



HAL
open science

Nouvelles techniques d'appariement dynamique dans un CNA multibit pour les convertisseurs sigma-delta

Esmail Najafi Aghdam

► **To cite this version:**

Esmail Najafi Aghdam. Nouvelles techniques d'appariement dynamique dans un CNA multibit pour les convertisseurs sigma-delta. Micro et nanotechnologies/Microélectronique. Université Paris Sud - Paris XI, 2006. Français. NNT : . tel-00278389

HAL Id: tel-00278389

<https://theses.hal.science/tel-00278389>

Submitted on 13 May 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**UNIVERSITÉ PARIS XI
UFR SCIENTIFIQUE D'ORSAY**

Thèse

Spécialité : Physique

présentée par

NAJAFI AGHDAM Esmael

pour obtenir

le GRADE de DOCTEUR en SCIENCES de
l'École Doctorale

"Sciences et Technologies de l'Information des Télécommunications et des Systèmes"

Sujet :

*NOUVELLES TECHNIQUES D'APPARIEMENT DYNAMIQUE DANS UN CNA MULTIBIT
POUR LES CONVERTISSEURS SIGMA-DELTA*

soutenance prévue le 30 Juin 2006 devant le jury composé de :

M. AMARA	Amara	Examineur
M. BENABES	Philippe	Directeur de thèse
M. DALLET	Dominique	Rapporteur
M. GAFFIOT	Frédéric	Examineur
M. GILLES	Jean-Paul	Président
M. LOUMEAU	Patrick	Rapporteur

*Je dédie ce mémoire à mon père
dont les mots encourageants étaient toujours une source d'inspiration . . .*

*. . . à ma mère :
il n'existe pas de rôle aussi important que le sien . . .*

*. . . à ma belle mère et à son fils Asgar
martyr dans la défense de mon pays l'IRAN en 1987*

Remerciements

Ce travail de thèse a été réalisé au département Signaux et Systèmes Électroniques (*SSE*) de SUPELEC. Je remercie Monsieur Jacques OKSMAN, chef du département, pour son accueil au sein de l'équipe, pour son encouragement et son soutien moral.

J'ai bénéficié d'une bourse d'étude pour cette thèse de la part de l'université de technologie SAHAND, Tabriz, j'en suis très reconnaissant.

Je remercie Monsieur Jean-Paul GILLES, Professeur à l'Institut d'Électronique Fondamentale (*IEF*) de l'université Paris 11, d'avoir présidé mon jury de thèse. J'exprime ma reconnaissance à Monsieur Dominique DALLET, Professeur au Laboratoire d'Étude de l'Intégration des Composants et Systèmes Électroniques (*IXL*), et à Monsieur Patrick LOUMEAU, Maître de Conférence à l'École Nationale Supérieure des Télécom (*ENST*), habilité à diriger des recherches, d'avoir accepté de juger ce travail en qualité de rapporteurs. J'adresse mes remerciements à Monsieur Frédéric GAFFIOT, Professeur à l'École Centrale de Lyon (*ECL*), et Monsieur Amara AMARA, Maître de Conférence à l'Institut Supérieur d'Électronique de Paris (*ISEP*), habilité à diriger des recherches, d'avoir fait partie du jury en qualité d'examineurs.

Je remercie tout particulièrement Monsieur Philippe BÉNABES, mon directeur de thèse, Professeur au département *SSE*, pour son soutien efficace et sa disponibilité tout au long de cette étude. Je remercie également Monsieur Richard KIELBASA, Professeur au département *SSE* et responsable du groupe Systèmes Électroniques, de m'avoir accepté au sein de son groupe de recherche et pour les nombreuses discussions que j'ai eues avec lui qui m'ont permis d'améliorer mon travail.

Je ne saurais exprimer toute ma gratitude à Monsieur Philippe PORQUET, attaché scientifique, pour avoir accepté de relire et de corriger ce rapport.

Je voudrais bien évidemment remercier tous les collègues du département, tout particulièrement Monsieur Jérôme JUILLARD, Mademoiselle Sylvie GUESSAB, Madame Caroline LELANDAIS-PERRAULT, Monsieur Alain GAUTIER pour leur soutien indéfectible, Monsieur Francis TRELIN pour son soutien technique et sa bonne humeur et Mademoiselle Fabienne SURAUD pour son extrême gentillesse. Ma reconnaissance va à Madame Sorore BENABID qui a récemment soutenu sa thèse au département *SSE*, pour son aide sur le plan technique que sur le plan humain. Ce travail a été grandement facilité grâce à l'ambiance qui règne au département *SSE* et plus généralement à SUPELEC. Je remercie l'ensemble des thésards de contribuer à celle-ci : Davud, Yoan, Tudor, Émilie, Ali, Éric, Alexia. Je leur souhaite bonne chance et bonne continuation.

Cette thèse est le résultat d'un parcours inhabituel qui a mêlé ma vie professionnelle et ma formation universitaire de nombreuses années. Les mots ne peuvent pas exprimer ma reconnaissance à mon épouse, Nassin YEKROUI. Elle m'a soutenu de tant de manières. Je lui dois une dette éternelle pour ses encouragements qu'elle seule était à même de me donner. Je voudrais également remercier mon fils et ma fille pour leur gentillesse. J'attends avec impatience le moment de rendre à ma famille les nombreuses heures qui leur ont manquées. J'ai une dernière dette naturelle à l'égard de mes parents, sans lesquels je ne serais pas là, sans oublier mes sœurs et aussi mon cher frère Jabraeil, pour le soutien affectif qu'ils m'ont manifesté en toutes circonstances.

Résumé

Les convertisseurs analogiques-numériques fondés sur le principe de la modulation $\Sigma\Delta$ sont capables de fonctionner à des résolutions très élevés. L'utilisation en interne d'un CAN et d'un CNA multibit permet de réduire le taux de suréchantillonnage, les contraintes imposées par les circuits actifs, améliore la stabilité de la boucle du modulateur, mais rend celui-ci très sensible aux imperfections des composants du convertisseur numérique analogique (CNA) interne situé dans le chemin de retour. Les erreurs statiques dues aux non idéalités des circuits constitutifs de ce CNA peuvent être corrigées au moyen de techniques d'appariement dynamique des composants (DEM).

Ce travail de thèse est consacré entre autre à l'étude théorique de ces techniques de correction des défauts des cellules des CNA multibits. Après avoir rappelé le principe de la conversion $\Sigma\Delta$ d'une part, et les différentes sources d'erreurs dominantes dans le cas multibit d'autre part, les techniques d'appariement existantes sont analysées et comparées. Nous soulignons les avantages, les inconvénients, et les domaines d'applications préférentiels de chacune.

Le cœur du travail consiste en la proposition de quatre nouvelles techniques d'appariement dynamique. Les deux premières dérivent de la méthode de la moyenne des données (DWA), l'une pour le cas passe-bas du premier ordre, l'autre dans le cas passe-bande du second ordre. Les deux dernières propositions (appelées MDEM et STDEM) dérivent des deux algorithmes de tri (SDEM) et d'arborescence (TDEM) : elles conviennent à une mise en forme des erreurs d'ordre élevé et sont destinées aux applications passe-bas et passe-bande de haute performance. Ces quatre méthodes proposées ont été mises en équation et leurs performances confirmées par diverses simulations.

Une implantation des algorithmes MDEM et STDEM a été faite au niveau cellule standard jusqu'à l'étape finale de routage en technologie CMOS $0.35 \mu m$. L'ensemble des résultats des simulations au niveau système et au niveau transistor confirme l'avantage des techniques développées dans ce travail en termes de surface occupée et aussi de fréquence maximale d'application, si on les compare avec les algorithmes conventionnels de SDEM.

Dans une dernière partie, les erreurs dynamiques du CNA, en particulier l'effet de la gigue d'horloge, le glitch, la dissymétrie des temps de transition, l'injection de charge (CFT) et la métastabilité du quantificateur sont également analysés. A l'issue de ces réflexions, une nouvelle cellule de CNA incluant un bloc limitant la plage dynamique de la commande d'entrée (SRD) est proposée. Elle possède une structure de remise à zéro partielle (semi-RZ) qui permet de bénéficier à la fois de l'avantage de la cellule RZ et non RZ. De plus, l'effet du retard du bloc de DEM est compensé par une modification dans l'architecture convenant aux applications passe-bande haute fréquence.

*** **

Keywords : appariement dynamique, modulation sigma delta, multibit, CNA, CAN, DAC, ADC, DEM, STDEM, MDEM, TDEM, SDEM, MP-DWA.

Title

Novel dynamic element matching techniques for multibit DAC used in a delta-sigma modulators

Abstract

$\Delta\Sigma$ analog-to-digital converters can reach very high resolutions. The use of a multibit quantizer and of a multibit DAC reduces the oversampling ratio and the constraints imposed by active circuits, and improves the modulator loop stability. However, a multibit modulator is very sensitive to imperfections of its internal DAC located in feedback path. The static errors due to the non idealities of the DAC circuits can be corrected by means of dynamic element matching techniques (DEM).

The aim of this PhD thesis is the theoretical study of several mismatch shaping techniques for a multibit feedback DAC used in delta-sigma modulators. After a presentation of the principle of $\Delta\Sigma$ conversion and the main error sources in the multibit case, the existing DEM techniques are analyzed and compared. We underline advantages, disadvantages, and preferred application for each of them. The main part of this work consists in the development of four new DEM techniques. The first two derive from data-weighted averaging (DWA), one for first-order lowpass application, and the other for the case of second-order bandpass application. The last two proposals (called MDEM and STDEM) derive from the vector-based (SDEM) and tree-structured (TDEM) algorithms: they are more suitable for a high-order mismatch-shaping algorithm for both lowpass and bandpass applications. These four suggested methods are put in equation and their performances are confirmed by system-level simulations.

A standard cell-level implementation of MDEM and STDEM techniques is performed in CMOS $0.35\mu m$ technology. Simulations confirm the advantage of the developed techniques in terms of area and maximum frequency, compared with the conventional SDEM algorithms.

In the last part, the DAC's dynamic errors such as clock jitter effects, glitch, non symmetrical transition times and clock-feedthrough (CFT), as well as quantizer metastability, are also analyzed. A new kind of DAC cell including a swing-reduced driver (SRD) is proposed. It has a partial return to zero structure (semi-RZ) which makes it possible to benefit at the same time from the advantages of RZ and non-RZ cells.

Moreover, the DEM delay effect is compensated by a modification in the modulator architecture for high frequency bandpass applications.

*** **

Keywords : dynamic element matching, delta sigma modulation, multibit, mismatch shaping, DAC, ADC, DEM, STDEM, MDEM, TDEM, SDEM, MP-DWA.

