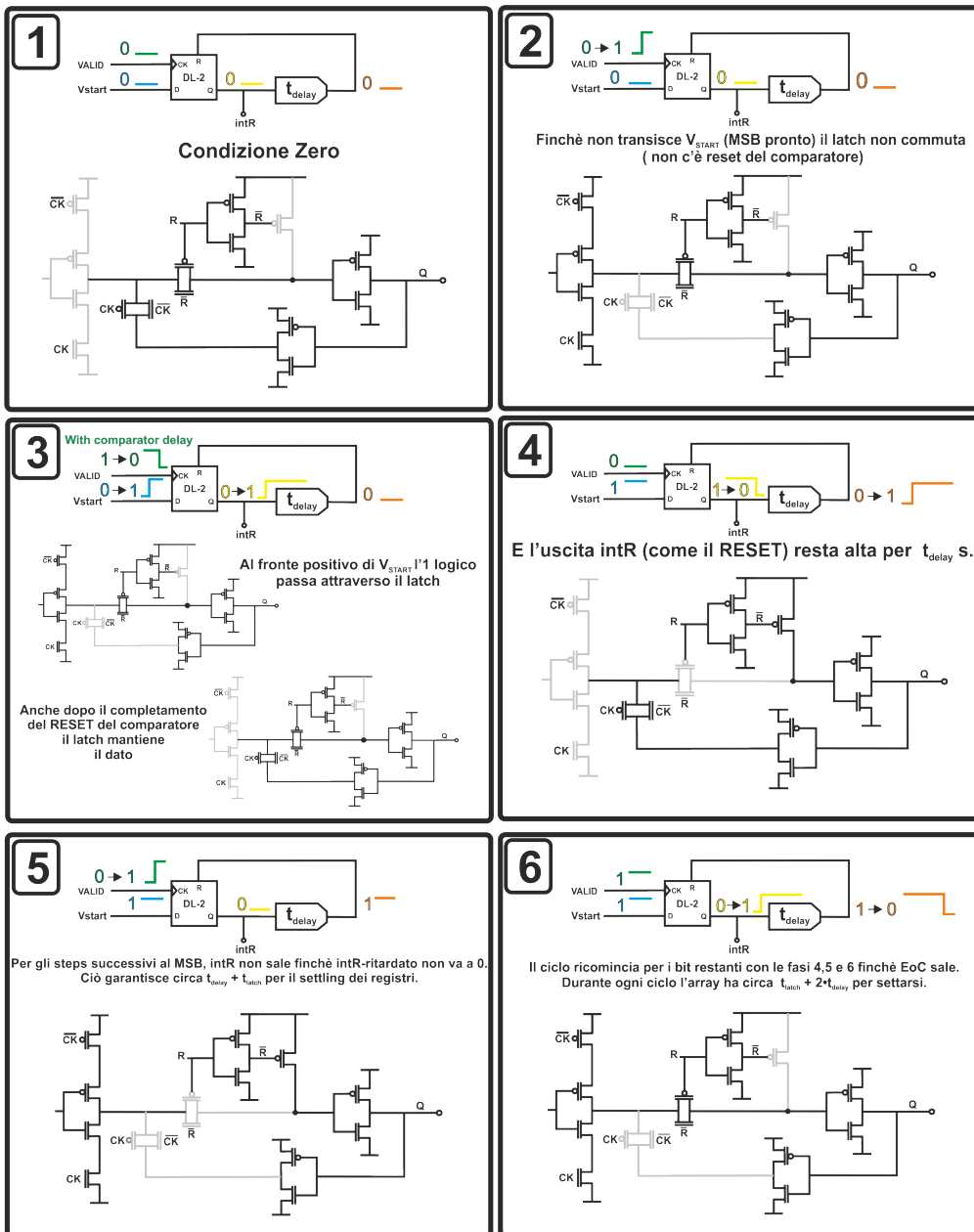


Figura 5.16: Schema di funzionamento del temporizzatore.

è poco superiore a 20ns (ampiamente sufficiente) e il tempo di settling offerto all'array è di oltre 40ns. In corrispondenza della valutazione del MSB, quando il ciclo non è ancora avviato a regime, è il segnale di inizio  $V_{start}$  ad innescare il temporizzatore subito dopo l'avvenuta memorizzazione del bit. In figura 5.17 è illustrata la sequenza di attivazione dei segnali intR, VALID e RESET durante il ciclo di conversione, mentre in figura 5.18 si può osservare il timing complessivo dei segnali logici del convertitore.

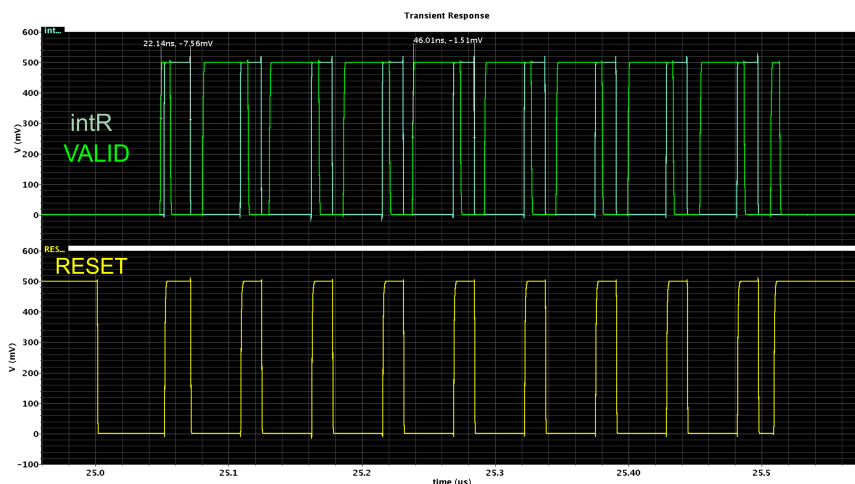


Figura 5.17: Timing diagram dei segnali logici del convertitore durante un ciclo di conversione.

Come si può osservare, l'intero ciclo di conversione avviene entro 510ns ed è inferiore ad un trentesimo del tempo disponibile con un campionamento a 30kHz e ipotizzando una suddivisione del 50% tra tempo di sampling e tempo di hold (o conversione). Così come è progettato il convertitore potrebbe quindi, potenzialmente, lavorare fino a 1Mps.

### 5.3.3 Flip-Flop e Latch Dinamici

La necessità di minimizzare consumo ed area anche per il blocco logico del convertitore ha portato ad una progettazione full-custom dei componenti con lo svantaggio di un layout più oneroso. La struttura del temporizzatore è stata già esaminata nel precedente paragrafo, mentre gli schemi circuitali dei flip-flop dinamici della prima fila e dei latch differenziali della seconda fila sono illustrati rispettivamente nelle figure 5.19(a) e 5.19(b).

L'architettura dei flip-flop presenta un nucleo  $C^2$ MOS con un numero di gate connessi al segnale di clock, cioè di fatto al VALID, pari a 4. Questo è importante perché il VALID è connesso al pin di clock di tutti i flip-flop e le sue commutazioni caricano e scaricano i gate dei MOS da loro offerti ad ogni passo. Tenendo conto che ad ogni step solo uno dei 10 flip-flop deve effettivamente svolgere una attività, il consumo di potenza al nodo di VALID rispetto al pilotaggio della logica è una fonte di grande inefficienza. Le varie architetture di flip-flop proposti dalla tecnologia UMC 130nm offrivano un numero minimo di gate al clock pari a 8, cioè il doppio rispetto a quelle realizzate. I transistori

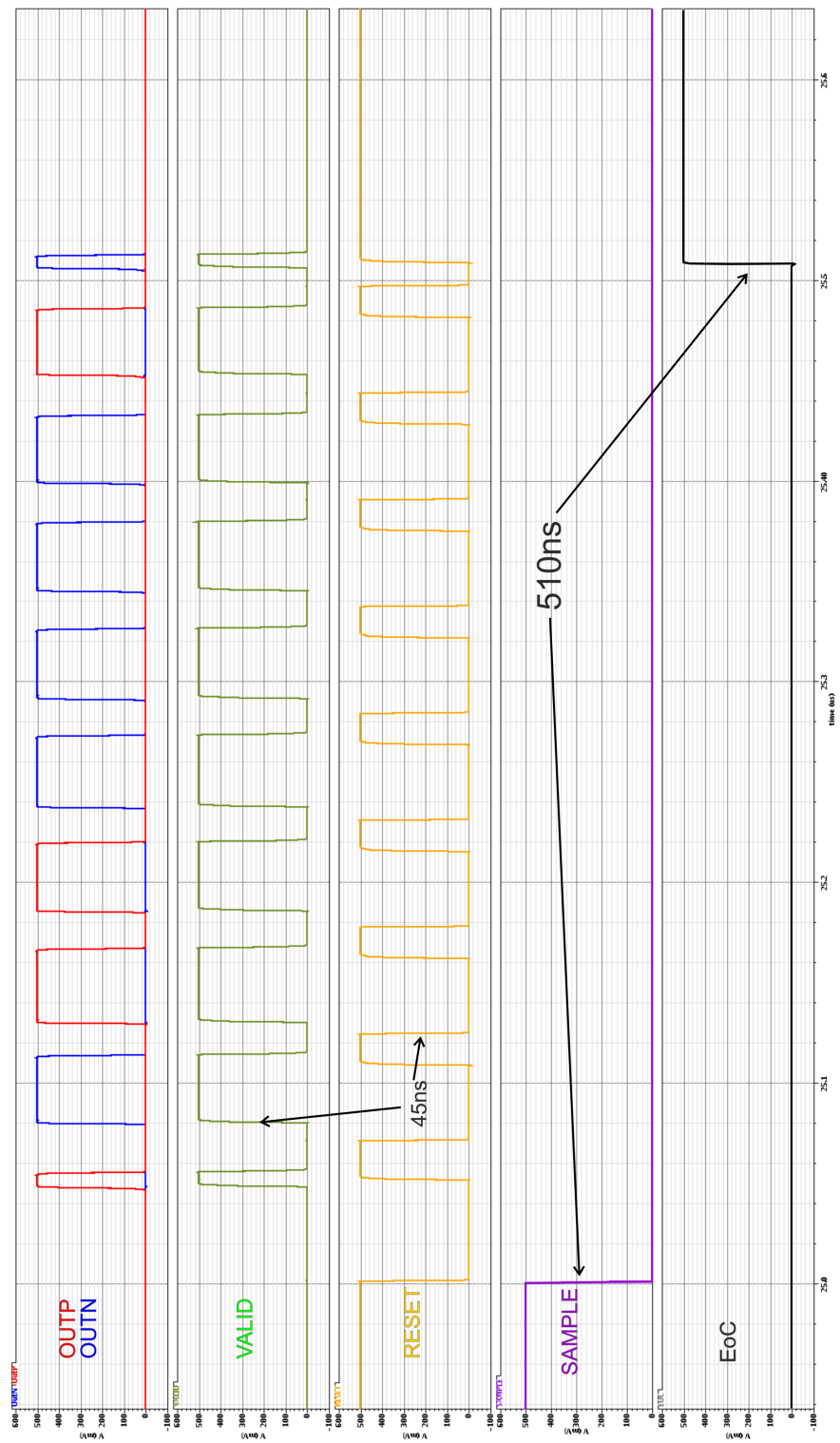


Figura 5.18: Timing diagram dei segnali logici del convertitore durante un ciclo di conversione.

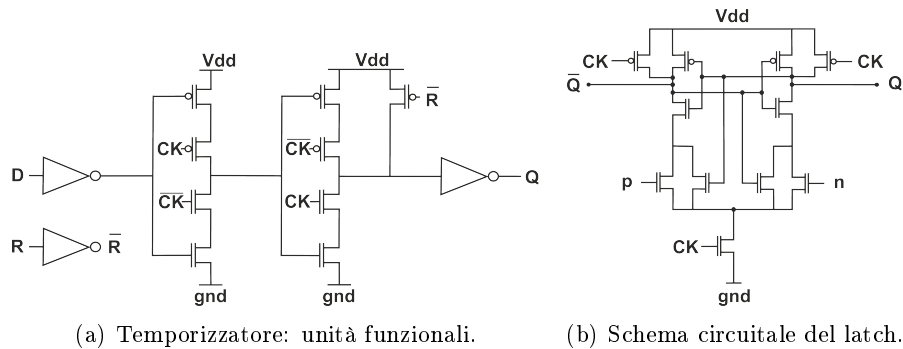


Figura 5.19: Circuiti logici dinamici full-custom.

utilizzati sono stati realizzati con le minime dimensioni possibili ed un inverter locale genera il segnale negato del reset per garantire una certa ripidità e coordinazione dei fronti del segnale SET/RESET ed evitare problemi di corse critiche e glitch che altrimenti potrebbero verificarsi nei circuiti  $C^2$ MOS [68].

I latch dinamici della seconda fila sono stati realizzati con una architettura differenziale perchè essa permette migliore efficienza [68] e per sfruttare il più possibile la natura differenziale del convertitore. La architettura circuitale è simile a quella del tradizionale latch differenziale ma è opportunamente modificata per ottenere una struttura che rappresenta una topologia nuova, progettata appositamente per l'applicazione. Come si può esaminare dallo schema circuitale e dall'illustrazione del suo funzionamento (Fig. 5.20) il latch differenziale commuta in base alle tensioni ai suoi ingressi in corrispondenza del fronte di salita del clock (cioè l'uscita del corrispondente flip-flop) ma non commuta se il clock è basso oppure se il clock è alto ma ha il circuito ha già svolto la sua commutazione utile. In questo modo, anche se tutti i latch risultano sempre connessi alle uscite del comparatore (come sempre avviene nella logica SAR), nessuno commuta se non quando è effettivamente sollecitato da un fronte di clock. Né quando gli ingressi sono entrambi a zero (comparatore resettato) né quando commutano nel verso opposto a quello in cui il latch è assestato (in seguito ad una precedente commutazione ad esempio), le uscite del latch cambiano e non vi è mai neppure un percorso conduttivo tra l'alimentazione e massa. Per tornare pronto a una nuova commutazione il circuito può solo essere resettato, e ciò avviene al termine della conversione. Questo garantisce il corretto funzionamento del convertitore evitando configurazioni errate dell'array di capacità e riduce al minimo le correnti di crossbar. Inoltre, l'architettura differenziale scelta minimizza, rispetto ad un latch tradizionale, il numero di gate connessi alle uscite del comparatore, imponendo un limite più basso alla potenza consumata ad ogni ciclo di conversione.