Abréviations:

<i>AOP</i>	Amplificateur opérationnel, <i>Opamp</i>
<i>BP</i>	Bandpass, Passe-bande
<i>BW</i>	Bande passant, Bandwidth
<i>CAN</i>	Convertisseur Analogique-Numérique
<i>CMOS</i>	Complementary Metal Oxide Semiconductor
<i>CNA</i>	Convertisseur Numérique-Analogique
<i>CT</i>	à temps continu, Continuous Time
<i>DEM</i>	Dynamic Element Matching, Appariement dynamique des composantes
- <i>RDEM</i>	Randomizing Dynamic Element Matching
- <i>SDEM</i>	Sorting Dynamic Element Matching (ou <i>Vector-Based DEM</i>)
- <i>TDEM</i>	Tree-structured Dynamic Element Matching
- <i>STDEM</i>	Shortened Tree-structured Dynamic Element Matching
- <i>DWA</i>	Data-Weighted Averaging Dynamic Element Matching
- <i>RDWA</i>	Rotated Data-Weighted Averaging Dynamic Element Matching
- <i>RnDWA</i>	Randomizing Data-Weighted Averaging Dynamic Element Matching
- <i>P – DWA</i>	Partitioned Data-Weighted Averaging Dynamic Element Matching
- <i>MP – DWA</i>	Modified Partitioned Data-Weighted Averaging Dynamic Element Matching
- <i>IDWA</i>	Incremental Data-Weighted Averaging Dynamic Element Matching
- <i>DIA</i>	Double Index Data-Weighted Averaging Dynamic Element Matching
<i>DR</i>	Dynamic Range
<i>DSP</i>	Densité Spectrale de Puissance
<i>DT</i>	à temps discret, Discret Time
<i>ENOB</i>	Nombre effectif de bit, Effective Number Of Bits
<i>GBW</i>	Produit gain- bande
<i>HZ</i>	Half-delay Zero
<i>INL</i>	Integral Non Linearity
<i>LP</i>	Lowpass, Passe-bas
<i>LSB</i>	Less Significant Bit
<i>MASH</i>	Multi stAge noise SHaping
<i>MOSFET</i>	Metal Oxide Semiconductor Field-Effect Transistor (ou <i>MOS</i>)
<i>MSB</i>	Most Significant Bit
<i>MSCL</i>	Multi Stage Closed Loop modulators
<i>MTF</i>	Mismatch Transfer Function
<i>NRZ</i>	Non Return to Zero
<i>NTF</i>	Noise Transfert Function
<i>OSR</i>	Oversampling Ratio
<i>RZ</i>	Return to Zero
<i>rms</i>	Racine carrée de la moyenne, Root-Mean-Square value
<i>SAH</i>	échantillonneur-Bloquer, Sample-And-Hold
<i>SFDR</i>	Spurious Free Dynamic Range
<i>SNDR</i>	Signal to Noise and Distortion Ratio
<i>SNR</i>	Signal to Noise Ratio
<i>SRD</i>	Swing Reduce Driver
<i>STF</i>	Signal Transfert Function
<i>THD</i>	Total Harmonic Distortion
$\Sigma\Delta$	Sigma Delta
<i>var(u)</i>	Variance de la variable <i>u</i>

Notations

α_i	Erreur de gain de la cellule numéro i
$\bar{\alpha}$	Gain moyen du CNA interne
B	Résolution du CAN et CNA de la boucle du modulateur sigma-delta
CC	Cancelling-Cycle, Cycle d'annulation d'erreur d'appariement des cellules
Δ	Pas de quantification
ΔV_{max}	Dynamique d'entrée du convertisseur
Do_t	Signale analogie de sortie du CNA
$e(n)$	Erreur du CNA désapparié
ϵ_i	Erreur de gain de la cellule numéro i
ϵ	Offset final du CNA
f_0	Fréquence centrale du modulateur
f_e	Fréquence d'échantillonnage
f_N	Fréquence de <i>Shannon</i> ; Fréquence de <i>Nyquist</i>
$F(z), H(z), G(z)$	transmittance en temps discret du filtre de la boucle du modulateur
$G(s), H(s)$	transmittance en temps continu du filtre de la boucle du modulateur
g_m	transconductance d'un transistor <i>MOS</i>
G_m	Gain d'un amplificateur à transconductance
M	Nombre de cellules du CNA interne
n	Index de l'échantillonnage (après n période de T_e)
k	Numéro de couche dans la structure en arbre
K_B	Coefficient de Boltzmin
$q(n)$	Bruit de quantification à la période d'index n
q_{rms}	<i>rms</i> du bruit de quantification $q(n)$
Q	Facteur de qualité du résonateur
σ_q	Écarte-type de bruit d'échantillonnage
$sv(n)$	Selected-Vector, Vecteur-Sélectionné pour activer une cellule de CNA
$v(n)$	Signale de sortie du modulateur
t_{on}	Temps de transition de l'état "off" à l'état "on"
t_{off}	Temps de transition de l'état "on" à l'état "off"
T_e	Periode d'échantillonnage
T	Température de Kelvine, en degré
V_{dd}	Tension d'alimentation

Table des matières

1	Introduction	1
1.1	Définition du problème et motivations	2
1.2	Objectif de la recherche	3
1.3	Organisation du mémoire	4
2	Conversion analogique numérique Sigma-Delta	6
2.1	Introduction	6
2.2	Conversion classique	7
2.3	Concepts élémentaires de la conversion sigma-delta	9
2.3.1	Suréchantillonnage et mise en forme du bruit de quantification	9
2.3.2	Structure du modulateur $\Sigma\Delta$	11
2.3.3	Modèle linéaire du modulateur $\Sigma\Delta$	12
2.3.4	Stabilité du modulateur	14
2.3.5	Du passe-bas au passe-bande	17
2.3.6	Stabilité d'un modulateur $\Sigma\Delta$ passe-bande	19
2.3.7	Du modulateur sigma-delta à temps discret au modulateur à temps continu	19
2.3.8	Comparaison entre les modulateurs à temps discret et les modulateurs à temps continu	22
2.3.9	Résolution du quantificateur	24
2.3.10	Calcul général du SNR_{max} d'un modulateur $\Sigma\Delta$ passe-bas et passe-bande	25
2.3.11	Optimisation de fonction de transfert du bruit de quantification	28
2.3.12	Critères de performance du modulateur:	29
2.4	Architecture de modulateurs $\Sigma\Delta$	33
2.4.1	Architecture de base : un seul-étage rebouclage-multiple(<i>CIFB/CRFB</i>)	33
2.4.2	Architecture avec un seul rebouclage anticipatif <i>CIFF/CRFF</i>	35
2.4.3	Architecture en cascade	36
2.4.4	Structure <i>MASH</i>	38
2.4.5	Structure <i>nFOC</i>	39
2.4.6	Structures mixtes	39
2.4.7	Structure <i>MSCL</i>	40
2.4.8	Architecture en parallèle Π	41
2.4.9	Réduction des tons en utilisant un "dither" ou un pôle "Chaotique"	42
2.4.10	Réalisation intégrée d'un modulateur $\Sigma\Delta$	43
2.5	Historique et évolution des modulateurs $\Sigma\Delta$	46
2.6	Conclusion	48
3	Sources d'erreur	50
3.1	Introduction	50
3.2	Le bruit thermique et le bruit en $1/f$	52
3.2.1	Effet de bruit thermique au niveau de l'échantillonneur	53

3.2.2	Bruit thermique dans un amplificateur opérationnel (AOP)	59
3.3	Bruit des filtres	60
3.4	Non-linéarité du quantificateur (CAN)	63
3.5	Convertisseur numérique-analogique interne (CNA)	68
3.5.1	Architecture générale du CNA interne	69
3.5.2	Erreurs statiques du CNA	71
3.5.3	Performances d'un modulateur avec CNA multibit réel sans correction	74
3.5.4	Erreurs de non linéarité de la tension de sortie du CNA	77
3.5.5	Influence de l'impédance finie de sortie du CNA	77
3.5.6	Erreurs dynamiques des CNA	80
3.6	Effet de la gigue d'horloge	83
3.6.1	Effet de la gigue d'horloge dans un modulateur à temps discret	84
3.6.2	Effet de la gigue d'horloge dans un modulateur CT	86
3.7	Retard de la boucle	90
3.8	Conclusion	92
4	Méthodes d'appariement dynamique de composantes	94
4.1	Introduction	94
4.2	Structure thermométrique du CNA	95
4.3	Brassage aléatoires des sources, RDEM	97
4.4	Clock Averaging, CLK-DEM	102
4.5	Brassage de source ILA-DEM	104
4.6	Appariement dynamique DWA	107
4.6.1	DWA, solutions complémentaires	110
4.7	Brassage de sources d'ordre élevé par DWA	121
4.8	Algorithme de SDEM	121
4.9	Algorithme de TDEM	126
4.9.1	Problème d'instabilité de l'algorithme TDEM	129
4.9.2	Structure TDEM segmentée	132
4.10	Techniques dérivées	134
4.10.1	Double quantificateur	134
4.10.2	CNA série	136
4.11	Ajustement physique des éléments du CNA	137
4.12	Techniques d'étalonnage numérique	137
4.13	Conclusion	139
5	Amélioration de la linéarité du CNA, techniques proposées	141
5.1	Introduction	141
5.2	algorithme MP-DWA	141
5.3	MP-DWA passe-bande	148
5.4	algorithme MDEM	151
5.4.1	MDEM Première option	152
5.4.2	Algorithme MDEM généralisé, deuxième option	157
5.5	Méthode STDEM	162
5.5.1	Description des circuit de STDEM	168
5.6	Conclusion	174

6	Implémentation de l'étage CAN/CNA	176
6.1	Introduction	176
6.2	Fonctionnement idéal du modulateur	177
6.3	Effet des erreurs du CNA	181
6.3.1	Compensation des erreurs statiques d'un CNA multibit	182
6.3.2	Implantation de méthode DEM au niveau circuit	183
6.3.3	Compensation des erreurs dynamiques d'un CNA multibit	189
6.4	Compensation des erreurs dynamiques du CAN multibit	198
6.5	Conclusion	202
	Conclusion générale	203
7	Conclusion générale et perspectives	203
	Références	206

Chapitre 1

Introduction

Les convertisseurs analogiques numériques (A/N) et numériques analogiques (N/A) sont présents dans tous les équipements de traitement du signal (acoustique, numérisation d'image et de parole, transmission des données, instrumentation de mesure, biomédical, etc.). Ils constituent de fait les parties les plus critiques des systèmes électroniques actuels et cela persistera dans le futur parce que la numérisation des systèmes électroniques progressera encore plus. Plusieurs types de convertisseurs peuvent être utilisés selon les applications. Ils se différencient par leurs performances : précision, vitesse de traitement, consommation et prix.

Les convertisseurs A/N de type *flash* fonctionnant à la fréquence de *Nyquist* permettent de traiter des signaux rapides. Afin d'accroître leur précision, il est nécessaire d'utiliser un filtre très sévère destiné à éviter les phénomènes de repliement de spectre avant le bloc d'échantillonnage, et de quantifier avec une résolution élevée. La réalisation de ces deux fonctions analogiques augmente la complexité des circuits, donc leur sensibilité au bruit, ainsi que la surface qu'ils occupent ce qui par conséquent augmente le coût de réalisation. Pour des applications demandant une grande précision, on préfère utiliser des convertisseurs A/N à suréchantillonnage sigma-delta ($\Sigma\Delta$). Leur principale particularité provient d'une part de l'utilisation d'un filtre d'antirepliement moins sévère en raison du suréchantillonnage, et d'autre part du rejet du bruit de quantification en dehors de la bande utile grâce à une rétroaction autour du quantificateur. Ces convertisseurs sont composés d'une partie analogique appelée modulateur $\Sigma\Delta$ délivrant une sortie suréchantillonnée sur un faible nombre de bit, suivie d'une partie numérique réalisant la décimation et un filtrage numérique éliminant le bruit de quantification en dehors de la bande de fréquence utile.

Le modulateur $\Sigma\Delta$ est une boucle d'asservissement qui comprend un filtre, un quantificateur et un CNA sur le chemin de retour. Plus les caractéristiques (ordre du filtre de boucle L , facteur de suréchantillonnage OSR et résolution du quantificateur B) du modulateur présentent des valeurs élevées, meilleure est la précision du convertisseur. Cependant, pour préserver la stabilité de la boucle, respecter les contraintes de réalisation tout en maintenant la fréquence d'échantillonnage à une valeur raisonnable et atteindre la résolution recherchée, un compromis doit être réalisé entre ces différents paramètres. Dans une technologie donnée et pour une largeur de bande donnée, l' OSR est assujéti à la limite de fréquence maximale de technologie. Afin de préserver la stabilité, on ne peut gagner plus

en augmentant l'ordre du filtre de boucle de sorte qu'il est en pratique limité à quatre ($L \leq 4$ en passe-bas et $L \leq 8$ en passe-bande). En revanche, la qualité du modulateur (précision, vitesse de conversion, faible consommation, etc.) peut être améliorée en employant une quantification multibit [1–5].