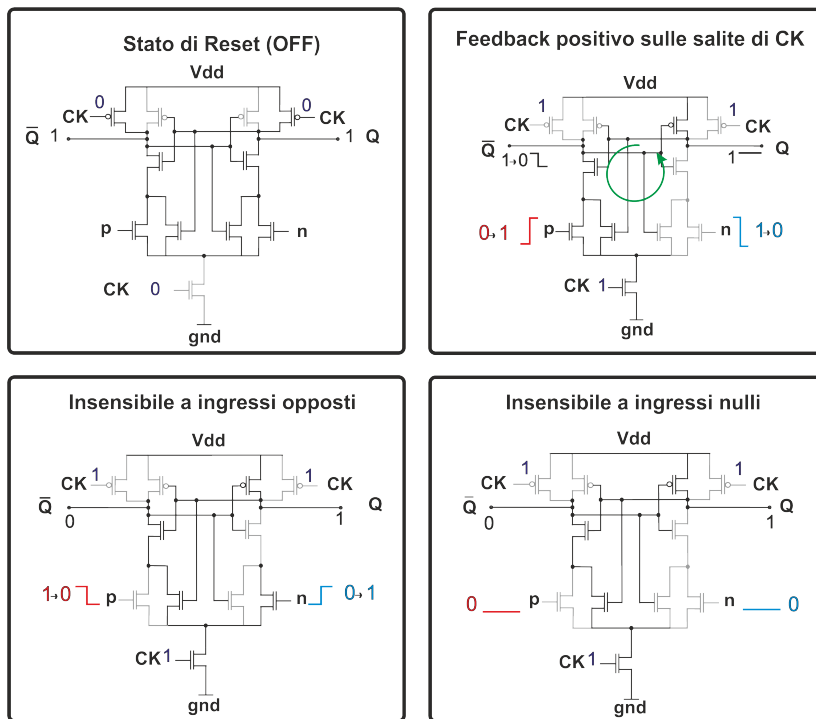


Figura 5.20: Schema di funzionamento del latch della seconda fila.



## 5.4 Risultati

Questa sezione è organizzata in due brevi paragrafi. Nel primo paragrafo sono riportate le immagini e le descrizioni del layout dei singoli componenti, dell'ADC complessivo e dell'intero front-end per il singolo canale di acquisizione; nel secondo e ultimo paragrafo sono riportate le prestazioni dell'ADC risultato di simulazioni post-layout e sono confrontate con i principali lavori in letteratura riguardanti ADC general-purpose e ADC per applicazioni biomediche.

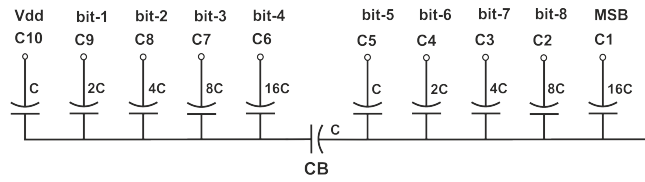
### 5.4.1 Layout

In un progetto full-custom il layout di ogni componente rappresenta una fase critica, ma l'array conserva comunque il ruolo più importante. Lo schema della sua configurazione, con riferimento alla sua struttura, è mostrato in figura 5.21(a), mentre la sua equivalente ed effettiva implementazione fisica è mostrata nell'immagine successiva.

Sebbene una disposizione a centroide comune sia potenzialmente in grado di minimizzare l'impatto del mismatch, essa non è stata utilizzata perchè aumenta la complessità e l'estensione delle connessioni con il rischio di incrementare oltremodo i parassiti ai nodi di top, che non dovrebbero essere superiori ai 30-40fF. L'organizzazione dell'array avviene comunque secondo una struttura dotata di un certo grado di simmetria per ridurre l'impatto del mismatch, ma è pensata principalmente per minimizzare i parassiti ai nodi di top dei due sub-array. L'estrazione dei parassiti effettuata in *CadenceVirtuoso<sup>TM</sup>* evidenzia una capacità parassita di 23fF all'array del LSB e di 25fF all'array del MSB. Il layout del comparatore è mostrato in figura 5.22 ed occupa una superficie di  $40 \times 40 \mu m^2$ . La sua implementazione è stata effettuata cercando di rispettare la massima simmetria possibile per ridurre al minimo le cause deterministiche di offset.

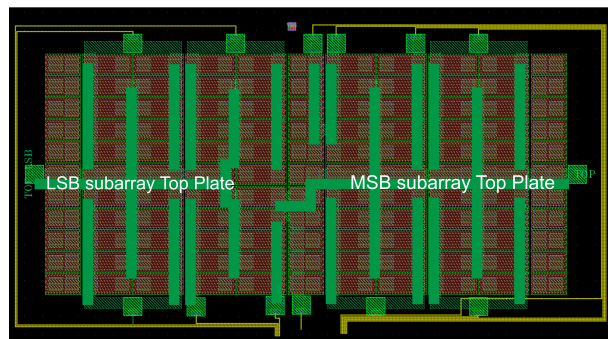
L'ADC occupa complessivamente un'area di  $190 \times 240 \mu m^2$ , cioè  $0.0456 mm^2$ , pari a circa il 30% della massima area imposta dalle specifiche. Il suo layout è mostrato in figura 5.23 e i dettagli relativi alla logica ed agli interruttori con bootstrap sono in figura 5.24.

Il layout dell'intero canale di acquisizione è illustrato in figura 5.25 insieme ai circuiti di servizio necessari al suo funzionamento. Il layout completo del chip consegnato alla fonderia UMC è, infatti, costituito anche da amplificatori per strumentazione appositamente progettati per la lettura dei segnali di uscita in fase di testing, dai riferimenti di corrente necessari al funzionamento e da undici compatti level-shifter necessari a traslare il livello delle uscite dell'ADC ad una logica 0-3V per un più semplice interfacciamento con dispositivi di testing. Come si può vedere dalla figura 5.25 i riferimenti di corrente, gli unici circuiti ausiliari facenti effettivamente parte del front-end, occupano uno spazio ridotto ed in ogni caso essi non dovranno essere replicati in ogni canale, dunque il loro impatto in termini di ingombro è trascurabile. Le singole dimensioni del canale sono di  $550 \times 250 \mu m^2$ , cioè  $0.1375 mm^2$ ; questo valore è inferiore al 75% dell'area massima imposta dalle specifiche.



D	D	D	D	D	D	D	D	D	D	D
D	C7	C7	C6	C6	D	C9	C3	C1	C1	D
D	C7	C7	C6	C6	D	C9	C3	C1	C1	D
D	C7	C7	C6	C6	D	C10	C3	C1	C1	D
D	C7	C7	C6	C6	D	D	C3	C1	C1	D
D	D	D	D	D	D	D	D	D	D	D
D	C6	C6	C8	CB	D	C1	C1	C2	C2	D
D	C6	C6	C8	C5	D	C1	C1	C2	C2	D
D	C6	C6	C8	C4	D	C1	C1	C2	C2	D
D	C6	C6	C8	C4	D	C1	C1	C2	C2	D
D	D	D	D	D	D	D	D	D	D	D

(a) Disposizione delle capacità.



(b) Layout.

Figura 5.21: Layout dell'array.

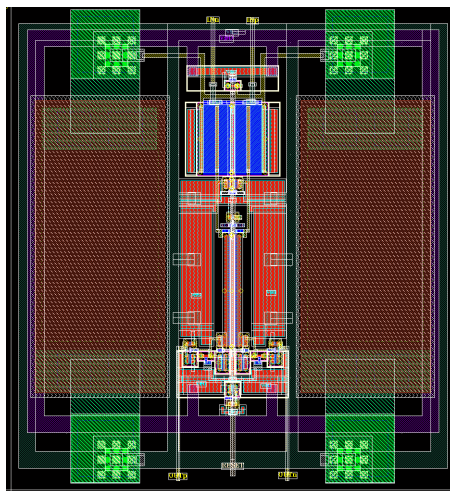


Figura 5.22: Layout del comparatore.

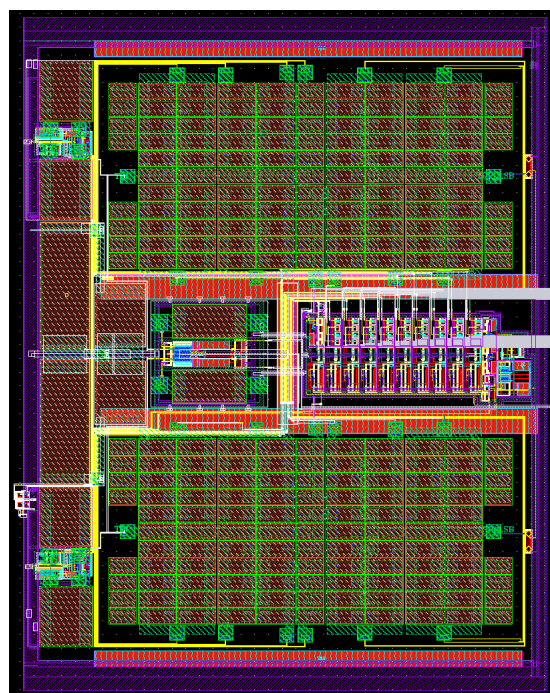
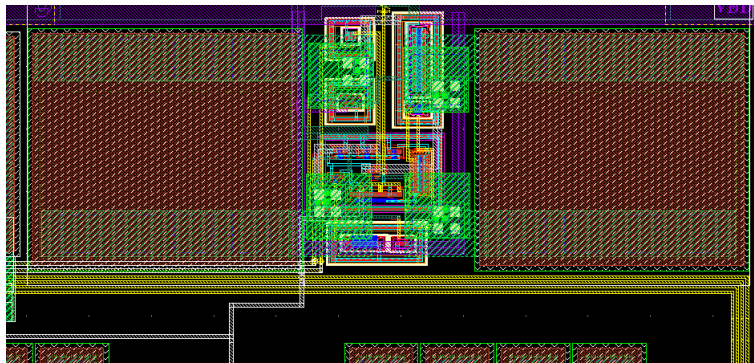


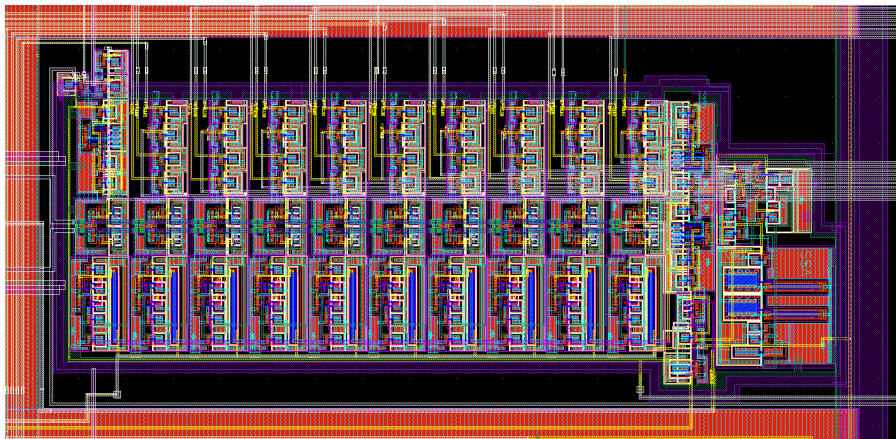
Figura 5.23: Layout dell'ADC.

#### 5.4.2 Prestazioni del Convertitore

Simulazioni post-layout sono state effettuate per verificare l'impatto dei parassiti sulla caratteristica di conversione e quantificare le prestazioni di velocità e consumo del



(a) Layout del bootstrap.



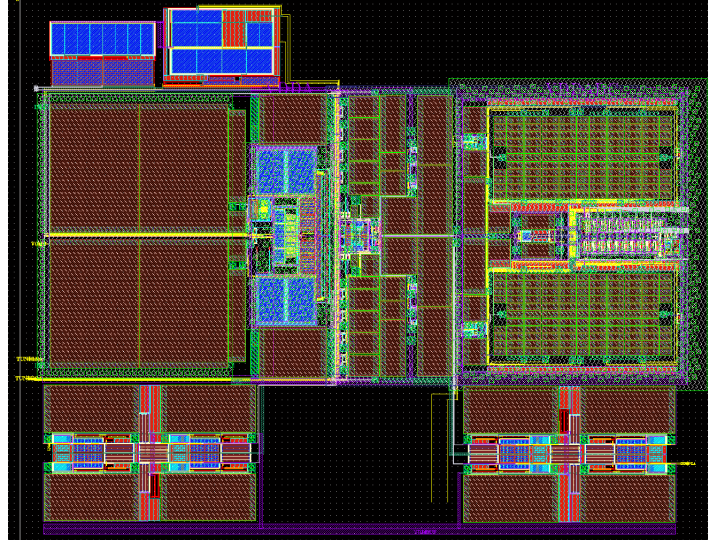
(b) Layout della logica.

Figura 5.24: Alcuni dettagli del Layout dell'ADC.