1.1 Définition du problème et motivations

Malgré les avantages remarquables de la quantification multibit, la linéarité du modulateur dépend de celle du CNA sur le chemin du retour. Beaucoup d'architectures de CNA multibit sont fondées en général sur des cellules appariées, par exemple des sources de courants. Cependant, en raison de la variation dans le processus de la fabrication des circuits (par exemple : défaut d'alignement de masque, épaisseur non uniforme d'oxyde, et densités non uniformes de dopage), du gradient de la température, du vieillissement, des composants parfaitement appariés sont pratiquement impossibles à fabriquer. Ainsi, les erreurs de disparité qui sont définies comme l'expression de la différence entre les valeurs des composants théoriques et des composants réels, sont inévitables. Les défauts d'appariement des cellules du CNA font apparaître des fréquences parasites dépendant du signal d'entrée et augmentant le niveau du bruit dans la bande utile. Les effets des défauts d'appariement d'un CNA multibit dégradent fortement les performances du système. Ces erreurs doivent être corrigées de sorte que la linéarité du CNA utilisé soit meilleure que la résolution finale du système souhaitée, ce qui est une tâche difficile.

Plusieurs techniques, y compris une technique de routage spéciale (*special-Layout*), l'ajustement par rayon laser, l'étalonnage (*calibration*), peuvent réduire les effets des erreurs d'appariement des cellules du CNA, mais, elles sont assez complexes et coûteuses. Une autre méthode, qui peut réduire les effets des erreurs des composants désappariés, est l'appariement dynamique des éléments (DEM) [6]. La technique DEM, qui a récemment été l'objet de plusieurs développements, est basée sur divers algorithmes de brassage des sources unitaires. Elle réarrange dynamiquement les interconnexions des composants désappariés de sorte que les moyennes des erreurs des composants soient égales ou presque égales dans la position de chacun des composants. Par conséquent, les erreurs de non linéarité de type statique peuvent être convenablement éliminées, réduites, ou rejetées en dehors de la bande utile. La méthode la plus simple de DEM change des interconnexions des cellules du CNA d'une manière aléatoire, ce qui convertit des tons dépendant du signal d'entrée en bruit blanc dont une partie est présente dans la bande [7]. D'autres méthodes de DEM (par exemple : *CLK* [8], *ILA* [9], *DWA* [10], *TDEM* [11], *SDEM* ou *vector-based mismatch-shaping* [12]) tentent de mettre en forme les erreurs d'appariement du CNA au moyen d'un filtrage d'ordre égal à un ou supérieur. Parmi les différentes méthodes qui supportent bien la mise en forme des erreurs d'appariement passe-bas d'ordre un, la méthode basée sur la moyenne de données *DWA* est à la fois simple et performante, mais souffre de la présence de tons parasites vis à vis des entrées périodiques ou DC. Certaines techniques supplémentaires peuvent compenser ce défaut présent dans la *DWA*. Cependant, seules deux méthodes existantes, *SDEM* et *TDEM*, peuvent être utilisées pour des applications d'ordre élevé passe-bas et passe-bande. La *SDEM* devient complexe pour un grand nombre de bits de quantification

($B > 2$). La TDEM souffre de la sensibilité de son algorithme aux caractéristiques stochastiques des entrées du CNA, par conséquent elle devient instable en pratique pour les applications dont l'ordre est élevé.

1.2 Objectif de la recherche

Ce travail de thèse a pour but d'étudier les différentes sources d'erreur dues à la non linéarité du modulateur $\Sigma\Delta$ multibit et aux méthodes possibles de correction, en particulier celles utilisant un CNA interne multibit.

Les éventuelles erreurs d'un modulateur à haute performance peuvent être de nature statique ou dynamique. Les éventuelles techniques de correction dérivent de la nature de l'application et de l'architecture du modulateur. Après avoir analysé les sources de non linéarité d'un modulateur $\Sigma\Delta$ réalisé en technologie CMOS, nous nous limiterons aux erreurs statiques et dynamiques les plus importantes dans le CNA utilisé dans un modulateur sigma delta multibit. Le travail actuel est consacré à étudier les différents aspects des diverses méthodes de DEM existantes ainsi qu'à développer des algorithmes plus efficaces, en particulier ceux qui concernent les méthodes de DEM d'ordre élevé. On cherche alors à mettre en évidence les contraintes liées aux méthodes de DEM et à des solutions éventuelles pour lever ces contraintes.

En ce qui concerne les erreurs dynamiques du modulateur, nous nous efforçons de limiter les effets de la gigue d'horloge et des phénomènes transitoires du CNA (comme le *glitch*, le *CFT* et le temps de transition non-équilibré) qui s'aggravent en général quand on utilise une méthode de DEM. De plus, afin d'améliorer la vitesse du quantificateur, une nouvelle architecture est proposée : elle permet de multiplier par deux la fréquence maximale d'un modulateur à temps continu sans réduire la performance du système.

Au cours de ce travail les points suivants seront développés :

- analyse des différentes sources d'erreurs statiques et dynamiques présentes dans un modulateur $\Sigma\Delta$ multibit,
- étude comparative des techniques de brassage de sources (DEM) existantes,
- développement de quatre nouveaux algorithmes pour : DEM passe-bas du premier ordre, DEM passe-bande du deuxième ordre, DEM passe-bas et passe-bande d'ordre élevé; ces nouveaux algorithmes bénéficient des points forts de certains algorithmes existants tout en évitant leurs points faibles. La performance de chacune de ces techniques est analysée analytiquement et validée ensuite par simulation au niveau système et au niveau transistor jusqu'à l'étape finale de routage en technologie CMOS $0.35\mu m$.
- compensation des erreurs dynamiques du CNA par des circuits proposés comme le "*Deglitcher*" et le "*Swing-Reduced-Deriver*" ou *SRD*,
- proposition d'une cellule de CNA de type *Semi-RZ*, apte à limiter l'effet de la gigue d'horloge,
- proposition d'une nouvelle architecture du CAN de type entrelacé, capable de repousser la limite de la fréquence maximale d'un modulateur $\Sigma\Delta$ à temps continu jusqu'à un facteur deux.

1.3 Organisation du mémoire

Les cinq chapitres constituant ce mémoire, après ce premier chapitre d'introduction, sont décrits brièvement ci-dessous :

Le chapitre 2 introduit les notions élémentaires de la modulation $\Sigma\Delta$, notamment le principe du suréchantillonnage et de mise en forme du bruit de quantification. Les différents paramètres caractérisant la structure d'un modulateur $\Sigma\Delta$ ainsi que les critères évaluant ses performances et sa stabilité y sont décrits. Les différentes architectures du modulateur $\Sigma\Delta$ sont ensuite présentées en passant en revue leurs particularités mais aussi leurs inconvénients en théorie comme en pratique. La modulation $\Sigma\Delta$ passe-bande à temps continu est étudiée. On rappellera les principes de la transformation passe-bas/passe-bande et de l'équivalence temps discret/temps continu.

Le chapitre 3 est consacré à la présentation et l'analyse des différentes sources d'erreur qui limitent la performance d'un modulateur $\Sigma\Delta$ multibit, comme le bruit thermique, la non idéalité des circuits analogiques et leurs éléments parasites. On insistera sur les erreurs statiques et dynamiques du CNA interne, le défaut d'appariement des cellules, le "*glitch*", la gigue d'horloge, l'effet du temps de transition non équilibré, l'effet d'injection du charge au sein du commutateur de courant (CFT) et l'impédance finie des sources de courant.

Le chapitre 4 a pour but de passer en revue les techniques de correction de défaut d'appariement les plus appropriées pour un CNA thermométrique employé dans un modulateur $\Sigma\Delta$. Nous nous limiterons aux diverses méthodes de brassage de sources. Le principe, les avantages, les inconvénients et les éventuelles solutions existantes de chaque méthode de DEM y sont détaillés.

Le chapitre 5 est consacré à la présentation des quatre nouvelles méthodes de DEM développées au cours de ce travail de thèse. Dans un premier temps, nous décrivons les deux algorithmes de brassage de sources qui conviennent le mieux pour les DEM du premier ordre passe-bas et passe-bande. Ensuite, nous introduisons les autres algorithmes destinés aux applications DEM d'ordre 2 et 3 passe-bas ainsi qu'aux applications DEM d'ordre 4 et 6 passe-bande. La performance de tous ces algorithmes est mise en équation puis validée par simulation.

Le chapitre 6 présente au niveau transistor les différents blocs numériques et analogiques constituant un CNA apparié, utilisé dans un modulateur $\Sigma\Delta$ en technologie *CMOS*, à savoir l'algorithme de brassage, le "*deglitcher*", le SRD, et la cellule du CNA à base de courants commutés. Dans ce chapitre, une nouvelle structure de CAN est présentée. Elle permet d'atteindre une fréquence d'échantillonnage égale à $1.2GHz$, ce qui a alors rendu possible l'intégration d'un convertisseur passe-bande autour de $300MHz$ en technologie *CMOS* $0.35\ \mu m$. Les résultats de la simulation sur une implémentation d'un modulateur $\Sigma\Delta$ multibit passe-bande du 6^{ième} ordre employant une méthode de DEM en technologie *CMOS* $0.35\ \mu m$ sont présentées. Ces simulations ont été effectuées dans l'environnement "*Cadence*". On a également utilisé des outils "*Verilog*" et "*Ambit*" qui nous ont servi à déterminer les routages optimaux pour les parties numériques. Elles mettent en évidence la fréquence de fonctionnement maximale ainsi que la surface qui est nécessaire à la mise en œuvre des différentes méthodes de DEM. Les simulations montrent que les méthodes DEM proposées dans ce travail de recherche né-

cessitent moins de circuits et fonctionnent plus rapidement que les méthodes conventionnelles. L'une d'entre elles peut être cadencée à une fréquence $330MHz$, d'environ 50% plus élevée que la méthode existante pour une application DEM passe-bande d'ordre 4.

Enfin, une conclusion générale dresse un bilan de nos contributions principales dans cet axe de recherche, et des perspectives dans le domaine de la conversion A/N à grande précision et à large bande.

Chapitre 2

Conversion analogique numérique Sigma-Delta

2.1 Introduction

De nombreuses fonctions de traitement du signal sont réalisées non plus par des circuits analogiques, mais par des circuits numériques. Ceux-ci sont en effet de plus en plus performants car la technologie VLSI (*Very Large Scale Integration*) permet leur implantation sur des surfaces très denses avec un faible coût. Cependant, il est toujours nécessaire d'effectuer une interface entre le monde analogique et le monde numérique. Le développement du numérique entraîne donc des besoins en conversion analogique-numérique qui devra être d'autant plus précis que les étapes suivantes, notamment dans les domaines de la mesure, de l'audio-vidéo et des télécommunications nécessitent des résolutions importantes. Or, il est avéré que les convertisseurs classiques n'atteignent pas des performances élevées sur une surface réduite sans un ajustement coûteux. C'est donc pour obtenir une grande résolution, tout en permettant une forte intégration, qu'est désormais utilisée la conversion sigma-delta ($\Sigma\Delta$).

Les convertisseurs sigma-delta forment une famille de convertisseurs analogiques-numériques à haute résolution dont les caractéristiques sont particulièrement attrayantes. Leur principal intérêt provient du fait que le bruit de quantification est étalé en dehors de la bande utile par un phénomène de mise en forme. Ces convertisseurs échantillonnent le signal analogique à une fréquence très élevée en comparaison avec les convertisseurs opérant à la fréquence de *Nyquist* f_N , et le signal numérique de sortie est restitué à une cadence plus faible.

Durant ces dernières années, de nombreuses recherches ont été consacrées à l'étude de nouvelles architectures des convertisseurs $\Sigma\Delta$ afin d'améliorer leurs performances. Les convertisseurs $\Sigma\Delta$ conventionnels sont implantés en temps discret et fonctionnent en bande de base; on parle alors de conversion $\Sigma\Delta$ passe-bas. Ce concept a été étendu à des fréquences intermédiaires donnant naissance aux convertisseurs passe-bande. Cependant, il est difficile d'implanter ces convertisseurs en temps discret avec une fréquence d'échantillonnage dépassant $100MHz$. On s'intéresse depuis un certain temps aux convertisseurs $\Sigma\Delta$ multibits, dont la structure permet de surmonter les limites rencontrées en monobit, spécialement pour les applications faible consommation et large-bande.