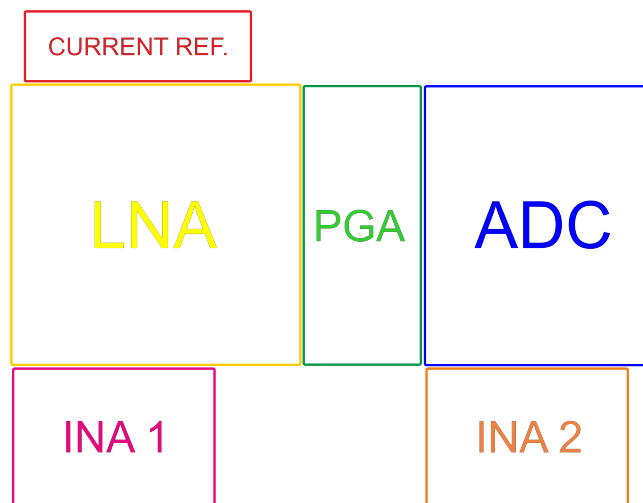
convertitore. Come spiegato anche nel precedente paragrafo, i parassiti estratti con il software Cadence *Assura*<sup>TM</sup> sono di 23fF e 25fF rispettivamente su top plate del MSB e LSB. Al primo si deve sommare la capacità di ingresso del comparatore ma questa non rappresenta una sorgente di non linearità. Non sono stati evidenziati accoppiamenti capacitivi tra i bottom plate delle capacità dell'array secondario e il top plate dell'array del MSB. I relativi diagrammi di DNL e INL del convertitore in funzione del livello (o codifica) di uscita sono mostrati in figura 5.26.

Sono quindi state svolte simulazioni in *MATLab*<sup>TM</sup> per valutare SNDR ed ENOB corrispondenti a questa caratteristica. In figura 5.27 è mostrato il SNDR in funzione dell'ampiezza di una sinusoide di ingresso. Si osserva un picco massimo di 57dB con un corrispondente ENOB di circa 9.18bit.

Le caratteristiche rappresentate non tengono conto dell'effetto dei mismatch pertanto

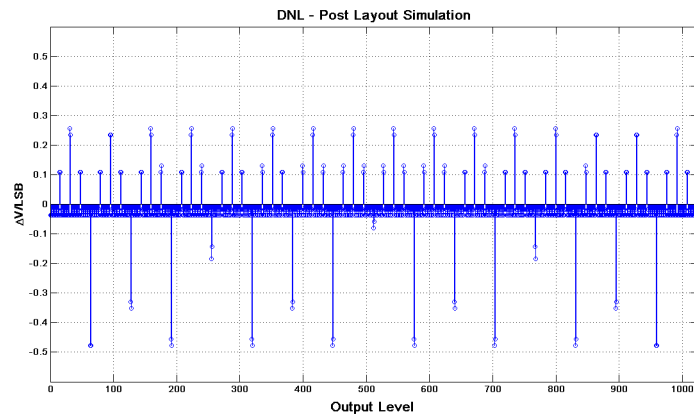


(a) Layout del singolo canale in uno screenshot da Cadence *Virtuoso<sup>TM</sup>*.

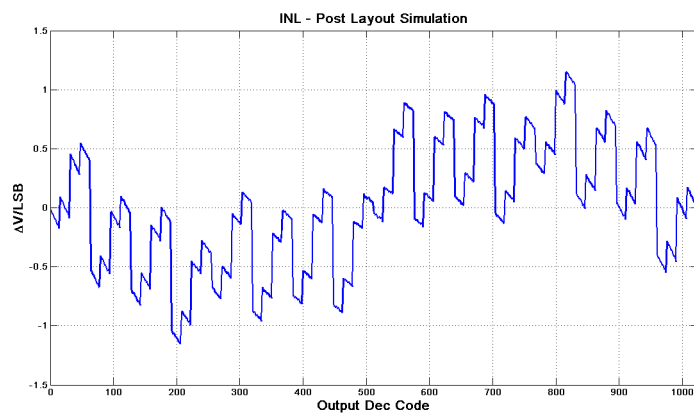


(b) Schematizzazione dei sottocircuiti.

Figura 5.25: Layout del front-end.



(a) DNL



(b) INL

Figura 5.26: DNL e INL simulate sull'estratto del convertitore in ambiente *Cadence<sup>TM</sup>*.



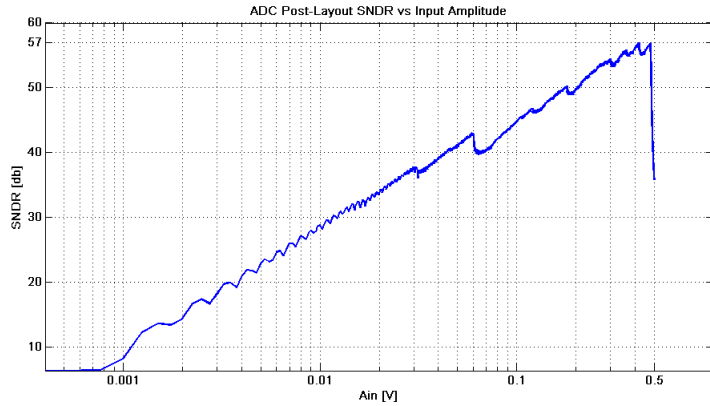


Figura 5.27: SNDR post-layout al variare dell'ampiezza di una sinusoide di ingresso.

si attende un degrado della linearità ma non così significativo da rappresentare un rischio per il rispetto delle specifiche di 8.2 bit minimi. Altre simulazioni sono state svolte per sovrapporre alla caratteristica DNL affetta da parassiti il contributo del mismatch tecnologico secondo le distribuzioni di  $\sigma_{DNL}$  per codifica tipiche della precisa architettura scelta e rappresentate in figura 5.28.

Per ognuna delle 1000 realizzazioni simulate è stato valutato l'ENOB, ottenendo un valore medio di 9.02 bit ed una deviazione standard di 0.082 bit (Fig. 5.29)

Il consumo complessivo del convertitore, quando questi lavora alla frequenza di campionamento di 30kHz fissata dalle specifiche, è di 73nW. Rispettivamente 40 nW sono da attribuire alla logica, 20nW al comparatore e 13nW all'array. Con l'ENOB realisticamente simulato e a questa frequenza di conversione, la Figura di Merito è circa pari a:

$$FoM = \frac{P_{tot}}{2^{ENOB} f_s} \approx 4.65 fJ/cstep \quad (5.15)$$

cioè molto vicina alla migliore figura di merito misurata sperimentalmente dell'attuale stato dell'arte, pari a 4.4fJ/cstep [61]. Per un confronto più equilibrato con lo stato dell'arte dei convertitori, ed in particolare dei SAR, occorre valutare la figura di merito al massimo della velocità alla quale può lavorare il convertitore. Come mostrato nella precedente sezione, il convertitore completa il ciclo di conversione in media dopo 510ns dal fronte negativo del segnale SAMPLE, a fronte dei 15μs resi disponibili dall'applicazione. L'ADC può quindi lavorare a frequenze anche 10 volte superiori. Dato che il consumo del convertitore realizzato durante il singolo ciclo di conversione non è influenzato dalla frequenza di campionamento (che dipende solo dalla fase di sample), la valutazione della FoM a frequenze superiori può offrire un risultato più favorevole perché ridurrebbe il consumo superfluo dovuto alle correnti di leakage (anche superiori a 35nA

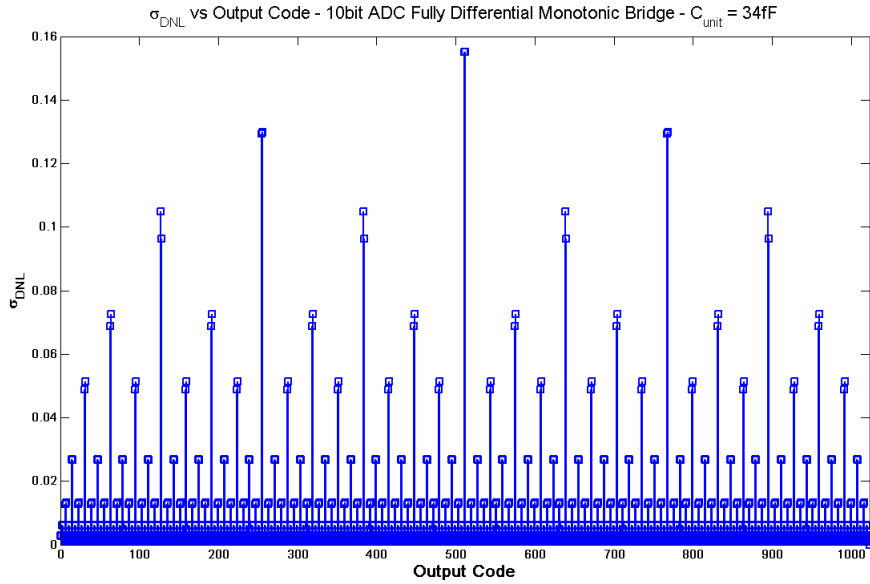


Figura 5.28: Deviazioni standard della DNL in funzione della codifica per l'ADC realizzato in base ai parametri di mismatch della tecnologia.

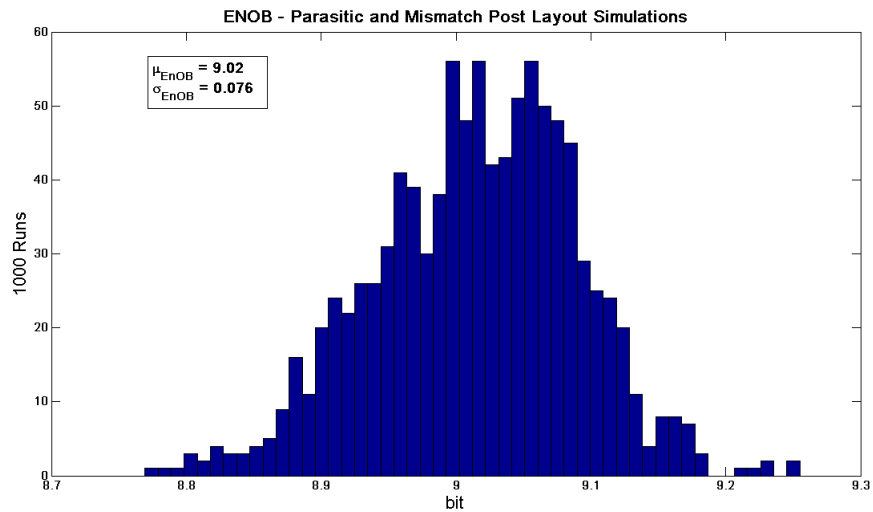


Figura 5.29: Istogrammi dell'ENOB associato alla caratteristica dell'ADC realizzato affetta da parassiti e mismatch.



complessivi) che, lavorando a 30kHz, sarebbero assorbite dall'alimentazione durante il lungo intervallo temporale che intercorre tra la fine della conversione e il nuovo fronte di salita del SAMPLE. In figura 5.30 è mostrato l'andamento della figura di merito con la frequenza di campionamento per il convertitore realizzato.

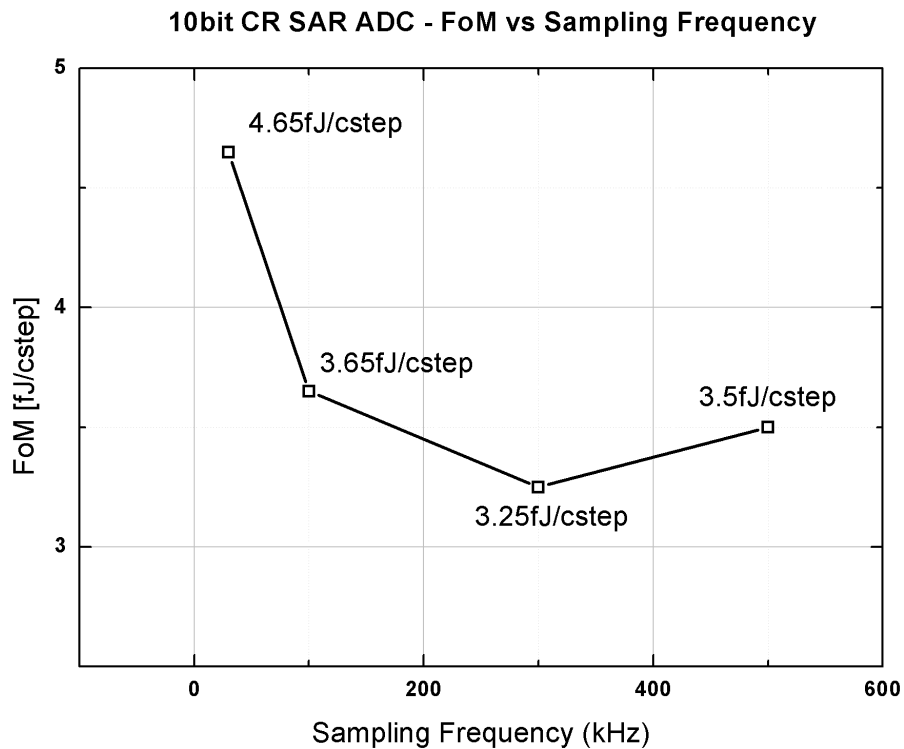


Figura 5.30: FoM del convertitore al variare della frequenza di campionamento.

Dalle simulazioni di corner si ottiene una frequenza massima con cui il convertitore può lavorare pari a circa 500kSps per un consumo totale di 925nW. Considerando che l'ENOB non subisce variazioni in seguito alla variazione del segnale di SAMPLE, la figura di merito del convertitore diventa pari a 3.5fJ/cstep. Una cifra simile si può ottenere già lavorando a frequenze di 100kSps, mentre la migliore si ottiene a 300kSps ed è pari a 3.25fJ/cstep. Nella tabella 5.2 le prestazioni del convertitore sono riassunte e confrontate con i principali lavori in letteratura.

La FoM simulata del convertitore è finora la migliore, sarà tuttavia necessaria una verifica sperimentali della accuratezza dei risultati di simulazione per un confronto equo con lo stato dell'arte. Questo sarà possibile nel prossimo Novembre, alla consegna del test chip realizzato da parte dei fabbricanti. Le prestazioni simulate appaiono comunque di livello molto buono e la FoM risulta inferiore anche a quelle di convertitori ad alta efficienza ed integrati con processi tecnologici più scalati. Se confrontata con i migliori convertitori realizzati con array di tipo bridge [62], la FoM ottenuta in simulazioni è

Tabella 5.2: Tabella riassuntiva e di confronto con i principali lavori internazionali delle prestazioni del convertitore.