Théoriquement, Les trois limites principales à la réalisation de modulateurs $\Sigma\Delta$, que sont la fréquence, le SNR et la dynamique d'entrée DR peuvent être améliorées si l'on augmente le nombre de niveaux de quantification (structure multibit). Malgré ces avantages très intéressants, cette structure demande une forte linéarité du convertisseur numérique-analogique (CNA) sur le chemin de rétroaction. Le problème essentiel est l'appariement des différents éléments du CNA . Ces imperfections demandent l'emploi d'au moins une méthode de correction. Un modulateur multibit à été réalisé pour le première fois en 1986 [13] dans lequel les éléments du CNA ont été ajustées de façon manuelle. Cependant, des méthodes externes et non-automatiques coûtent beaucoup trop cher pour que l'on puisse les utiliser. C'est pour cela que la plupart des convertisseurs fabriqués initialement étaient des convertisseurs monobit.

Au cœur de ces derniers années, les modulateurs multibit sont revenus à la mode grâce aux différentes méthodes de correction qui sont de plus en plus utilisables dans la partie numérique du modulateur, notamment les algorithmes de brassage des sources (*Dynamic Element Matching, DEM*). Nous nous attarderons sur ces algorithmes dans les chapitres suivants.

Après un bref rappel sur les convertisseurs classiques, ce chapitre présente les notions fondamentales de la conversion $\Sigma\Delta$. Nous décrirons tout d'abord la structure de base des modulateurs $\Sigma\Delta$, ainsi que les différents paramètres utilisés pour évaluer leurs performances. Ensuite, nous rappellerons les propriétés des modulateurs $\Sigma\Delta$ multibit en citant les avantages et les inconvénients par rapport à leurs homologues monobit. A partir de là, nous nous concentrerons plutôt sur les structures passe-bande multibit, ce qui est l'objectif de ce travail. Enfin, un bref état de l'art sera donné sur les modulateurs $\Sigma\Delta$ passe-bande multibit publiés récemment, rassemblant le nombre de bit du quantificateur interne, la méthode de correction de la non linéarité, la technologie utilisée et les performances maximales atteintes.

2.2 Conversion classique

Les convertisseurs analogique numérique (CAN) effectuent la transition entre le monde analogique et le monde numérique, en échantillonnant le signal de nature analogique et en le quantifiant sur un nombre de bit fini. La fonction général pour réaliser un CAN peut être représentée par la figure.2.1.

La chaîne de conversion se compose tout d'abord d'un filtre passe-bas qui évite le repliement de

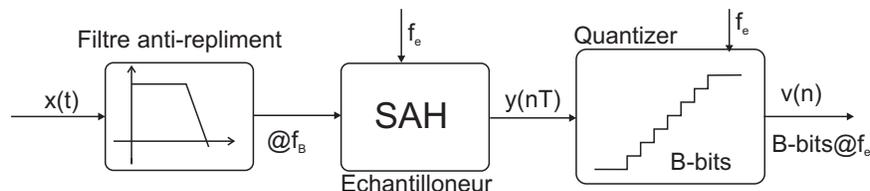


FIG. 2.1 – Schéma général d'un convertisseur analogique-numérique (CAN) classique.

spectre lors de l'échantillonnage en limitant la bande passante du signal d'entrée $x(t)$. Ensuite vient

un échantillonneur-bloqueur¹ qui mémorise la valeur du signal d'entrée pendant la phase de conversion pour permettre sa quantification, c'est-à-dire l'association à la valeur d'entrée d'un des niveaux discrets du quantificateur. La figure 2.2 montre la relation entre la sortie d'un quantificateur idéal par rapport à son entrée $y(n)$ ². La quantification est une étape essentielle de la conversion, elle entraîne l'adjonction au signal d'un bruit irréversible supposé blanc qu'on appelle bruit de quantification q . La fonction de transfert d'un quantificateur idéal est une loi en marche d'escalier uniforme qui lie l'entrée analogique au code numérique qui lui est affecté. Un quantificateur est caractérisé par un nombre de bits B (ou résolution) et une dynamique d'entrée [FS= $y_{max} - y_{min}$]. Pour une entrée en dehors de cet intervalle, la sortie du quantificateur sature. Le pas de quantification noté Δ appelé aussi quantum, désigne la plus petite variation de la tension analogique associée au bit de poids le plus faible *LSB* (*Less Significant Bit*), et est donné par l'équation 2.1. Plus le quantum Δ est faible, meilleure est la précision du convertisseur.

$$\Delta = \frac{y_{max} - y_{min}}{2^B - 1} \quad (2.1)$$

où, pour la raison de la simplicité, y_{max} est supposé égal à $-y_{min}$. Le bruit de quantification est

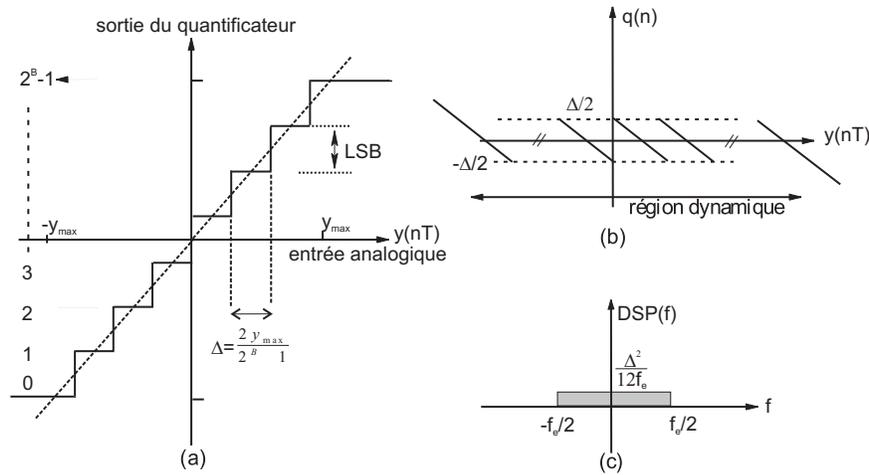


FIG. 2.2 – *Quantification: (a) fonction de transfert, (b) bruit de quantification, (c) puissance de bruit de quantification.*

donné par l'expression 2.2 :

$$q(n) = v(n) - y(n) \quad (2.2)$$

Le bruit de quantification q est un signal en dent de scie d'amplitude $\Delta/2$, réparti sur 2^B segments (figure 2.2-b). En supposant le bruit de quantification uniformément distribué dans ces segments sa puissance moyenne est égale à sa variance [1-4]. Celle-ci est définie par :

$$P_{quan.} = \delta_q^2 = Var(q) = q_{rms}^2 = \int_{-\Delta/2}^{+\Delta/2} q^2 P(q) dq = \int_{-\Delta/2}^{+\Delta/2} q^2 \frac{1}{\Delta} dq = \frac{\Delta^2}{12} \quad (2.3)$$

1. "Sample-And-Hold", SAH

2. On choisit respectivement "y" et "v" pour l'entrée et la sortie du quantificateur en imaginant que dans la sortie, $v(n)$ l'information perdue par rapport à l'entrée correspond d l'élimination de la queue de la lettre "Y" pour devenir la lettre "V".

où $P(q)$ représente la densité de probabilité du bruit.

Quand le quantum Δ est très petit devant la variation du signal, le bruit de quantification est généralement assimilé à un bruit blanc caractérisé par une densité spectrale de puissance (DSP) uniformément répartie entre $-\frac{f_e}{2}$ et $+\frac{f_e}{2}$ (f_e étant la fréquence d'échantillonnage) (fig.2.2-c). Son niveau se calcule à partir de la puissance du bruit comme suit :

$$\int_{-f_e/2}^{+f_e/2} DSP_q(f)df = q_{rms}^2 \Rightarrow DSP_q(f) = q_{rms}^2 \cdot \frac{1}{f_e} = \frac{\Delta^2}{12f_e} \quad (2.4)$$

Considérons un signal sinusoïdal d'amplitude égale à la pleine échelle du convertisseur ($2y_{max}$), et échantillonné en respectant le théorème de *Nyquist* ($f_e \geq 2 f_{max}$); le rapport signal sur bruit maximal SNR_{max} (*Signal to Noise Ratio*) vaut dans ce cas :

$$P_{sig} = \frac{1}{2} \left(\frac{y_{max} - y_{min}}{2} \right)^2 = \frac{\Delta^2 (2^B - 1)^2}{8} \quad (2.5)$$

$$SNR_{max}(dB) = 10 \log \left(\frac{P_{sig}}{P_{quan.}} \right) \simeq 6,02B + 1,76 \quad (2.6)$$

Ainsi, à partir du SNR_{max} nous pouvons déduire la résolution du convertisseur classique et vice versa. Selon cette relation indiquée ci-dessus, un bit supplémentaire rajoute 6 dB du SNR .

2.3 Concepts élémentaires de la conversion sigma-delta

Les convertisseurs sigma-delta sont basés sur une technique de suréchantillonnage et de mise en forme du bruit de quantification, ce qui permet de minimiser le bruit de quantification dans la bande de fréquence utile, et par conséquent, d'améliorer la résolution du convertisseur. La structure classique d'un convertisseur $\Sigma\Delta$, représentée sur la figure 2.3, est composée d'un modulateur suivi d'un filtre numérique de décimation. Un filtre antirepliement à temps continu est nécessaire à l'entrée du modulateur, afin d'éliminer les problèmes qui concernent au conflit spectral entrée le signal utile et les signaux en dehors de la bande ("*Aliasing*").

La théorie associée à cette technique est présentée dans le paragraphe suivant.

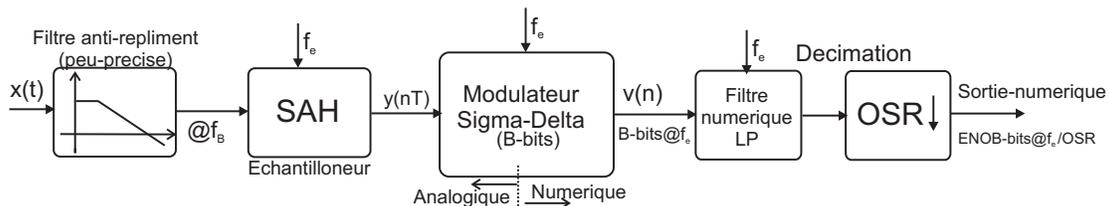


FIG. 2.3 – Chaîne de conversion analogique-numérique.

2.3.1 Suréchantillonnage et mise en forme du bruit de quantification

Quand on code un signal à très haute fréquence, bien au delà de la fréquence dite de *Nyquist*, on étale le spectre du bruit de quantification sur une plus grande gamme de fréquence, améliorant ainsi

le rapport signal sur bruit dans la bande utile. C'est donc par un échantillonnage à grande vitesse que l'on augmente la précision. L'autre avantage de cette technique est que les contraintes sur le filtre anti-repliement placé à l'entrée du convertisseur sont moins sévères.

Ce mécanisme est illustré sur la figure 2.4. Dans le cas (a), le signal est échantillonné à la fréquence de *Nyquist* (f_N). Comme cela a été mentionné dans la section précédente, la densité spectrale de puissance du bruit de quantification est uniformément répartie sur une largeur de bande égale à f_e . Si on échantillonne le même signal à une fréquence K fois supérieure à la fréquence de *Nyquist*, on va alors diviser la densité spectrale de puissance par K qui va cette fois s'étaler entre $\pm \frac{K f_N}{2}$ (dans le cas (b)), soit un rapport signal sur bruit (*SNR*):

$$\begin{aligned} q_B^2 &= \frac{q_{rms}^2}{K} = \frac{\Delta^2}{12K} \\ SNR(dB) &= 6,02B + 1,76 + 10 \log[K] \end{aligned} \quad (2.7)$$

où q_B^2 est la puissance du bruit dans la bande utile. On appelle K le rapport de suréchantillonnage qui est souvent noté *OSR* (*Over Sampling Ratio*) et défini par :

$$OSR = \frac{f_e}{f_N} = \frac{f_e}{2f_B} \quad (2.8)$$

où f_B est la bande utile du système.

Par exemple, si on utilise un *OSR* égal à 4, cela revient à diminuer le bruit de quantification dans la

Densité Spectrale de Puissance

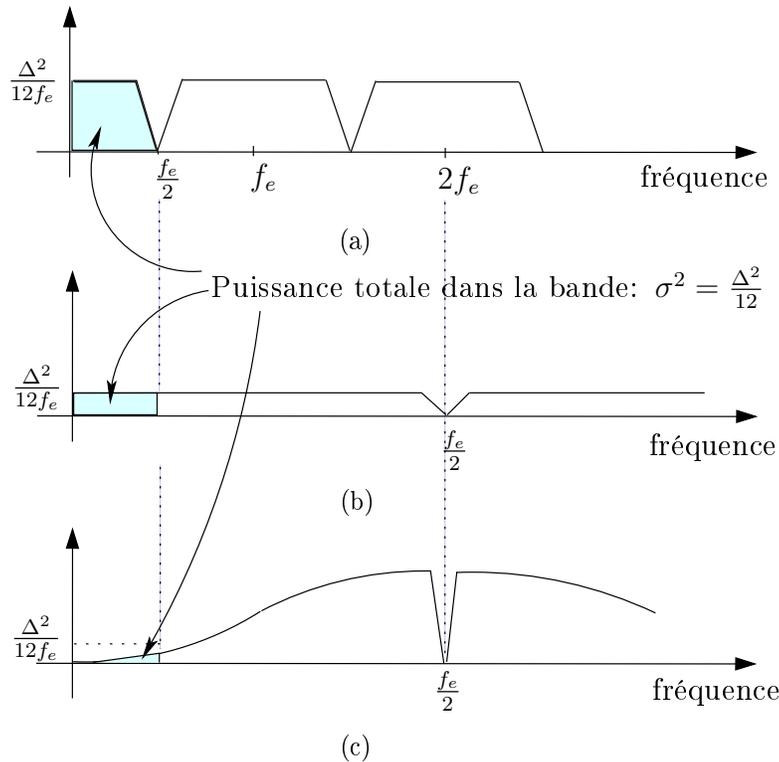


FIG. 2.4 – *Bruit de quantification dans le cas d'une conversion, (a) à la fréquence de Nyquist, (b) à suréchantillonnage, (c) sigma-delta.*

bande de 6 dB d'où un gain de 1 bit. Une augmentation de résolution de 10 bit revient à échantillonner le signal à une fréquence égale à $10^6 f_N$. Cependant, en considérant la limite de la fréquence maximum des circuits standards actuelles (par exemple *CMOS*), la bande passante du signal n'excéderait pas quelques centaines de Hertz.

Ainsi, pour obtenir une plus grande résolution sans compromettre la largeur de la bande, une opération de filtrage avec une rétroaction permet d'étaler le bruit de quantification en dehors de la bande utile, réduisant ainsi son énergie dans la bande du signal. Ceci est appelé mise en forme de bruit (*noise shaping*), et est illustré sur la figure 2.4.c. On parle alors de modulation $\Sigma\Delta$ dont le principe est décrit dans la section suivante.

2.3.2 Structure du modulateur $\Sigma\Delta$

Un modulateur $\Sigma\Delta$ code le signal analogique sur un faible nombre de bits suréchantillonnés. Sa structure générale est composée d'un filtre, d'un CAN et d'un CNA (figure 2.5). La mise en forme du bruit de quantification est réalisée en plaçant un filtre de gain élevé et une rétroaction autour du *CAN*. Cette étape est illustrée sur la figure 2.4.c.

Le filtre de boucle sert à minimiser l'écart moyen entre le signal d'entrée et sa valeur quantifiée; de

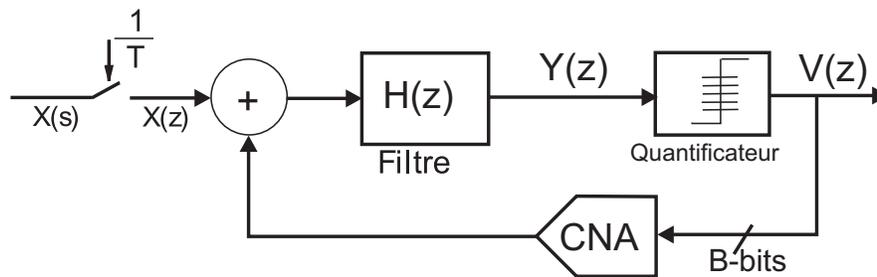


FIG. 2.5 – Modulateur $\Sigma\Delta$.

cette manière, le signal de sortie va tendre à suivre l'évolution du signal d'entrée. Ce filtre joue aussi un double rôle; comme dans toute rétroaction il assure le gain dans la boucle et détermine en plus la bande passante du bruit rejeté. Ce filtre peut être un filtre passe-bas (intégrateur) ou passe-bande (résonateur). Généralement, l'ordre du modulateur (L) est défini par celui du filtre de boucle. Celui-ci est directement lié au nombre d'intégrateurs ou de résonateurs (m) qu'il contient :

$$L = \begin{cases} m & \text{filtre passe-bas} \\ 2m & \text{filtre passe-bande} \end{cases} \quad (2.9)$$

Plus l'ordre du modulateur augmente, moins le bruit est important dans la bande ce qui améliore la résolution du modulateur. Cependant, dès que le nombre d'intégrateurs ou de résonateurs (m) est supérieur à 2, un problème de stabilité peut se poser.

Le modulateur peut être monobit en utilisant un comparateur dans la boucle ou multibit si le comparateur est remplacé par un *CAN*. Plus la résolution du *CAN* est grande, meilleure est la précision du modulateur $\Sigma\Delta$. Cependant, le nombre de niveaux du *CNA* dans le chemin de retour augmente. Ceci a pour effet d'induire des erreurs de non linéarité non négligeables qui vont dégrader les performances

globales du modulateur.

Au tout début, les modulateurs réalisés étaient limités à des modulateurs monobit de fait de leur simplicité et de leur linéarité. Aujourd'hui, la réalisation de modulateurs multibits devient une priorité pour obtenir des performances qui ne peuvent jamais être atteintes au moyen de structures monobit.

En résumé un modulateur $\Sigma\Delta$ est caractérisé par les paramètres suivants :

- le type de réalisation; discret ou continu,
- le type du filtre de boucle; passe-bas ou passe-bande,
- la résolution du CAN (B),
- le facteur de suréchantillonnage (OSR),
- l'ordre (L),
- la fréquence d'échantillonnage (f_e).

Le spectre du signal à la sortie du modulateur $\Sigma\Delta$ $v(n)$ contient le signal d'origine (x) plus le bruit de quantification mis en forme. Un filtre numérique placé à la sortie du modulateur est donc nécessaire pour supprimer le bruit de quantification en dehors de la bande et faire passer uniquement le signal dans la bande utile. En combinant le modulateur et le filtre numérique on obtient la structure d'un convertisseur $\Sigma\Delta$ qui représente l'une des architectures les plus utilisées actuellement pour la numérisation des signaux analogiques.

Il existe un autre domaine d'utilisation et d'architecture des modulateurs $\Sigma\Delta$ dans lesquelles le filtre de boucle se partage avec les caractéristique naturelle d'un capteur électrique, souvent intégré pour des systèmes mesures électroniques. Ce domaine de recherche n'entre pas dans notre travail mais l'on peut se référer aux autres documents de notre laboratoire ainsi que dans la littérature ouverte au monde de la mesure électronique et du micro-système [14].

2.3.3 Modèle linéaire du modulateur $\Sigma\Delta$

Le modèle linéaire du modulateur $\Sigma\Delta$, représenté dans la figure 2.6, consiste à substituer au quantificateur une source de bruit blanc additive et uniforme $q(n)$. Cette approximation est valable si les conditions suivantes sont vérifiées [15] :

1. Le quantificateur ne sature pas (fig.2.2.b),
2. les niveaux de quantification sont suffisamment grands et équirépartis dans l'intervalle de fonctionnement,
3. l'amplitude du signal d'entrée est uniformément répartie sur toute la plage d'utilisation.

Dans le cas d'un modulateur $\Sigma\Delta$, ces hypothèses ne sont pas totalement vérifiées, mais dans beaucoup de cas, les résultats obtenus forment un bon modèle de la réalité surtout pour des modulateurs d'ordre 2 et plus, ainsi que pour des modulateurs multibit. Pour l'instant, nous supposons que le CNA du chemin de retour est idéal sans aucun défaut d'appariement. Alors, la sortie du modulateur $V(z)$ peut s'exprimer à partir de la fonction de transfert du signal $STF(z)$ ainsi que celle du bruit $NTF(z)$.

Par :

$$V(z) = STF(z)X(z) + NTF(z)Q(z) \quad (2.10)$$

d'où $STF(z) = \frac{V(z)}{X(z)} = \frac{H(z)}{H(z) + 1}$, $NTF(z) = \frac{V(z)}{Q(z)} = \frac{1}{H(z) + 1}$

Selon le choix du filtre de boucle $H(z)$ (passe-bas ou passe-bande), La STF laisse passer le signal

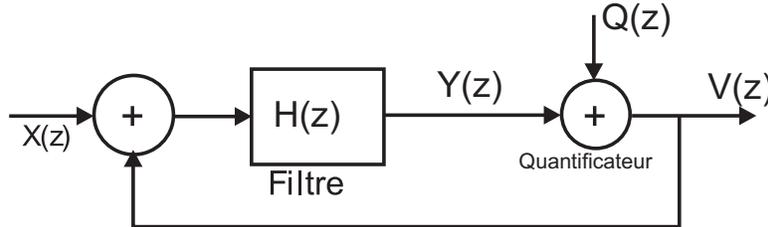


FIG. 2.6 – Modèle linéaire du modulateur $\Sigma\Delta$.

dans une certaine bande de fréquences, alors que la NTF atténue le bruit de quantification dans cette même bande. Généralement dans un modulateur d'ordre L où $L < 3$, la STF est une fonction retard de la forme z^{-L} et la NTF est une fonction coupe-bande de la forme $(1 - z^{-1})^L$ pour le cas passe-bas et $(1 + z^{-2})^{L/2}$ pour le cas passe-bande. Par exemple, un modulateur passe-bas du premier ordre peut être réalisé selon le schéma suivant (figure 2.7). La fonction de transfert du signal $STF(z)$ ainsi que celle du bruit $NTF(z)$ sont :

$$\begin{aligned} STF(z) &= z^{-1} && ; \text{ simple retard} \\ NTF(z) &= (1 - z^{-1}) && ; \text{ passe-haut} \end{aligned} \quad (2.11)$$

La figure 2.8 décrit la sortie numérique d'un modulateur passe-bas du premier ordre et son spectre. On voit bien que la probabilité de la valeur de sortie numérique est liée à l'amplitude d'entrée. De plus, le spectre de la sortie numérique du modulateur qui est une courbe passe-haut, montre le phénomène de mise en forme de bruit de quantification dans le boucle $\Sigma\Delta$.

Pour des modulateurs d'ordre plus élevé, la formule 2.10 reste toujours valable, mais l'expression des fonctions $NTF(z)$ et $STF(z)$ change. Il y existe plusieurs architectures de modulateurs. Certaines seront décrits prochainement dans ce chapitre.