Parametri	Risultati	Nauta '10	Harpe '11	Liu '10	Maloberti '08
Rif.	NA	[61]	[71]	[58]	[62]
<i>Tecnologia</i>	130nm	65nm	90nm	130nm	180nm
$V_{DD}$	0.5V	1V	1V	0.6V	1V
N bit	10	10	8	10	12
$C_{unit}$	34fF	$0.5fF^*$	$0.5fF^*$	$4.8fF^*$	120fF
$C_{tot}$	4.28pF	512fF	300fF	5pF	15.24pF
Area	$0.0456mm^2$	$0.0258mm^2$	$0.0205mm^2$	$0.0517mm^2$	$0.24mm^2$
$P_{cons}$	550nW	$1.9\mu W$	$26.3\mu W$	$46\mu W$	$3.8\mu W$
ENOB	9.02	8.75	7.7	8.91	9.4
FoM	3.25fJ/cstep	4.4fJ/cstep	12fJ/cstep	9.6fJ/cstep	56fJ/cstep
$f_{max}$	300kSps	1MSps	10.24MSps	10Msps	100kSps

\* Full-Custom Capacitors

circa un ordine di grandezza inferiore. Inoltre, si tenga presente che il convertitore è stato realizzato una tecnologia meno scalata rispetto a quelli con cui sono stati realizzati molti dei convertitori con le migliori FoM [61] [71] e senza utilizzare capacità custom. In figura 5.31 è mostrato il trend di evoluzione delle FoM in funzione del nodo tecnologico ed è evidenziato il punto in cui si colloca il convertitore realizzato, mentre in figura 5.32 il risultato ottenuto in simulazione è confrontato con le migliori figure di merito di ADC SAR realizzati con le differenti tecnologie submicrometriche.

Il progetto di un SAR ad altissima efficienza richiede nella maggior parte dei casi ed anche con i processi tecnologici più scalati, il ricorso ad approcci tecnologici full-custom per i vari componenti. L'array è in genere la struttura del convertitore più importante per la qualità del segnale ed è intorno a questo elemento che prende forma il progetto di un convertitore CR SAR. La Binaria e la Bridge sono le due strutture più utilizzate in questo campo e da una lettura dei principali lavori pubblicati in campo internazionale si potrebbero dire appartenenti a due differenti filosofie progettuali.

La prima sfrutta la grande linearità dell'array binario e le relativamente buone caratteristiche di mismatch delle tecnologie submicrometriche per realizzare convertitori ad alta velocità ed efficienza. Ottenere le prestazioni desiderate e circuiti compatti con questo approccio è tuttavia possibile solo accettando di realizzare i vari componenti e soprattutto le capacità unitarie, in modalità full-custom. Lo svantaggio è naturalmente un grande incremento della complessità progettuale. Spingere l'array binario al limite della sua sensibilità al mismatch per minimizzare capacità totale, aumentare la velocità e ridurre consumi impone, infatti, di utilizzare capacità minime un ordine di grandezza inferiore a quelle in genere offerte dagli standard tecnologici. In quel caso, però, anche l'effetto dei parassiti, normalmente meno problematico, può diventare critico.

Il bridge capacitor array presenta per sua natura una superiore sensibilità al mismatch

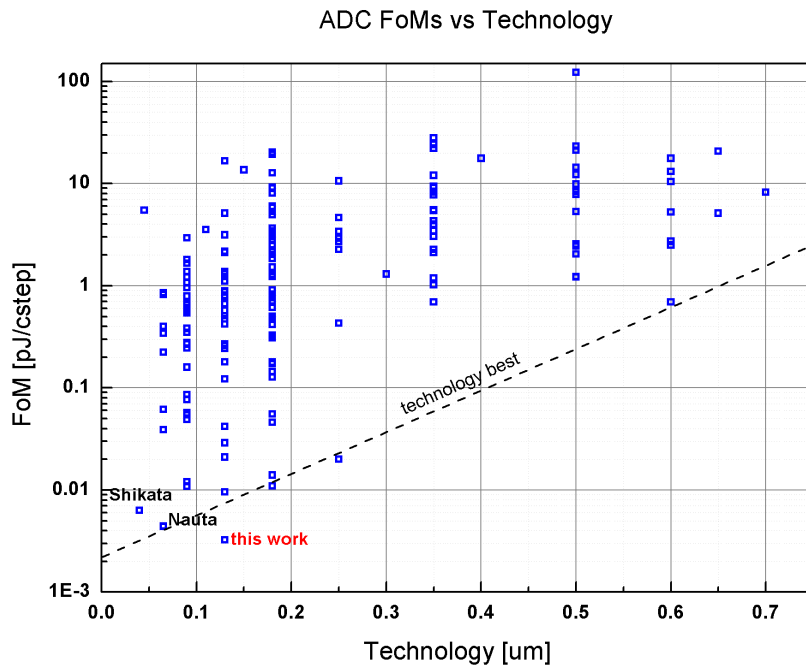


Figura 5.31: Trend di evoluzione delle FoM in funzione del nodo tecnologico.

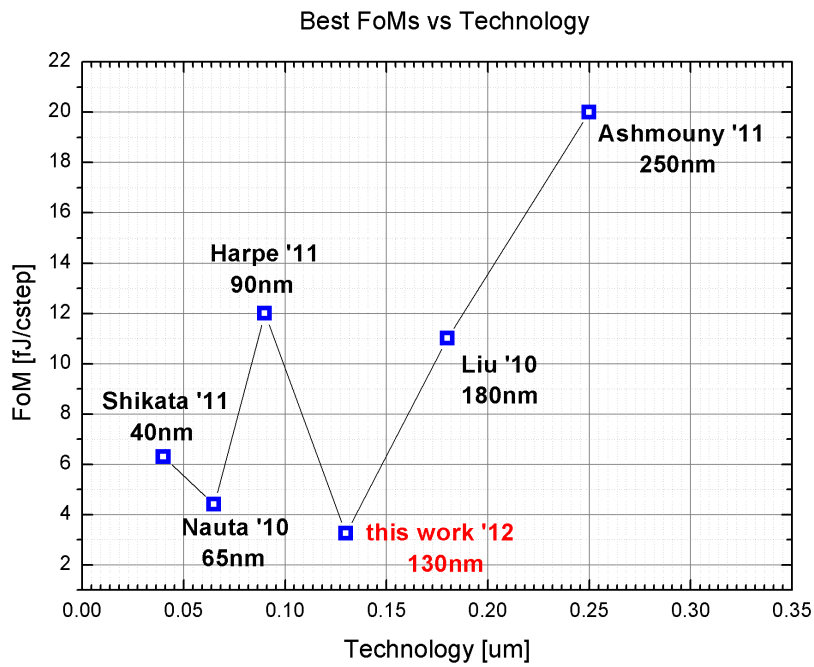


Figura 5.32: Migliori FoM in funzione del nodo tecnologico.

tecnologico e ai parassiti ma permette di realizzare strutture compatte e con una capacità totale contenuta senza il ricorso ad una progettazione full-custom, semplicemente utilizzando condensatori poco superiori a quelli minimi offerti dalla tecnologia. Raggiungere frequenze operative superiori ai 10Msps con questa classe di array ed alimentazioni scalate è più difficile. Sarebbe infatti necessario ridurre la capacità totale fino a usare capacità unitarie così piccole da degradare a livelli inaccettabili la qualità della conversione. Tuttavia per velocità inferiori si possono comunque ottenere alte efficienze.

Laddove si desideri ridurre quanto più possibile i consumi e se si è disposti ad affrontare una complessità progettuale di poco superiore ma comunque inferiore a quella richiesta per la realizzazione di nuove capacità, è possibile realizzare una logica asincrona e full-custom. In queste architetture, infatti, la logica è uno dei componenti più dispendiosi e ci si deve spesso adeguare alle prestazioni spesso poco ottimizzate offerte dalle librerie digitali messe a disposizione dalla tecnologia. Una scelta simile è attuabile utilizzando comunque transistori di libreria e non costituisce uno sforzo eccessivo dato la ristretta varietà di unità logiche che costituiscono un ADC SAR ed è questa la scelta che ha reso possibile raggiungere i risultati ottenuti.

## Capitolo 6

# Il Test Chip

Questo breve capitolo conclusivo è dedicato all'illustrazione del *Miniasic* effettivamente realizzato riportandone le prestazioni attese e descrivendone i circuiti ausiliari. Alcuni di questi, come i riferimenti di corrente, sono parte integrante del sistema e sono necessari al suo funzionamento, altri hanno il solo scopo di facilitare l'attività di testing una volta che il sistema sarà integrato. Sebbene rivestano un'importanza secondaria rispetto ai circuiti descritti fino a questo punto, questi componenti hanno richiesto una progettazione adeguata e fanno parte a tutti gli effetti dell'attività di tesi. In Figura 6.1 è illustrato lo schema a blocchi del circuito completo delle sue unità di servizio e con evidenziate le connessioni ai PAD dei principali segnali di ingresso e di uscita.

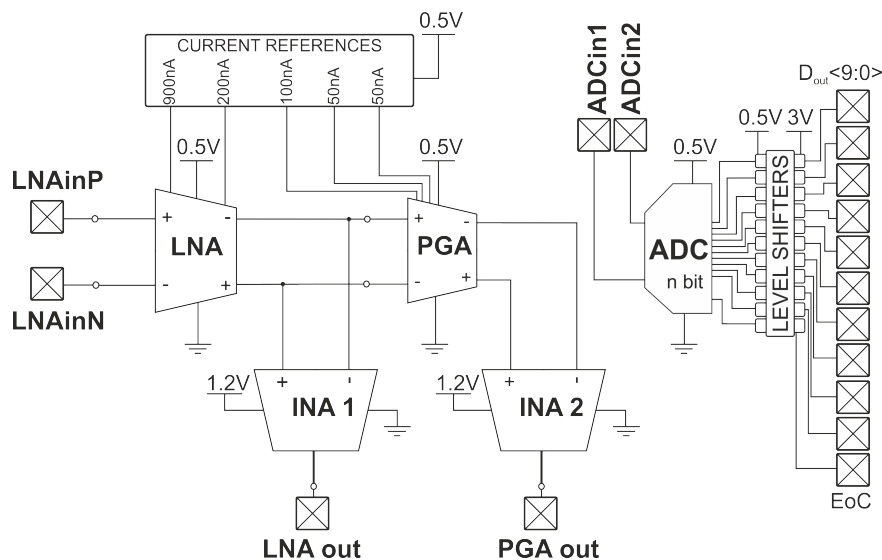


Figura 6.1: Schema del test chip e dei circuiti di servizio.

Il layout di alcuni dei blocchi di servizio qui descritti compare nella figura 5.25 del precedente capitolo insieme a quello dell'intero canale. Nel prossimo paragrafo sono

riportate le prestazioni attese per l'intero front-end realizzato accompagnate da alcune considerazioni sui risultati ottenuti; i paragrafi successivi sono invece dedicati alla descrizione dei circuiti di servizio.

Tra questi, solo i riferimenti di corrente faranno effettivamente parte del canale e sono necessari a garantire la corretta polarizzazione degli stadi amplificanti. Essi in generale devono dimostrare una buona reiezione dei disturbi di alimentazione (in realtà poco presenti in questa applicazione) e la minima sensibilità possibile alle variazioni di processo. Testare un sistema significa innanzitutto verificarne il buon funzionamento e in secondo luogo quantificarne le prestazioni. Ciò dovrebbe idealmente essere fatto per ciascun sottocircuito e compatibilmente con la quantità di pin a disposizione nel package che si è scelto di utilizzare per il test chip. Per verificare la qualità del segnale di uscita di LNA e PGA sono stati realizzati due buffer in grado di leggere il segnale differenziale della sorgente (uscite di LNA o PGA) e uscire in modalità single-ended. Per garantire ciò sono stati utilizzati degli OTA configurati in modo tale da realizzare un INA i cui dettagli saranno approfonditi in seguito. Infine è stato necessario implementare dei traslatori di livello (level shifters) per rigenerare le ridotte transizioni dei segnali logici di uscita dell'ADC (0-0.5V) fino a valori di tensione di 0-3V più facili da interfacciare con componenti elettronici off-chip per la verifica delle prestazioni (FPGA, microcontrollori, DAC). L'ultimo paragrafo illustra brevemente la loro struttura e i loro layout. In figura 6.2 è illustrato il layout complessivo del test chip consegnato alla fabbrica. Il front-end di acquisizione e digitalizzazione è situato nella regione inferiore sinistra del chip. Al di sotto del canale di acquisizione, rispettivamente alle due estremità si possono osservare gli INA; i riferimenti di corrente sono posizionati all'angolo superiore sinistro della regione occupata dal front-end. I level shifters sono situati alla destra dell'ADC e sono vicini a un'altra serie di componenti, prevalentemente capacità di decoupling necessarie al corretto funzionamento dell'integrato. I due oscillatori che si possono notare nella regione superiore del chip non fanno parte del lavoro di tesi ma appartengono ad un altro progetto di ricerca sviluppato con lo stesso processo tecnologico e che per motivi logistici e di costo è stato implementato nello stesso miniasic di  $1.5 \times 1.5 \text{mm}^2$ .

## 6.1 Prestazioni del front-end

Nella tabella 6.1 sono riassunte le prestazioni dell'intero front-end e confrontate con i principali lavori nel campo dei sistemi di neural monitoring. Come si può verificare dai numeri riportati nella tabella il front-end soddisfa le specifiche di efficienza offrendo un consumo complessivo inferiore al 30L'amplicazione è adeguata all'applicazione e il numero di bit effettivi sono tali da non determinare un peggioramento del rapporto segnale rumore attraverso la digitalizzazione. Se confrontate con quelle dei principali lavori presenti in letteratura per l'amplicazione, il filtraggio e la digitalizzazione di segnali neurali, le prestazioni si dimostrano tra le migliori in quasi tutti gli aspetti considerati.