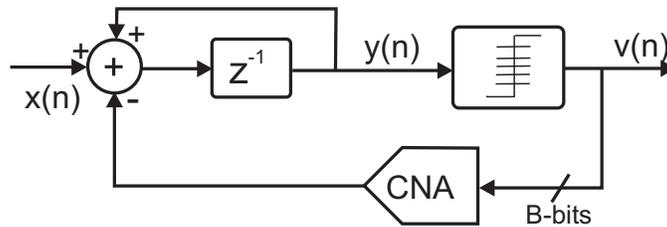


FIG. 2.7 – schématique-simple d'un modulateur d'ordre un passe-bas

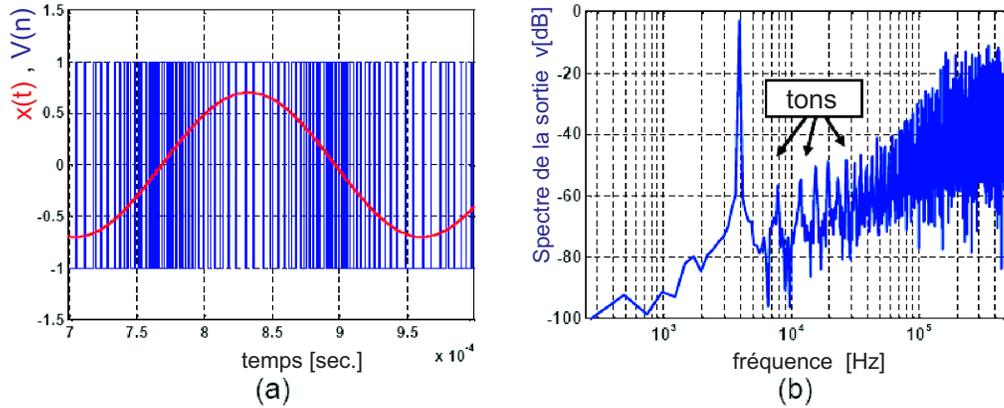


FIG. 2.8 – sortie numérique et spectre de la sortie d'un modulateur passe-bas monobit du premier ordre

2.3.4 Stabilité du modulateur

Comme tout système à contre réaction, les modulateurs $\Sigma\Delta$ peuvent être sujet à de l'instabilité : c'est-à-dire qu'ils peuvent entrer dans un régime d'oscillations entretenues et y rester, le signal de sortie ne dépendant alors plus du signal d'entrée. Par conséquent, la conception du modulateur $\Sigma\Delta$ réclame une condition de stabilité robuste. Actuellement, aucun outil mathématique ne permet de concevoir une architecture de modulateur $\Sigma\Delta$ d'ordre élevé (m supérieur à 2) qui soit stable. La raison principale est que le modulateur $\Sigma\Delta$ est un système qui présente une grande non linéarité due à la présence du quantificateur dans la boucle. Pour concevoir et analyser un modulateur $\Sigma\Delta$, on utilise habituellement un modèle simplifié du quantificateur : le quantificateur alors est interprété comme un additionneur de bruit blanc non corrélé à son signal d'entrée. Bien que cette modélisation du quantificateur ne représente pas la réalité, elle offre une évaluation suffisante des performances atteintes avec un modulateur $\Sigma\Delta$ d'ordre L en fonction de son taux de suréchantillonnage. Cependant, elle ne permet pas d'analyser son comportement non-linéaire et interdit toute étude de stabilité satisfaisante.

Plusieurs tentatives pour trouver un critère rigoureux pour définir la stabilité du modulateur $\Sigma\Delta$ ont été proposées dans la littérature [2, 16–19]. Quelques critères sont rappelés ci dessous :

Le critère BIBO (*Bounded Input Bounded Output*): un modulateur est stable si le filtre de boucle produit des oscillations bornées pour toute entrée bornée. Il est totalement instable si ces oscillations divergent pour un signal d'entrée nul. Cette condition devient plus difficile à réaliser quand l'entrée est proche de la valeur pleine échelle du quantificateur, puisque la saturation augmente la nature non linéaire du modulateur. Ce phénomène est accentué quand l'ordre du modulateur est grand (nombre d'intégrateurs ou de résonateurs dans la boucle augmente). Quand le quantificateur sature, le signal de contre-réaction appliqué à l'entrée est bloqué et le gain de la rétroaction se dégrade, le quantificateur ayant besoin de plus de temps pour suivre le signal d'entrée, la boucle peut éventuellement devenir instable. D'après [16, 17, 19], le modulateur $\Sigma\Delta$ passe-bas sera stable si l'amplitude du signal en entrée du modulateur ne dépasse pas un certain pourcentage de l'amplitude

du signal de sortie $v(n)$ (typiquement 60% à 80%). Cette raison est directement liée à la condition de non-saturation du premier intégrateur.

Le critère basé sur le modèle quasi linéaire du modulateur: le quantificateur est modélisé par un amplificateur ayant un gain variable η suivi d'une source de bruit blanc additif. Ce modèle est basé sur le théorème de *Kalman* sur la stabilité des systèmes non linéaires [10, 18, 20]. Dans ce cas le schéma bloc du modulateur en absence du signal d'entrée est illustré sur la figure 2.9. La fonction

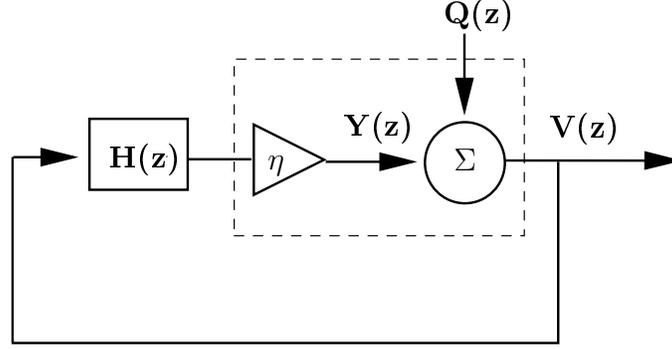


FIG. 2.9 – Modèle quasi linéaire du modulateur en absence du signal d'entrée.

de transfert du bruit devient alors :

$$NTF(z) = \frac{1}{1 + \eta H(z)} \quad (2.12)$$

En faisant intervenir le gain η dans la fonction de transfert du bruit de modulateur, le domaine de stabilité peut être déterminé dans le plan complexe par l'emplacement des pôles et des zéros.

Quand le signal d'entrée est nul, la puissance moyenne du signal de sortie V est égale à la puissance totale du bruit du modulateur dans la bande utile :

$$\overline{V^2} = \overline{(NTF \cdot Q)^2} \quad (2.13)$$

Dans le cas d'un modulateur monobit où le signal de sortie V est codé sur 1 bit, sa puissance moyenne est égale à l'unité, et la puissance moyenne du bruit égal à $\frac{\Delta^2}{12}$. L'équation 2.13 devient alors :

$$\overline{(NTF \cdot B)^2} = \frac{2}{3} \int_0^{\frac{1}{2}} \left| \frac{1}{1 + \eta H(e^{j2\pi\nu})} \right|^2 d\nu = 1 \quad (2.14)$$

où ν étant la fréquence normalisée $\left(\frac{f}{f_e}\right)$.

La résolution de l'équation 2.14 permet de déterminer la valeur du gain du comparateur η .

En présence du signal d'entrée, on obtient le modèle linéaire général du modulateur représenté à la figure 2.10. Dans ce cas la puissance moyenne du bruit totale du modulateur monobit est égale à :

$$\overline{(NTF \cdot B)^2} = 1 - \overline{(STF \cdot X)^2} \quad (2.15)$$

Ce qui nous ramène à établir une condition nécessaire de validité du modèle linéaire du modulateur :

$$\frac{2}{3} \int_0^{0.5} \left| \frac{1}{1 + \eta H(e^{j2\pi\nu})} \right|^2 d\nu \leq 1 \quad (2.16)$$

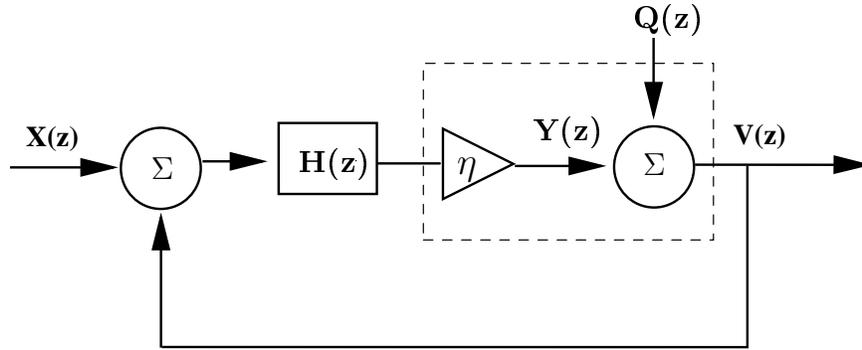


FIG. 2.10 – Modèle quasi linéaire du modulateur.

Le calcul de η peut se faire en pratique, en déterminant la valeur pour laquelle l'erreur et le signal sont décorrélés. Cette valeur optimale de η est indiquée par l'équation suivante [2, 21].

$$\eta_{opt} = \frac{\langle y, v \rangle}{\langle y, y \rangle} = \frac{cov(y, v)}{var(y)} = \lim_{N \rightarrow \infty} \frac{\sum_{n=0}^N y(n)v(n)}{\sum_{n=0}^N y(n)^2} \quad (2.17)$$

La formule ci-dessus montre que η_{opt} dépend de y , qui dépend lui même du signal d'entrée du modulateur x . Comme la stabilité du modulateur dépend de son signal d'entrée, il est impossible de tester toutes les amplitudes et les formes de ce signal. De plus, la plupart des modulateurs d'ordre élevé (supérieur à 2) ne sont jamais totalement stables, ce qui signifie que dans certaines conditions les signaux intermédiaires peuvent diverger. D'autres méthodes d'analyse (diagramme de Nyquist, plan de phase...) permettent d'obtenir des critères de stabilité. Cependant, elles ne permettent pas de construire une architecture de manière analytique. La principale difficulté est que pour définir une architecture de modulateur $\Sigma\Delta$, il faut évaluer le gain η du quantificateur, mais ce gain dépend de l'architecture choisie et surtout du signal d'entrée. L'architecture d'un modulateur d'ordre élevé peut se déterminer de manière empirique à partir du gabarit de sa fonction de mise en forme de bruit et en considérant que le gain en hautes fréquences ne doit pas dépasser une certaine valeur [21].

La valeur de η obtenue pour une entrée nulle est le paramètre le plus intéressant. Soit η_0 cette valeur. Elle peut être obtenue par exemple en résolvant l'équation 2.16, qui impose que la puissance du signal de sortie et la puissance du bruit de sortie soient égales. On note que l'équation 2.16 ne peut pas être résolue formellement pour un ordre supérieur à 2, des méthodes numériques doivent être utilisées.

Le lieu des pole-zéros:

La variation de η , en fonction du signal d'entrée peut mener à de l'instabilité bien que le modèle linéaire simple prévoit une architecture stable. Le lieu des zéros de l'expression $\frac{1}{1+\eta H(z)}$ apporte des informations concernant le comportement d'un modulateur. Un modulateur sera :

- complètement instable si un de ses pôles est hors du cercle d'unité pour $\eta = \eta_0$,
- inconditionnellement stable si le lieu de ses pôles sont complètement à l'intérieur du cercle d'unité pour $\eta < \eta_0$,
- conditionnement stable dans tout autre cas.

La figure.2.11 montre les lieux des pôles d'un modulateur du second degré totalement stable pour

un η variant de 2 à 0. Il est intéressant de noter que tous les pôles se situent à l'intérieur du cercle d'unité ce qui montre la stabilité du modulateur. Cependant dans la pratique, lors d'une réalisation par des circuits électroniques, il est préférable de réduire le gain de la boucle afin d'assurer une marge de stabilité suffisant en tenant compte des déviations des caractéristiques de composants réels par rapport à leurs caractéristiques idéales.

La figure 2.12 montre le lieu des pôles (dans le plan complexe) de la fonction NTF d'un modulateur

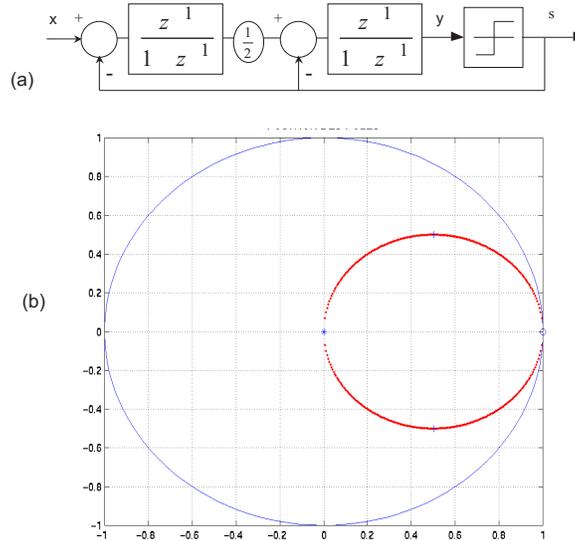


FIG. 2.11 – a) Modulateur passe-bas d'ordre 2, b) lieu des pôles du modulateur d'ordre 2

du 3^{ième} ordre passe-bas quand le gain η varie de 0 à 1. Comme on peut le constater, les pôles se déplacent à l'extérieur du cercle unité quand $\eta < 0.572$ ce qui conduit à l'instabilité du modulateur. Par conséquent, le gain de boucle doit être strictement limité pour un modulateur d'ordre trois ou supérieur [1, 15].

2.3.5 Du passe-bas au passe-bande

La façon la plus simple d'introduire la modulation passe-bande consiste à s'appuyer sur la théorie développée précédemment avec le passe-bas. Après avoir mis en équations la mise en forme de bruit en passe-bande, les performances du modulateur seront calculées analytiquement. Enfin, les problèmes de stabilité seront analysés.

La différence entre un modulateur $\Sigma\Delta$ passe-bas et un modulateur passe-bande réside essentiellement dans la fonction de mise en forme de bruit $NTF(z)$. Le bruit de quantification doit être minimal non plus à la fréquence nulle, mais à une fréquence f_o correspondant le plus souvent à la fréquence centrale du signal d'entrée. Le passage du passe-bas au passe-bande s'obtient en remplaçant le filtre passe-bas de la boucle $H(z)$ du modulateur $\Sigma\Delta$ (voir Figure.2.13) par un filtre passe-bande.

Les modulateurs passe-bas ont une mise en forme du bruit de quantification de la forme passe-haut et sont réalisés avec un filtre composé d'intégrateurs. Si nous construisons le filtre de boucle, $H(z)$, à partir de résonateurs, le bruit sera rejeté en dehors de la fréquence centrale du filtre. Dans ce cas,

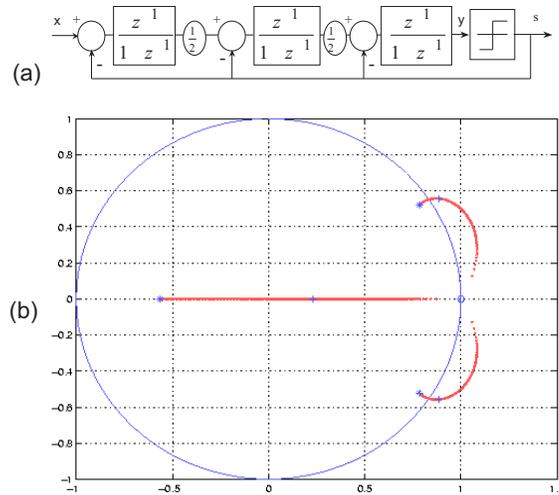


FIG. 2.12 – (a) Modulateur passe-bas d'ordre 3 instable, (b) Position des pôles en fonction du gain du comparateur d'un modulateur $\Sigma\Delta$ du 3^{ième} ordre.

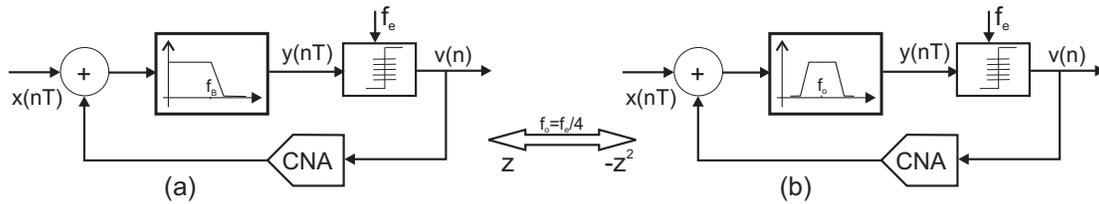


FIG. 2.13 – Modulateurs $\Sigma\Delta$: (a) passe-bas et (b) passe-bande

la mise en forme du bruit de quantification est de type coupe-bande [22].

Un convertisseur passe-bande peut être obtenu à partir d'un convertisseur passe-bas en appliquant la substitution $z^{-1} \rightarrow -z^{-2}$ dans le filtre de boucle; il en résulte un convertisseur pour lequel le bruit de quantification est rejeté en dehors de la fréquence centrale (celle ci est alors égale au quart de la fréquence d'échantillonnage) avec les mêmes propriétés de performances et de stabilité qu'un passe-bas, l'ordre du filtre étant cependant multiplié par deux [22,23].

Une application typique d'un tel convertisseur est la conversion analogique numérique d'un signal *RF* ou *IF* d'un récepteur de radio communication, dont le traitement est représenté à la figure.2.14. En effet, les multiplications par les fonctions sinus et cosinus se résument alors à une multiplication par $\{0, 1 \text{ ou } -1\}$. Ainsi l'étage complexe de démodulation analogique à réaliser est remplacé par quelques portes logiques.

Le spectre à la sortie d'un convertisseur sigma delta passe-bande est représenté sur la figure.2.14. On peut remarquer que le bruit de quantification est minimal aux alentours de la fréquence centrale égale à $\frac{f_e}{4}$. En général, la possibilité d'un modulateur $\Sigma\Delta$ à convertir un signal en bande étroite à une fréquence non nulle, le rend particulièrement attractif dans les applications radio-fréquence. De plus, un modulateur $\Sigma\Delta$ à temps continu peut être suffisamment rapide pour permettre la conversion de signaux vers une centaine de méga-hertz et même au delà. Le facteur *OSR* d'un convertisseur

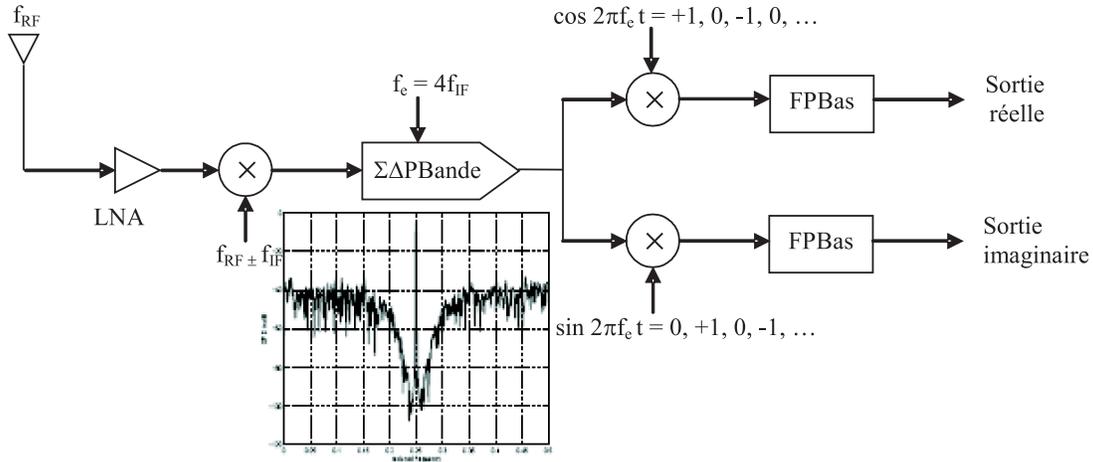


FIG. 2.14 – Chaîne de réception radio utilisant un passe-bande

Le passe-bande est défini comme la moitié du rapport entre la fréquence d'échantillonnage et la largeur de bande utile [2]. Ainsi, un convertisseur travaillant à $\frac{f_e}{4}$ avec un signal occupant l'intervalle de fréquence $(\frac{f_e}{4} - \frac{f_e}{2OSR}, \frac{f_e}{4} + \frac{f_e}{2OSR})$ a une largeur de bande de $2f_B = \frac{f_e}{OSR}$, et par conséquent l' OSR est égal à $\frac{f_e}{2f_B}$.

2.3.6 Stabilité d'un modulateur $\Sigma\Delta$ passe-bande

L'étude de la stabilité des modulateurs $\Sigma\Delta$ [15] passe-bande s'inspire fortement des travaux effectués dans le cas du passe-bas. Comme pour tout système re-bouclé, le problème se pose pour une fonction d'ordre $2m \geq 6$: c'est-à-dire qu'un modulateur passe-bande d'ordre $L = 2m > 4$ peut être instable. Cependant, les simulations semblent montrer que si un modulateur $\Sigma\Delta$ passe-bande est réalisé à partir de son homologue passe-bas en effectuant la substitution $z^{-1} \rightarrow -z^{-2}$, les conditions de stabilité du modulateur passe-bande seront identiques à celles du passe-bas. Cette constatation peut s'expliquer par le fait que le passage du passe-bas au passe-bande correspond uniquement à une variation de phase des pôles et des zéros dans le plan complexe. Or, les conditions de stabilité sont liées au module des pôles.

2.3.7 Du modulateur sigma-delta à temps discret au modulateur à temps continu

La vitesse du filtre de boucle du modulateur est une limitation majeure de la fréquence d'échantillonnage. Bien que les circuits à capacités commutées aient été l'approche dominante pour implanter les filtres de la boucle à temps discret du modulateur $\Sigma\Delta$ en raison de leur souplesse et leur précision, leur vitesse est limitée à quelques dizaines de MHz en raison de la non-idéalité des commutateurs. D'où l'intérêt des filtres à temps continu qui peuvent être implantés, entre autre, par une technique G_m - C ou en utilisant des composants passifs intégrés LC [24] ou des composants ligne de transmission (TL : *transmission-Line*) [25–27], permettant d'opérer en haute fréquence.

Une illustration du modulateur $\Sigma\Delta$ à temps discret et à temps continu est représentée à travers la

figure 2.15. La différence entre les deux modulateurs se trouve d'une part à l'entrée du modulateur. Celle-ci, dans le modulateur à temps continu, est un signal qui varie dans le temps $x_c(t)$, alors que dans le modulateur à temps discret, le signal d'entrée est échantillonné avant d'entrer dans la boucle : $x_d[n] = x_c(nT)$. D'autre part dans le modulateur à temps continu, l'échantillonnage s'effectue juste avant le quantificateur comme cela est indiqué sur la figure 2.15.b.

Le CNA du chemin de retour n'est qu'un interrupteur qui, à chaque période T , affecte au code numérique appliqué à son entrée un signal analogique qui peut être une tension ou un courant. Ce signal peut être constant durant le cycle d'horloge (*NRZ* : *Non-Return to Zero*), ou peut changer durant le cycle (*RZ* : *Return to Zero*, *HZ* : *Half-delay return to Zero*). Dans le cas des modulateurs discrets c'est le signal de sortie du CNA moyenné qui est important et non pas sa forme qui peut varier entre deux fronts d'horloge. Cependant dans le cas du modulateur à temps continu, la forme de sortie du convertisseur numérique analogique (*CNA*) sur le chemin de retour est aussi importante que sa valeur moyenne, car la sortie de CNA est appliquée directement à l'entrée du filtre continu.

Les modulateurs $\Sigma\Delta$ à temps continu sont des systèmes mixtes : l'entrée et le filtre de boucle sont

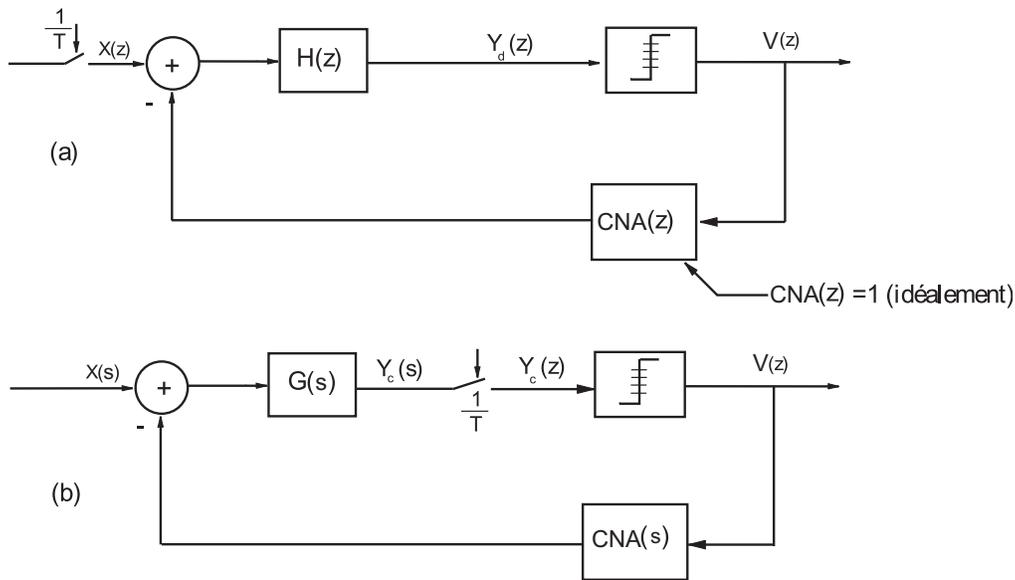


FIG. 2.15 – Modulateur $\Sigma\Delta$ (a) à temps discret, (b) à temps continu.