In lavori come [18] il consumo per il singolo equivalente front-end è dominato dalla necessità di utilizzare dei multiplexer e dei buffer a causa della ripartizione di più stadi

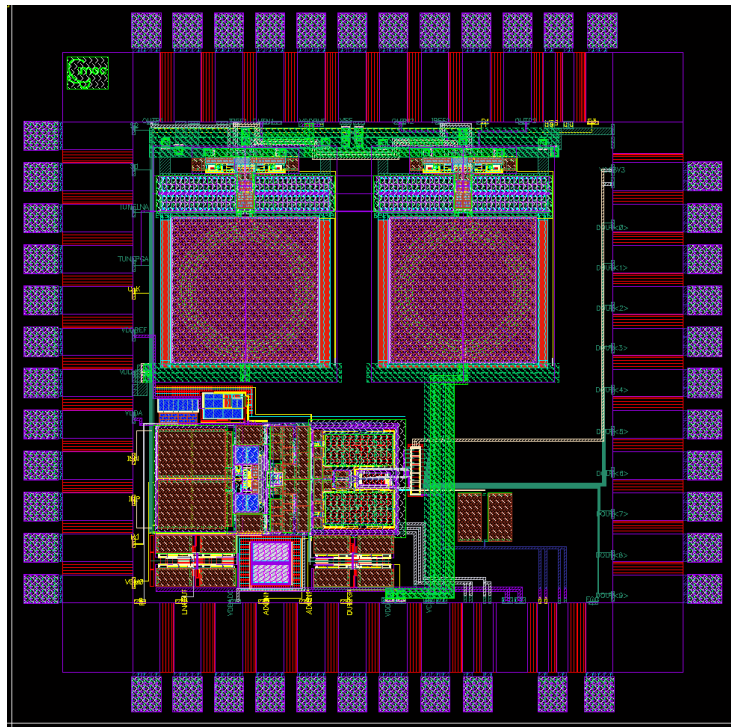


Figura 6.2: Immagine del Miniasic.

di amplificazione e filtraggio tra un numero di ADC inferiore al numero totale di canali. In [31] invece la NEF e la PEF risultano scadenti perchè le ottime prestazioni di rumore riferito in ingresso sono state ottenute a spese di diversi  $\mu A$  di corrente assorbita dal primo stadio di amplificazione della catena. Si può affermare con una certa confidenza che l'adozione di una filosofia progettuale con due amplificatori ed un ADC per singolo canale si è rivelata vincente sia in termini di efficienza che in termini di prestazioni. E' qui opportuno sottolineare tuttavia che il risultato non sarebbe stato raggiunto senza un adeguato sforzo per la realizzazione di ciascun blocco funzionale (LNA e ADC in particolare) ad alta efficienza. Le prestazioni e l'efficienza ottenute per i vari circuiti progettati si sono rivelate sia nel contesto applicativo che in assoluto di livello molto buono. In attesa delle verifiche sperimentali previste per il prossimo novembre e di poter replicare ed integrare il front-end realizzato con il sistema di trasmissione UWB progettato dal team del Prof. Bevilacqua dell'Università di Padova, le prospettive riguardo la realizzazione di un sistema di neural monitoring di efficienza superiore a qualsiasi altro lavoro realizzato finora appaiono favorevoli.

Tabella 6.1: Tabella riassuntiva e di confronto con i principali lavori internazionali delle prestazioni del front-end.

Parametri	Risultati	Liew'11	Chae '08	Rabaey'12	Genov '10
Rif.	NA	[50]	[18]	[40]	[42]
<i>Tecnologia</i>	130nm	130nm	350nm	65nm	130nm
$V_{DD}$	0.5V	0.5V	$\pm 1.65V$	0.5V	1.2V
Area/ch.	$0.1375mm^2$	$0.073mm^2$	NA	$0.013mm^2$	$0.09mm^2$
Power/ch.	$0.93\mu W$	$1.13\mu W$	NA	$5\mu W$	$\approx 7\mu W$
NEF	2.1	3.09	8.45	5.99	7.2
PEF	2.22	4.77	235	17.96	62.2
Amplifiers and Filters					
$P_{cons}$ 0.86	nW	$0.86\mu W$	$2.54\mu W$	$0.8\mu W$	$4.8\mu W$
BW	0.1Hz-10kHz	40Hz-400Hz	0.1-2kHz	0.1Hz-10kHz	0.1Hz-10kHz
IRnoise(rms)	$4\mu V_{rms}$	$5.32\mu V_{rms}$	$4.9\mu V_{rms}$	$4.3\mu V_{rms}$	$6.5\mu V_{rms}$
Gain	40/59dB	48/54dB	57/60dB	32dB	54/60dB
CMRR	86dB	NA	90dB	75dB	78dB
ADC					
N bit	10	8	6-9	7-9	8
$P_{cons}$	73nW	103nW	$21\mu W$	$0.24\mu W$	$1.8\mu W$
ENOB	9.02	7.4	NA	7.8	7.8
FoM	4.65fJ/cstep	21fJ/cstep	NA	84fJ/cstep	141fJ/cstep
$f_{sampling}$	30ksps	30ksps	40ksps	20ksps	57ksps

## 6.2 Circuiti Ausiliari

### 6.2.1 I Riferimenti di Corrente

I riferimenti di corrente richiesti per la polarizzazione degli stadi interni LNA e PGA e delle reti di reazione di modo comune sono di quattro diversi tipi:

- 900nA per il primo stadio del LNA;
- 200nA per il secondo stadio del PGA;
- 100nA per la rete di reazione di modo comune del PGA;
- 50nA per il primo stadio del PGA e per la rete di modo comune del secondo stadio del LNA;

La prima corrente di polarizzazione tra quelle elencate è significativamente più elevata di tutte le altre ed è stata generata separatamente per non ricorrere a rapporti dimensionali tra MOS troppo elevati. Le correnti di 50nA, 100nA e 200nA sono invece state generate con lo stesso circuito. La topologia utilizzata per la generazione delle correnti è il riferimento a  $G_m$  costante rappresentato in figura 6.3.



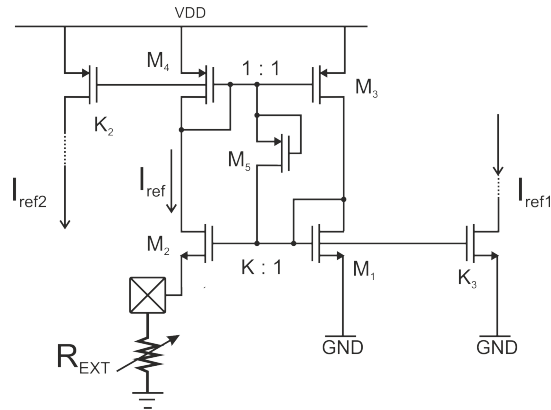


Figura 6.3: Riferimento di corrente utilizzato.

Questa topologia permette di generare una corrente dipendente dalla resistenza  $R$  connessa al source di  $M_2$  (fig. 6.3) e dal rapporto di forma tra quest'ultimo e il suo omologo connesso al gate (oltre che dai parametri tecnologici). Nell'ipotesi di utilizzare transistori polarizzati in forte inversione la corrente portata da ciascun ramo del circuito è definita dalla relazione:

$$I_{ref} = \frac{2}{\mu_n C'_{ox} (\frac{W}{L})_1 R^2} \left(1 - \frac{1}{\sqrt{K}}\right) \quad (6.1)$$

In realtà a causa della alimentazione ridotta pari a 0.5V i transistori dello stadio operano in debole inversione, in questo caso la corrente di riferimento è [49]:

$$I_{ref} \approx \frac{nU_T}{R} \ln(K) \quad (6.2)$$

I morsetti di source dei due diversi circuiti di riferimento sono stati connessi a due PAD del chip per potervi collegare resistenze esterne  $R_{EXT}$  regolabili e avere un maggiore controllo delle polarizzazioni, gli specchiamenti per ottenere le correnti utili sono effettuati con PMOS oppure NMOS a seconda del verso richiesto. Per ottenere la corrente di 900nA è stato utilizzato un rapporto di forma unitario del MOS  $M_1$  e un rapporto dimensionale  $K$  pari a 2 con  $30.8k\Omega$  di resistenza esterna. Attraverso simulazioni Montecarlo delle variazioni di processo e mismatch si osserva una buona stabilità del valore generato, come evidenziato dai risultati riportati nella tabella 6.2.

Il circuito dedicato alla generazione di tutte le altre correnti di riferimento è stato realizzato per portare una corrente di 100nA e specchiarla con rapporti dimensionali pari a 0.5 e 2 allo scopo di ottenere rispettivamente correnti di 50nA e 200nA. Il rapporto di forma del MOS  $M_1$  di questo riferimento è stato fissato a 1/10 mentre il parametro  $M$  è

sempre pari a 2. Con una resistenza esterna di 300k si ottengono le correnti desiderate con buona stabilità rispetto alle variazioni di processo e ai mismatch (Tabella 6.2).

Tabella 6.2: Risultati delle simulazioni Montecarlo (500 realizzazioni) per le correnti di riferimento generate.

Corrente	Media	Sigma
Primo Stadio LNA	902.95nA	13.61nA
Rete CMFB LNA	49.78nA	1.34nA
Primo Stadio PGA	51.39nA	1.44nA
Secondo Stadio PGA	202.498nA	3.83nA
Rete CMFB PGA	102.28nA	1.9nA

Il MOS M5 mostrato in figura 6.3 ha lo scopo di ausilio allo start-up e impedisce che all'accensione il circuito raggiunga il punto di equilibrio stabile caratterizzato da correnti nulle in ogni ramo.

### 6.2.2 Gli INA

Lo schema interno degli INA 1 e 2 di figura 6.1 è mostrato in figura 6.4.

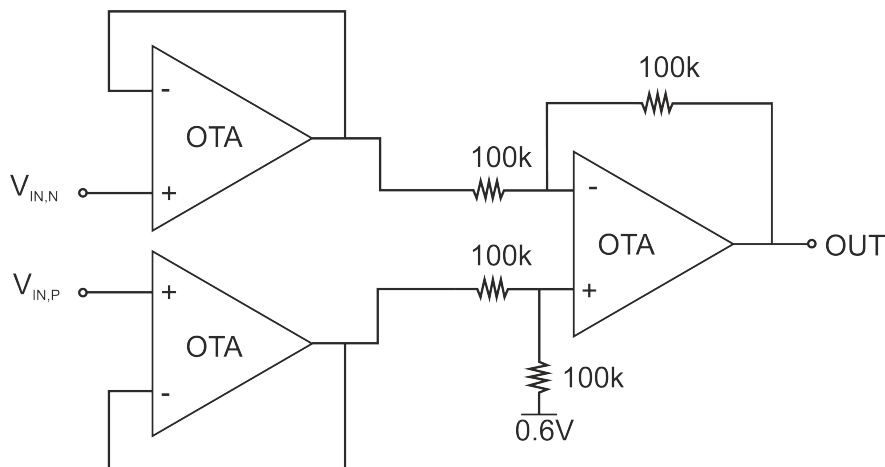


Figura 6.4: Struttura degli INA utilizzati come buffer di lettura delle uscite di LNA e PGA.

L'alimentazione dello stadio è differente da quella degli altri circuiti ed è di 1.2V. Lo stadio è pensato per rilevare un segnale completamente differenziale e tradurlo in un segnale single-ended polarizzato a metà dinamica, cioè circa 0.6V. L'architettura è costituita da tre identici OTA la cui struttura interna è quella classica a due stadi con ingresso a PMOS e uscita in classe A ed è illustrata in figura 6.5.

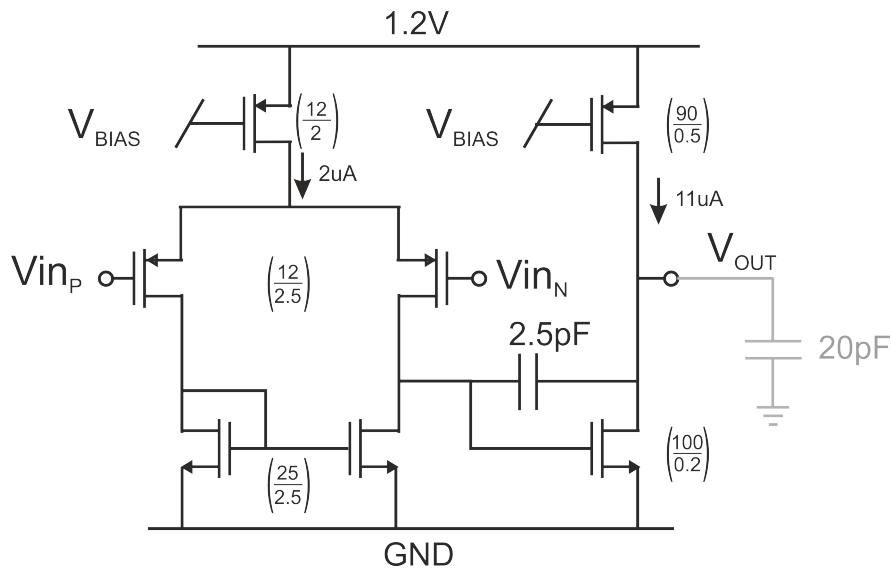


Figura 6.5: Struttura d'interna degli OTA utilizzati per l'INA.

Dato che lo stesso tipo di OTA è utilizzato all'uscita dell'INA, è stato dimensionato in previsione di pilotare carichi capacitivi di 20pF, circa pari alla capacità parassita attesa al pin di uscita e dovuta al pin stesso, alle piste di interconnessione afferenti e agli strumenti di misura. La tensione di polarizzazione dei generatori è comune ai due stadi ed è generata localmente attraverso la serie di due transdiodi e una resistenza. Nella tabella 6.3 sono elencate le prestazioni dell'INA.

Tabella 6.3: tabella riassuntiva delle prestazioni degli INA (Simulazioni MonteCarlo).

Parametro	Media	Sigma
$V_{OS}$	2.2mV	3mV
CMRR	88dB	10dB
$P_{TOT}$	250 $\mu$ W	12.8 $\mu$ W
BW	487kHz	21.7kHz
Gain	0dB	28m dB

La banda è passa basso ed è più ampia di quella prevista per i circuiti analogici del front-end e raccoglie tutta l'informazione utile potenzialmente presente all'uscita degli amplificatori. Il consumo risulta elevato ma essendo l'INA un componente di servizio utile solo in fase di testing non rientra nel calcolo dell'efficienza nel front-end.

### 6.2.3 I Traslatori di Livello

La funzione dei traslatori di livello è rigenerare il valore delle uscite digitali rappresentate dai bit di uscita dell'ADC e il segnale di fine conversione con un livello logico alto associato a una tensione di alimentazione superiore rispetto a quella del convertitore e pari a 3V. La loro struttura è rappresentata in figura 6.6. Essi ricevono all'ingresso i segnali logici a 0.5V ed i loro negati e attraverso un latch rigenerano il segnale.

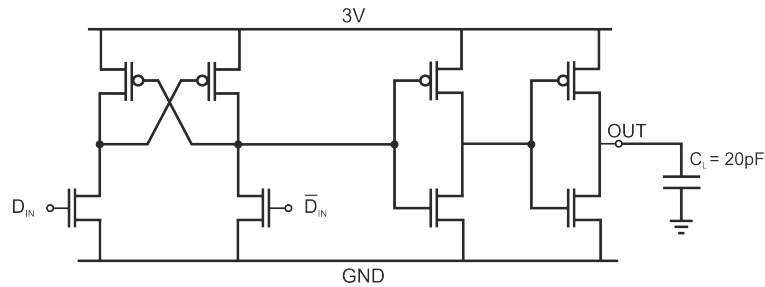


Figura 6.6: Traslatore di livello.

La catena di inverter in uscita è stata dimensionata per pilotare carichi capacitivi elevati offerti dalle piste di interconnessione di una scheda a discreti come quella di testing e per farlo in tempi compatibili con il testing dell'ADC anche a frequenze di campionamento superiori ai 30kHz, almeno fino a 500kHz. Perchè ciò si garantito è necessario che la commutazione dei level shifters si completi in meno di un  $\mu s$ . La figura 6.7 mostra il risultato di simulazioni Montecarlo (processo e mismatch) per il transitorio dell'uscita del level shifter in corrispondenza di segnali di ingresso con un fronte di 2.5ns tra 0 e 0.5V e con un carico di 20pF. Come si può osservare i tracciati delle 300 realizzazioni mostrano che il completamento della commutazione avviene in meno di 100ns, tempo ampiamente sufficiente per il testing del convertitore.

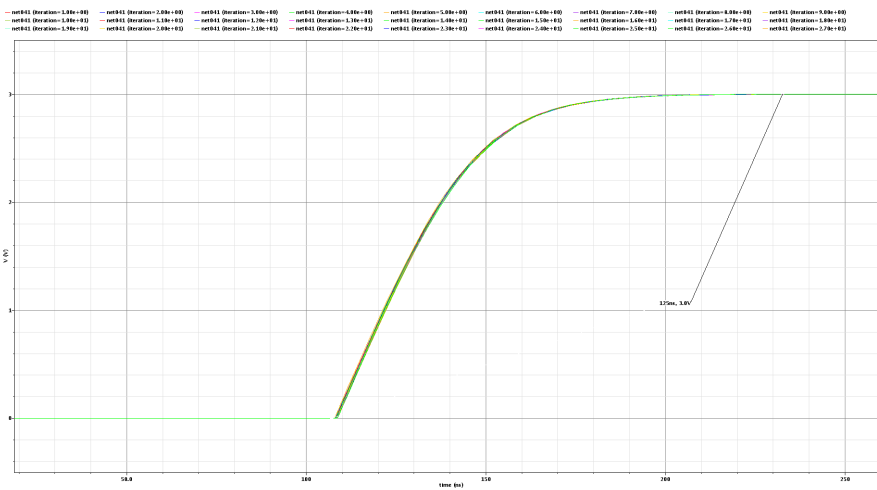


Figura 6.7: Traslatore di livello: commutazione delle uscite (analisi transitoria Montecarlo).

# Conclusioni

In questo lavoro è presentato un front-end progettato in tecnologia UMC 130nm ed alimentato a 0.5V per l'acquisizione, l'elaborazione e la digitalizzazione di segnali neuronali. Il sistema è stato simulato fino alla fase post-layout ed è attualmente in fabbricazione presso le fonderie UMC. Il front-end è parte di un sistema integrato multicanale per il monitoraggio neuronale e che sarà dotato anche di una unità wireless UWB per la trasmissione dei dati acquisiti a unità di elaborazione remote. Il progetto realizzato costituisce l'architettura di ognuno dei 64 canali previsti per il chip finale ed è composto dalla cascata di un amplificatore a basso rumore (LNA), un amplificatore a guadagno programmabile (PGA), un ADC SAR a redistribuzione di carica e i relativi riferimenti di corrente necessari al funzionamento. Al solo scopo di facilitare il testing del front-end, sono stati progettati anche alcuni circuiti ausiliari: due buffer per la lettura delle uscite degli amplificatori e dei traslatori di livello per traslare a tensioni più alte i segnali logici di uscita del convertitore analogico-digitale. Ogni blocco è realizzato in configurazione fully differential. LNA e PGA offrono una amplificazione programmabile con quattro bit di selezione tra 40dB e 59dB in una banda compresa tra 0.1Hz e 10kHz e con un rumore riferito in ingresso di circa  $4\mu V_{rms}$ . Il primo stadio del LNA è realizzato con una topologia tail-current-reuse per la riduzione del rumore riferito in ingresso fino a circa  $4\mu V_{rms}$  mentre gli stadi di uscita di entrambi gli amplificatori sono progettati per garantire la massima dinamica. Con queste caratteristiche è possibile la rilevazione e l'amplificazione sia dei potenziali d'azione che dei LFPs. Ogni stadio è progettato con una opportuna di rete per il controllo del segnale di modo comune ed il CMRR complessivo è di circa 86dB. Il consumo dei due amplificatori è di 860nW per una corrente assorbita pari a circa  $1.7\mu A$ .

Il convertitore analogico digitale è posto a valle degli amplificatori ed è di tipo SAR a redistribuzione di carica. Per il DAC interno al convertitore è stata proposta ed utilizzata una nuova topologia realizzata attraverso l'unione di una procedura di commutazione monotona e una struttura di array con capacità di attenuazione già presenti separatamente in letteratura. E' inoltre proposto un nuovo criterio per la selezione della topologia di DAC ed il dimensionamento della capacità unitaria. In particolare, attraverso la derivazione di nuove formule chiuse che legano l'impatto di parassiti e mismatch tecnologico alla DNL del convertitore, è stato possibile effettuare una stima preliminare dell'ENOB in funzione della capacità unitaria per diverse topologie di DAC e differenti procedure di commutazione. Successivamente, attraverso l'ausilio di modelli *MATlab*<sup>TM</sup> appositamente implementati, è stato indagato l'impatto di particolari distribuzioni di DNL

sull'ENOB di un convertitore per effettuare un confronto tra diverse topologie. Attraverso gli stessi modelli è stato possibile effettuare per la prima volta una stima del consumo delle topologie di DAC con capacità di attenuazione. La capacità unitaria scelta per l'ADC progettato è di 34fF ed è stata realizzata con le capacità MIM del designkit. Il comparatore è stato realizzato con una topologia completamente dinamica mentre la logica progettata è di tipo asincrono ed è realizzata con un approccio full-custom. In particolare essa è costituita da flip-flop  $C^2MOS$  dinamici, da un nuovo tipo di latch dinamico differenziale e una unità di temporizzazione realizzati appositamente per l'applicazione. Nel contesto del front-end il convertitore realizzato lavora ad una frequenza di campionamento di 30kSps consumando 73nW con un ENOB di 9.02 bit ed ottiene una FoM di 4.65fJ/cstep. Le caratteristiche dell'ADC realizzato gli permettono tuttavia di funzionare anche a frequenze superiori, fino a 500kSps, e di ottenere in simulazione una FoM di 3.25fJ/cstep, inferiore a qualsiasi altra presentata in letteratura. Complessivamente il front-end occupa un'area di  $0.1375mm^2$  compatibile con le specifiche e consuma complessivamente  $0.93\mu W$ . Il rumore riferito in ingresso alla catena di amplificazione e filtraggio è dominato dal contributo del primo amplificatore ed è di circa  $4\mu V_{rms}$ . I risultati di simulazione post-layout indicano una NEF e una PEF per tutto il sistema rispettivamente pari a 2.1 e 2.22 ed inferiori rispetto a quelle degli equivalenti integrati di neural monitoring presentati in letteratura.

## Appendice A

# Inaccuratezza dei Modelli di Rumore in Regime di Debole Inversione

I modelli di rumore dei dispositivi della tecnologia forniti dal DesignKit si dimostrano inaccurati in regime di funzionamento di sottosoglia. Per questo motivo, nel corso del capitolo dedicata agli amplificatori è stato spiegato che i dimensionamento sono stati effettuati in modo analitico. In questa sezione mostriamo alcuni risultati di simulazione che evidenziano l'inadeguatezza dei modelli di rumore termico per il funzionamento in sottosoglia e il superamento dei valori massimi teorici.

Nel funzionamento in debole inversione e sotto soglia il rumore dei dispositivi MOSFET assume le caratteristiche di un rumore shot. La relativa espressione della densità spettrale di rumore in corrente misurato all'uscita (drain o source) di un MOS in saturazione dipende dalla corrente di drain secondo la relazione:

$$\bar{I}_{noise}^2 = 2qI_D \tag{A.1}$$

Differenti simulazioni sono state effettuate per diversi tipi di transistori appartenenti a due diverse tecnologie: una tecnologia consolidata come la AMS  $0.35\mu m$  e la tecnologia utilizzata per il progetto cioè la UMC  $0.13\mu m$ . La configurazione delle simulazioni per ogni dispositivo è illustrata in figura A.1.

Le tensioni di alimentazione e bias sono state scelte per lavorare sempre con i MOS in saturazione. Fissato un adeguato rapporto di forma (in genere elevato per poter lavorare in sottosoglia con correnti tra il nA e il  $\mu A$ ) è stata fatta variare la corrente di saturazione per misurare il rumore rilevato da un generatore sonda. Contemporaneamente è stata calcolata la tensione tra gate e source per verificare il regime di funzionamento. L'andamento del rumore simulato per un NMOS standard della tecnologia AMS035 al variare della corrente di polarizzazione e in debole inversione (o sotto soglia) è mostrato in figura A.2.



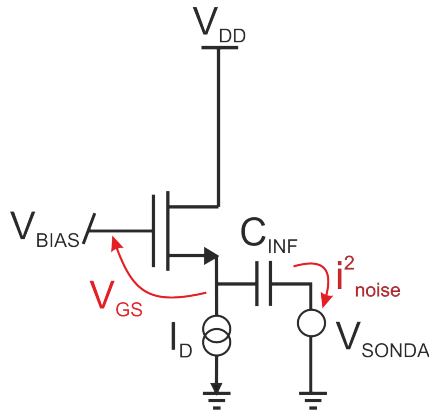


Figura A.1: Schema di simulazione per la verifica del rumore dei dispositivi.

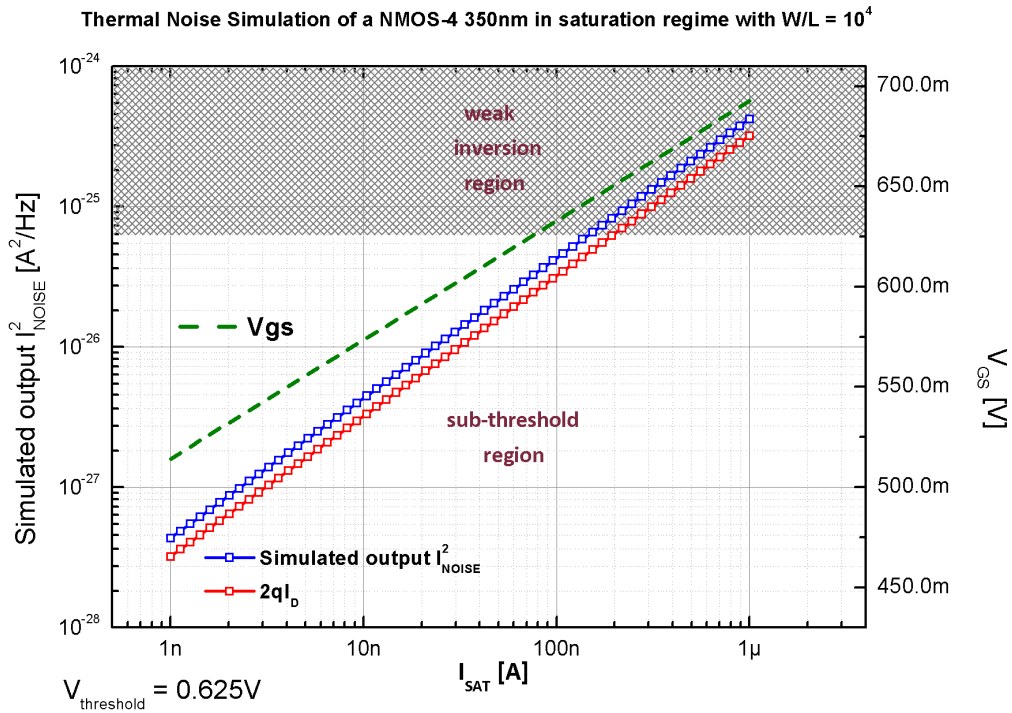


Figura A.2: Densità spettrale di rumore di un NMOS AMS035 in debole inversione e sottosoglia, confrontata con la stime teoriche.

Nel grafico si osserva un andamento coerente con i valori teoricamente attesi, perciò il modello si può considerare affidabile. Non si può affermare lo stesso per gli NMOS ad ossido spesso della tecnologia UMC130nm utilizzati per lo stadio di ingresso del LNA.

Come dimostra la figura A.3, il rumore risulta essere significativamente superiore ai

Thermal Noise Simulation of a NMOS HGLV 130nm in saturation regime with  $W/L = 10^4$

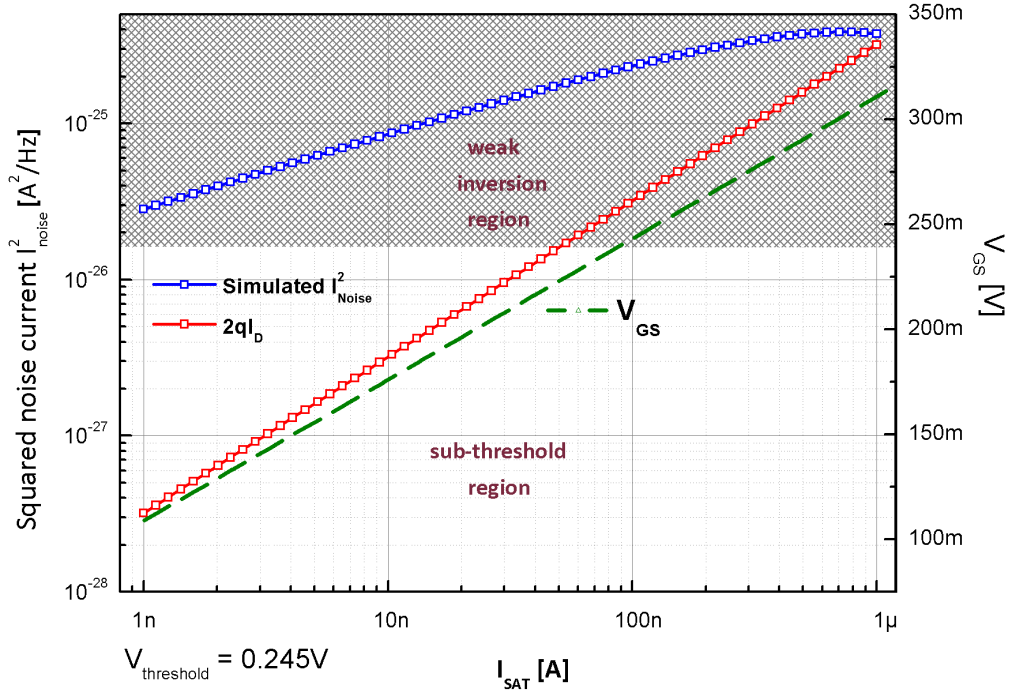


Figura A.3: Densità spettrale di rumore di un NMOS HGLV UMC130nm in debole inversione e sottosoglia, confrontata con la stime teoriche.

valori teoricamente possibili. Analoghe simulazioni sono state effettuate con altri dispositivi della stessa tecnologia e si evidenziano in ogni caso scostamenti simili. Altri gruppi di ricerca internazionali [52] che hanno lavorato con la stessa tecnologia ne hanno sottolineato l'inaffidabilità dei modelli e l'incoerenza con le misure effettuate, che dimostrano sempre valori in accordo con le stime teoriche.

## Appendice B

# Effetto del Mismatch sull'ENOB nelle Varie Topologie di CR SAR

In questa sezione sono riportate le curve ottenute dalle simulazioni effettuate per stimare il legame tra capacità unitaria e a degradazione dell'ENOB per effetto del mismatch, attraverso la caratteristica DNL di ogni singola topologia. La struttura con cui sono organizzate le simulazioni è quella schematizzata in figura 4.27 del capitolo dedicato all'ADC.

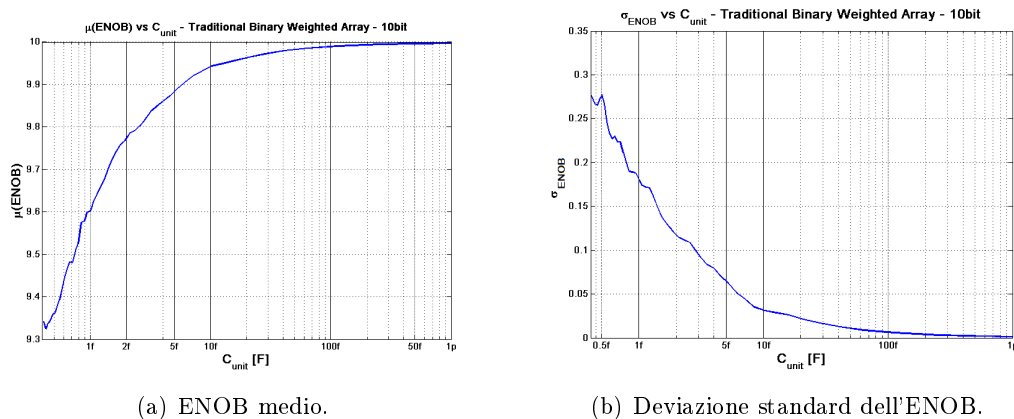
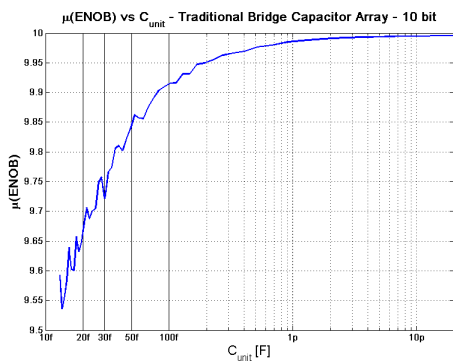
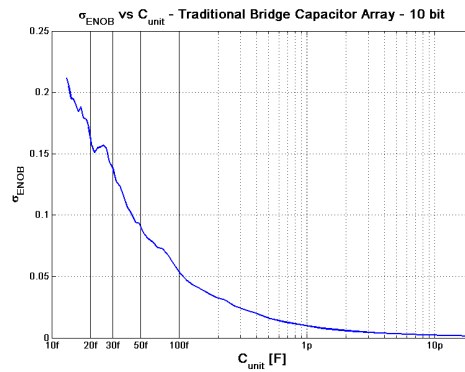


Figura B.1: Topologia Binaria Classica (Single Ended).

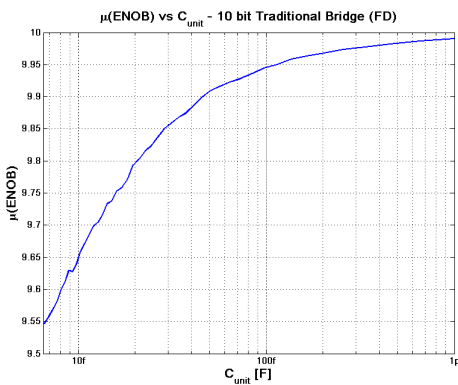


(a) ENOB medio.

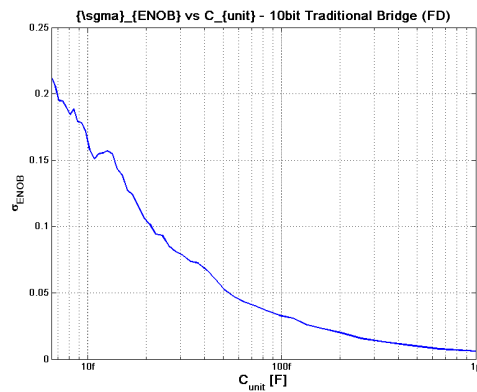


(b) Deviazione standard dell'ENOB.

Figura B.2: Topologia Bridge Tradizionale (Single Ended).

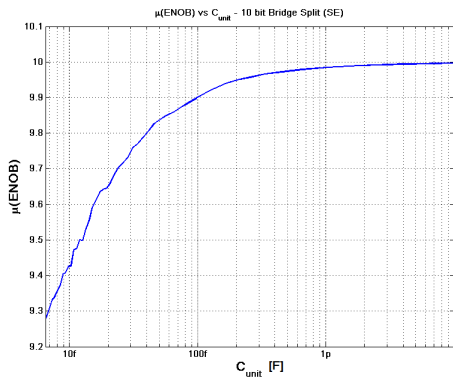


(a) ENOB medio.

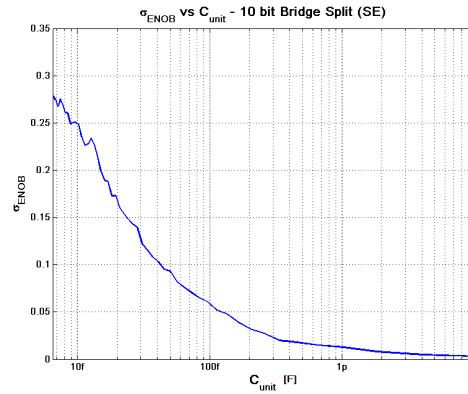


(b) Deviazione standard dell'ENOB.

Figura B.3: Topologia Bridge Tradizionale (Fully Differential).

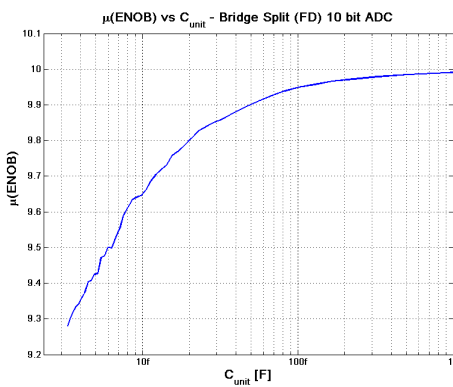


(a) ENOB medio.

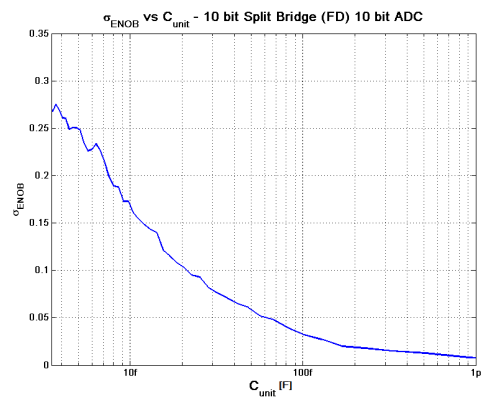


(b) Deviazione standard dell'ENOB.

Figura B.4: Topologia Split Bridge (Single Ended).

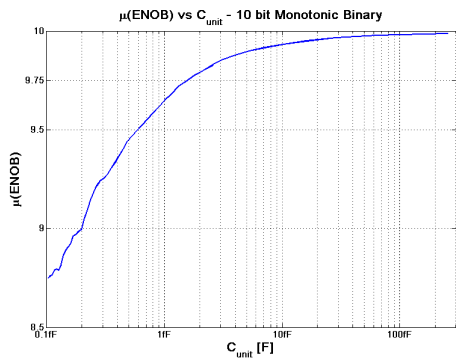


(a) ENOB medio.

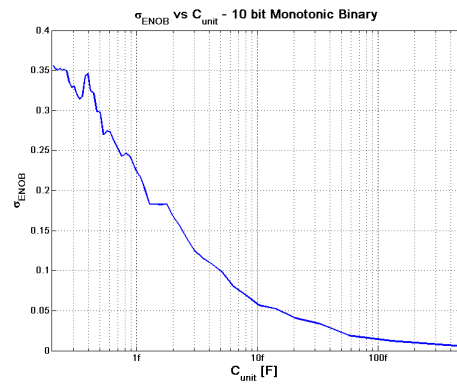


(b) Deviazione standard dell'ENOB.

Figura B.5: Topologia Split Bridge (Fully Differential).

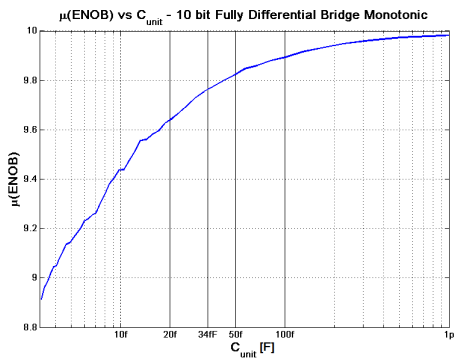


(a) ENOB medio.

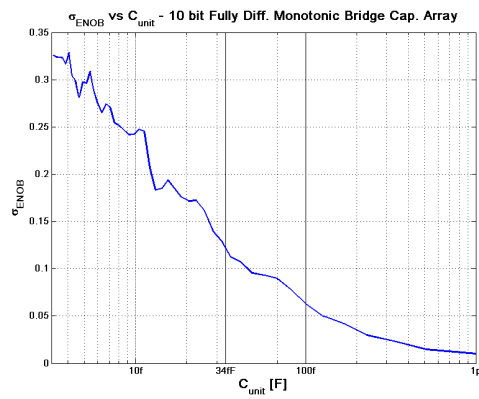


(b) Deviazione standard dell'ENOB.

Figura B.6: Topologia Binaria Monotona.



(a) ENOB medio.



(b) Deviazione standard dell'ENOB.

Figura B.7: Topologia Bridge Monotona.

# Bibliografia

- [1] P.Andlin-Sobocki, B.Jonsson,H.Wittchen, and J.Olesen, ” Costs of disorders of the brain in Europe ”, *Eur. J. Neurol.*, vol.12, pp.1-27, 2005.
- [2] WHO, *Neurological Disorders: Public Health Challenges*. Genève: WHO Press, 2007.
- [3] S.Tong and N.Thakor, ” Quantitative EEG Analysis Methods and Clinical Applications”. Norwood, MA: Artech House, 2009.
- [4] S.Breit, J. Schulz, and A.Benabid, ” Deep Brain Stimulation”, *Cell Tissue Res.*, 318, pp.275-288, 2004.
- [5] K.Iniewszki, ” CMOS Biomicrosystems, Where Electronics Meets Biology”, John Wiley & Sons, Inc., 2011.
- [6] K. Wise, D. Anderson, J. Hetke, D. Kipke and K. Najafi, ”Wireless Implantable Microsystem: High Density Electronic Interfaces to the nervous system”, *Proceedings of the IEEE*, pp. 76-97, 2004.
- [7] P. Campbell, K. Jones, R. Hubert, K. Horch, and R. Normann, ”A Silicon-Based, Thre Dimensional Neural Interface: Manufacturing Processes for an Intracortical Electrode Array”, *IEEE Trans. Biomed. Eng.*,38(12), pp. 758-768, 1991.
- [8] F. Solzbacher et al., ”Wafer Scale Fabrication of Penetrating Neural Electrode Arrays, *Biomedical Microdevices*, 2010.
- [9] T. Seese, H. Harasaki, G. Saidel, and C. Davies, ”Characterization of Tissue Morphology, Angiogenesis and Temperature in the Adaptive Response of Muscle Tissue to Chronic Heating”, *Lab. Invest.*, 78, pp1553-1562,1998.
- [10] K. Silay, C. Dehollain, and M. Declerq, ”Numerical Analysis of Temperature Elevation in the Head due to Power Dissipation in a Cortical Implant”, *Proceedings of the 2008 International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 951-956, 2008.
- [11] G. Webster, ”*Medical Instrumentation: Application and Design*” , John Wiley & Sons, 1998.

- [12] Reid R. Harrison, "The Design of Integrated Circuits to Observe Brain Activity", Proceedings of the IEEE , Vol. 96, No. 7, 2008.
- [13] Reid. R. Harrison, P. Watkins, R. Kier, R. Lovejoy, D. Black, B. Geiger, F. Solzbacher, "A Low-Power Integrated Circuit for a Wireless 100-electrode Neural Recording System", IEEE Journal of Solid-State Circuits, vol. 42, no.1, pp 123-133, 2007.
- [14] R. Harrison and C. Charles , "A low - power low - noise CMOS amplifier for neural recording applications", IEEE J. Solid - State Circuits , 38( 6 ), pp. 958 – 965 , 2003 .
- [15] A. M. Sodagar, G. E. Perlin, Y. Yao, K. Najafi, K. Wise. "An Implantable 64-channel Wireless Microsystem for Single Unit Recording". IEEE J. Solid State Circuit 2009, 9, 2591-2604.
- [16] Y. Perelman and R. Ginosar , "An Integrated System for Multichannel Neuronal Recording with Spike/LFP Separation, Integrated A/D Conversion and Threshold Detection", IEEE Trans. Biomed. Eng. , 54 ( 1 ), pp. 130 – 137 , 2007.
- [17] W. Wattanapanitch , M. Fee , and R. Sarpeshkar , "An Energy - Efficient Micropower Neural Recording Amplifier", IEEE Trans. Biomed. Eng. , 1 ( 2 ), pp. 136 – 147 , 2007.
- [18] M.S. Chae, W. Liu, Z. Yang, T. Chen, J. Kim, M. Sivaprakasam, M. Yuce. "A 128-Channel 6mW Wireless Neural Recording IC with On-the-Fly Spike Sorting and UWB Transmitter". IEEE Int. Solid-State Circ. Conf. 2008, 146-148.
- [19] Ghovanloo et al. , "An Inductively Powered Scalable 32-channel Wireless Neural Recording System-on-a-chip for Neuroscience Applications". IEEE Int. Solid-State Circ. Conf. , 2010, 330-333.
- [20] L. R. Hochberg et al. , "Neuronal Ensemble Control of Prosthetic Devices by a Human with Tetraplegia". Nature, 442, 164-171, 2006.
- [21] "International Technology Roadmap for Semiconductors", 2003.
- [22] "International Technology Roadmap for Semiconductors" , Overall Technology Roadmap Characteristics, 2011.
- [23] R. M. Swanson and J. D. Meindl, "Ion-Implanted Complementary MOS Transistors in Low-Voltage Circuits, IEEE Journal of Solid-State Circuits, vol. 7, no. 2, pp. 146-153, April. 1972.
- [24] E. Vittoz and J. Fellrath, "New Analog CMOS ICs Based on Weak Inversion Operation", in Proceedings of the European Solid-State Circuits Conference, Toulouse, Sept. 1976, pp. 12-13.



- [25] E. Vittoz et al. "CMOS Analog Integrated Circuits Based on Weak Inversion Operation", IEEE Journal of Solid-State Circuits, vol. 12, no. 3, pp. 224-231, June 1977.
- [26] A.B. Schwartz, "Cortical Neural Prosthetics" , Annual Review of Neuroscience, vol. 27, pp 487-501, 2004.
- [27] K. Wise, D.Anderson, J.Hetke, D. Kipke, K. Najafi, "Wireless implantable micro-system: High density electronic interfaces to the nervous system", in Proceedings of the IEEE, pp. 76-97, 2004.
- [28] D. Henze, Z. Borhegyi, J. Csicsvari, A. Mamiya, K. Harris, and G.Buzsaki, "Intracellular Features Predicted by Extracellular Recordings in the Hippocampus in-vivo, J. Neurophysiology , 84, pp. 390-400, 2000.
- [29] D.Henze, K. Harris, J. Csicsvari, H. Hirase and G. Buzsaki, "Accuracy of Tetrode Spike Separation as Determined by Simultaneous Intracellular and Extracellular Measurements", J. Neurophysiology, J. Neurophysiology, 84, pp. 401-414, 2000.
- [30] Ghovanloo et al., "An Inductively Powered Scalable 32-Channel Wireless Neural Recording System-on-a-Chip for Neuroscience Applications", IEEE Trans. Biomed. Circ. , Vol. 4, No. 6, 2008.
- [31] T. H. Meng et al., " HermesE: A 96-channel Full Data RAte Direct Neural Interface in 0.13 $\mu\text{m}$  CMOS", IEEE J. Solid-State Circ., vol. 47, No. 4, pp 1043-1055, 2012.
- [32] F. Zappa, R. Zappa, "Complementi di Elettronica", cap. 9, pp. 338-339 , Soc. Ed. Esculapio, 2006.
- [33] Z.Tang, B. Smith, J. Schild and P. Peckham, " Data Transmission from an Implantable Biotelemeter by Load-shift Keying sing CiUrcuit Configuration Modulator", IEEE Trans. Biomed. Eng., 42, pp.524-528, 1995.
- [34] K. Guillory, A. Misener, and A. Pungor, " Hybrid RF/IR Transcutaneous telemetry for Power and High Bandiwdth Data", Proceedings of the IEEE EMBS Conference, pp.4338-4340, 2004.
- [35] M. Baghaei-Nejad et al., " Remote-powered RFID Tag with 10Mb/s UWB Uplink and 18.5dbm sensitivity UHF downlink in 0.18 $\mu\text{m}$ ", IEEE Int. Solid-State Circ. Conf., pp 198-199,2009.
- [36] M. Klemm, G. Troester, " EM Energy Absorption in the Human Body Tissues due to UWB Antennas", Prog. Electromagnetics Res., 62, pp. 261-280, 2006.
- [37] T. Borghi et al., " A Power-Efficient Analog Integrated Circuit For Amplification and Detection of Neural Signals", Proceedings of the EMBC 2008, pp. 4911-4914, Vancouver, 2008.

- [38] T. Borghi et al., “ An Integrated Low Noise Multichannel System for Neural Signals Amplification”, Proceedings of the IEEE ESSCIRC, pp. 456-459, 2007.
- [39] A. Bonfanti et al., “ A Multi-Channel low-power IC for Neural Spike Recording with Data Compression and Narrowband 400MHZ MC-FSK Wireless Transmission ”, Proceedings of the ESSCIRC, 2010.
- [40] R. Muller, S. Gambini, J. Rabaey, “ A  $0.013mm^2$   $5\mu W$  DC-Coupled Neural Signal Acquisition IC with 0.5V Supply”, ISSCC, Session 17, Biomedical and Displays, 2011.
- [41] J. Steyaert and W. Sansen , “ A Micropower Low - Noise Monolithic Instrumentation Amplifier for Medical Purposes”, IEEE J. Solid - State Circuits , 22 , pp. 1163 – 1168 , 1987.
- [42] K. Abdelhalim, R. Genov, “ 915-MHz Wireless 64-Channel Neural Recording SoC with Programmable Mixed-Signal FIR Filters”, Proceedings of the ESSCIRC, 2011.
- [43] B. Murmann, “ A/D Converter Trends: Power Dissipation, Scaling and Digitally Assisted Architectures”, CICC 2008.
- [44] Tauro et al., “ Common Mode Stability in Fully Differential Voltage Feedback CMOS Amplifiers”, 2003.
- [45] J.F. Duque-Carrillo, “ Continuous-Time Common-Mode Feedback Networks for Fully Differential Amplifiers: a Comparative Study”, ISCAS, 1993, vol. 2, pp. 1267-1270.
- [46] P. M. Van Peteghem, J. F. Duque-Carrillo, “ A General Description of Common-Mode Feedback in Fully Differential Amplifiers”, ISCAS, 1990, vol. 4 pp.320-312.
- [47] J. Holleman, B. Otis, “ A Sub-Microwatt Low-Noise Amplifier for Neural Recording”, Proc. of the 29th Annual International Conference of the IEEE EMBS, 2007.
- [48] C. Enz , F. Krummenacher and E. Vittoz , “ An Analytical MOS Transistor Model Valid in All Region of Operation and Dedicated to Low - Voltage and Low - Current Applications”, Analog Integr. Circuits Process. , 8 , pp. 83 – 114 , 1995.
- [49] A. Wang, A. Chandrakasan, B. Calhoun, “ Sub-Threshold Design For Ultra Low-Power Systems”, Springer , 2006.
- [50] W. Liew, X. Zou, Y. Lian, “A 1V  $22\mu W$  32 Channel Implantable EEG Recording IC”, ISSCC, 2010.
- [51] W. Liew, X. Zou, Y. Lian, “ A 0.5V  $1.13\mu W$ -channel Neural Recording Interface with Digital Multiplexing Scheme”, ISSCC, 2011.
- [52] Genat et al., “ Silicon Strips Readout using Deep Sub-Micron Technologies, 12th Workshop on Electronics for LHC and Future Experiments, Valencia, 2006.

- [53] G. Giustolisi, G. Palumbo, “ Introduzione ai Dispositivi Elettronici ”, FrancoAngeli, 2009.
- [54] S. Rabbii, B. Wooley, “ A 1.8-V Digital-Audio Sigma–Delta Modulator in 0.8 $\mu$ m CMOS”, Journal of Solid State Circuits, Vol. 32, No. 6, June 1997.
- [55] M. Chae, J. Kim, W. Liu, “ Fully-differential self-biased bio-potential amplifier”, Electronic Letters, November, Vol.44, No.24, 2008.
- [56] W. Kester, “ The Data Conversion Handbook”, Analog Devices, Inc., 2005.
- [57] A. Chandrakasan and B. Ginsburg, “ An Energy-Efficient Charge Recycling Approach for a SAR Converter With Capacitive DAC” , ISCAS, 2005.
- [58] Liu et al., “ A 10-bit 50-MS/s SAR ADC With a Monotonic Capacitor Switching Procedure”, IEEE Journal of Solid-StateCircuits, Vol. 45, No. 4, April 2010.
- [59] Chandrakasan et al., “ A Low-power Area-efficient Switching Scheme for Charge-sharing DACs in SAR ADCs”, CICC, 2010.
- [60] J. Lee,C. Park, “ Capacitor Array Structure and Switch Control for Energy-Efficient SAR Analog-to-Digital Converters”, IEEE International Symposium on Circuits and Systems, ISCAS 2008.
- [61] Nauta et al., “ A 10-bit Charge-Redistribution ADC Consuming 1.9  $\mu$ W at 1 MS/s”, ISSCC, 2010.
- [62] A.Agnes, E. Bonizzoni, P. Malovati, F. Maloberti, “ 12.5 A 9.4-ENOB 1V 3.8 $\mu$ W 100kS/s SAR ADC with Time-Domain Comparator”, ISSCC, 2008.
- [63] A. Agnes, F. Maloberti, “ Design of an Ultra-Low Power Time Interleaved SAR Converter”, PRIME, 2008.
- [64] I. Svensson, J. G. Koller, “ Driving a Capacitive Load Without Dissipating  $CV^2$ ”, IEEE Symposium on Low Power Electronics, 1994.
- [65] D. Johns, K. Martin, “ Analog Integrated Circuit Design”, Wiley and Sons, 1997.
- [66] I. Svensson, J. G. Koller, “ Adiabatic charging without inductors”, IEEE Symposium on Low Power Electronics , 1994.
- [67] A. Shikata, R. Seikimoto, H. Ishikuro, “ A 0.5V 65nm-CMOS Single Phase Clocked Bootstrapped Switch with Rise Time Accelerator” , IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), 2010.
- [68] J. Rabaey, A. Chandrakasan, B. Nikolic, “ Circuiti Integrati Digitali, l’Ottica del Progettista”, Prentice Hall, II ed., 2005.
- [69] Nauta et al. , “ A Low-Offset Double-Tail Latch-Type Voltage Sense Amplifier”, IEEE Transactions on VLSI Systems, 2009.

- [70] B. Razavi and B. A. Wooley, "Design techniques for high-speed, high-resolution comparators", *Solid-State Circuits, IEEE Journal*, vol.27, no.12, pp.1916-1926, Dec. 1992.
- [71] Harpe et al. , "A  $26\mu\text{W}$  8 bit 10 MS/s Asynchronous SAR ADC for Low Energy Radios", *IEEE Journal of Solid State Circuits*, 2011.
- [72] P. M. Figueiredo and J. C. Vital, "Low kickback noise techniques for CMOS latched comparators", *Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium*, vol.1, pp.I- 537-40 Vol.1, May 2004.
- [73] A. Chandrakasan, B. Ginsburg, "A 500-MS/s 5-bit ADC in 65-nm CMOS With Split Capacitor Array DAC", *JSSCC*, 2007.
- [74] J. Kim, S. Lim, " Design of a 12-b Asynchronous SAR CMOS ADC", *ISOC*, 2011.