à temps continu et leur sortie est à temps discret. Afin de surmonter les problèmes liés à la synthèse des circuits mixtes en raison des difficultés de simulation, un modulateur $\Sigma\Delta$ à temps continu peut être entièrement étudié dans le domaine discret puis être réexaminé en temps continu.

Afin d'obtenir le modulateur, nous allons introduire l'équivalence temps discret - temps continu. Cette équivalence existe en raison de l'échantillonnage qui s'effectue à l'entrée du quantificateur et qui ne change pas le comportement global du modulateur. Pour montrer l'équivalence, nous allons annuler l'entrée et ouvrir la boucle du modulateur $\Sigma\Delta$ autour du quantificateur. La figure 2.16 représente la boucle ouverte qui entoure le quantificateur dans le cas discret et continu.

Le modulateur à temps continu va produire le même code de sortie $v(n)$ que le modulateur en temps discret si l'entrée du quantificateur dans les deux cas est identique aux instants d'échantillonnage,

c'est à dire $y_d[n] = y_c[nT]$. Ceci sera vrai si la condition suivante est vérifiée [15,28–31] :

$$H(z) = Z_T \{L^{-1}[CNA(s) \cdot G(s)]\}, \quad (2.18)$$

avec

$CNA(s)$: transformée de Laplace de la réponse à un échantillon du CNA ,

Z_T : la transformée en Z , après échantillonnage à la période T

L^{-1} : la transformée de Laplace inverse.

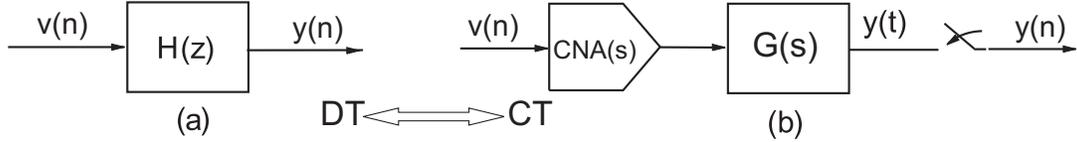


FIG. 2.16 – Boucle ouverte du modulateur (a) à temps discret, (b) à temps continu.

D'une façon générale, durant une période T_e , le signal rectangulaire à la sortie du CNA représenté sur la figure 2.17, peut être décrit dans le domaine temporel au moyen de la relation suivante :

$$cna(t) = u(t - t_d) - u(t - t_d - \tau), \quad (2.19)$$

où $u(t)$ un échelon unité, t_d et τ sont le retard et la largeur de la pulsation. En appliquant la

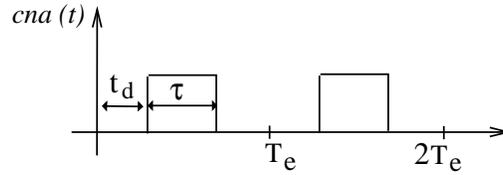


FIG. 2.17 – Réponse temporelle du CNA.

transformée de Laplace, nous obtenons :

$$CNA(s) = \frac{e^{-t_d s} - e^{-(t_d + \tau)s}}{s} \quad (2.20)$$

Selon les valeurs de t_d et τ , on peut modifier la largeur de la pulsation, ce qui conduit à trois type de CNA : NRZ , RZ ou HZ (fig.2.18). Dans le cas d'un CNA de type NRZ ($t_d = 0$ et $\tau = T_e$), par substitution dans l'équation 2.20 nous obtenons la fonction de transfert classique d'un bloqueur d'ordre zéro.

$$CNA(s) = \frac{1 - e^{-T_e s}}{s} \quad (2.21)$$

Dans ce cas simple, l'équation exprime d'équivalence temps discret-temps continu (équation 2.18) s'écrit :

$$H(z) = (1 - z^{-1})Z_T \left[L^{-1} \left\{ \frac{G(s)}{s} \right\} \right] \quad (2.22)$$

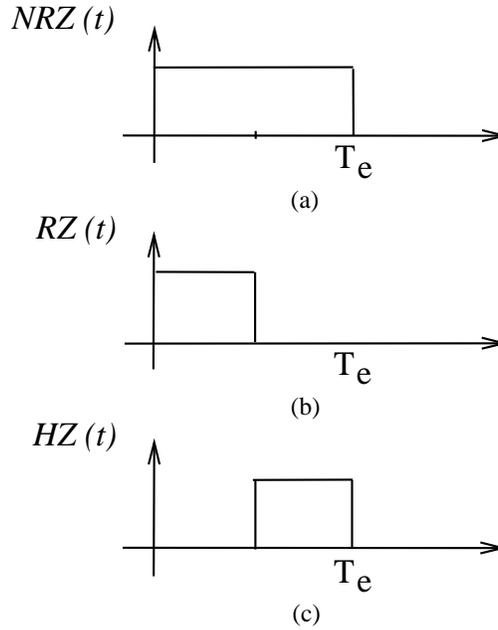


FIG. 2.18 – Réponse temporelle du CNA : (a) NRZ, (b) RZ, (c) HZ.

2.3.8 Comparaison entre les modulateurs à temps discret et les modulateurs à temps continu

Les principaux avantages et inconvénients des modulateurs $\Sigma\Delta$ à temps continu par rapport à leurs équivalents à temps discret sont :

1. **La fréquence d'échantillonnage :** dans les circuits à capacités commutées, la fréquence d'échantillonnage est limitée par la bande passante de l'amplificateur opérationnel et par les erreurs produites dans l'échantillonneur (SAH). Ces erreurs sont dues aux non idéalités des commutateurs, au transfert de charge, à l'injection de charges à travers l'horloge (*Clock Feedthrough*, *CFT*), à l'effet de gigue d'horloge et au temps d'établissement. Ceci a pour effet d'augmenter le bruit dans la bande et par conséquent de limiter la fréquence d'échantillonnage dans les modulateurs à temps discret. Tandis que dans les modulateurs à temps continu, ces erreurs sont rejetées en dehors de la bande d'intérêt par la fonction de transfert NTF, car l'échantillonnage s'effectue dans la boucle l'entrée du quantificateur.
2. **La consommation de puissance :** La diminution des contraintes sur les parties actives du modulateur, notamment sur l'ampli op, offre une réduction possible de la consommation et du niveau de la tension d'alimentation des circuits du modulateur. Cela est plus important si le filtre compose d'éléments passifs, c.-à-d. R, L, C . D'autre part, l'échantillonnage introduit dans la boucle du modulateur, réduit fortement le bruit thermique (par exemple $\frac{KT}{C}$ dans la bande de fréquence utile (voir le chapitre 3), et le repliement des signaux en dehors de la bande, ce qui permet d'éliminer le filtre antirepliement. Donc, pour les mêmes spécifications, un modulateur à temps continu consomme moins de courant que son équivalent à temps discret.
3. **Faible tension d'alimentation :** La tendance à la réduction de la tension d'alimentation des

circuits, notamment en technologie CMOS a limité les performances des circuits à capacités commutées (réduction de la dynamique de sortie).

Cependant, les modulateurs $\Sigma\Delta$ à temps continu sont sensibles à la gigue d'horloge et aux caractéristiques de commutation (la forme du signal de sortie de *CNA*) dans le chemin de retour contrairement aux modulateurs à temps discret :

1. **La gigue d'horloge :** La gigue d'horloge notée δT_e est la variation aléatoire des fronts d'horloges. Cette variation constitue une source de distorsion non linéaire. Les modulateurs à temps discret sont peu sensibles à la gigue d'horloge par rapport à leurs équivalents à temps continu. Le signal de retour dans les modulateurs à temps discret possède une forme exponentielle (fig.2.19.a). Le plus grand transfert de charge se produit au début de la période d'horloge, ce qui fait que la quantité de charge δQ_d perdue due à la gigue d'horloge δT_e est relativement faible. A l'inverse, la sortie du *CNA*, dans le cas du modulateur à temps continu, est un signal rectangulaire (figure 2.19.b), où le transfert de charge est constant sur une période d'horloge et donc la quantité de charge perdue δQ_c due à la même gigue d'horloge est beaucoup plus importante que dans le cas des modulateurs à temps discret.

Nous traiterons l'effet de la gigue d'horloge en détail dans le chapitre suivant.

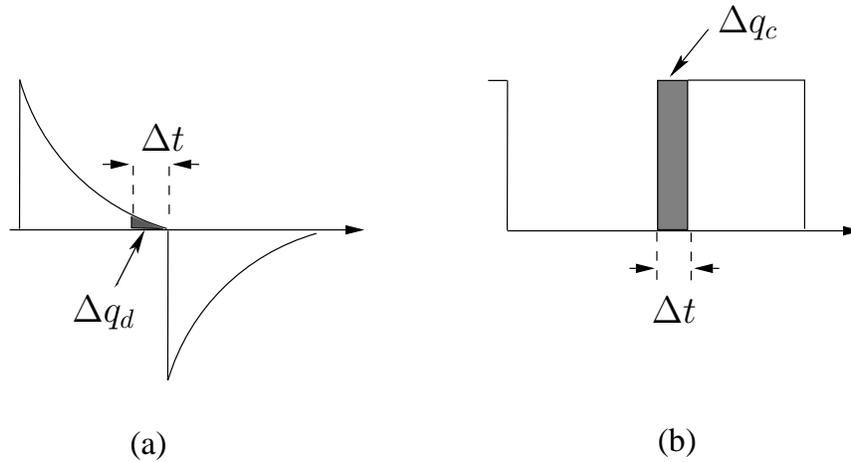


FIG. 2.19 – L'effet de la gigue d'horloge dans les modulateurs (a) à temps discret, (b) à temps continu.

2. **Le retard de boucle :** Le retard de boucle dans les modulateurs à temps continu est dû aux temps de propagation des différents blocs (le filtre, le CAN et le *CNA*). Un incrément de ce retard modifie la fonction de transfert du bruit *NTF* et risque donc de déplacer les pôles en dehors du cercle unité, ce qui réduit le *SNR* et peut aussi conduire le modulateur à l'instabilité [28, 32–36]. Utiliser un *CNA* de type *RZ* peut donner assez du temps au CAN à établir une sortie, et supprimer par conséquent l'influence du temps d'établissement du CAN sur les performances et la stabilité du modulateur. Il est aussi possible d'appliquer un certain retard artificiel entre le CAN et le *CNA* (par exemple $0.25T_e$ ou $0.5T_e$). Cette technique peut permettre d'éviter le risque d'incertitude de la valeur du temps de propagation entre eux et le risque d'instabilité. Ce retard artificiel qui est supposé supérieur à ceux du CAN et du *CNA*,

devrait être compensé au niveau système. Afin de compenser le retard dans la boucle de sigma delta, la méthode la plus simple qui a été produite dans notre laboratoire est l'utilisation d'un rebouclage supplémentaire de la sortie du CNA à l'entrée du CAN [29, 30].

3. **La forme du signal de sortie du CNA:** L'inégalité des temps de montée et de descente et de dérivation non prévus des formes des sources de courant commutées du CNA, introduit des harmoniques dans le signal de retour et dégrade par conséquent le *SNDR* du modulateur. L'effet de l'asymétrie peut être réduit en utilisant un CNA du type *RZ*.

Malgré les désavantages des modulateurs à temps continu, on peut dire que ces derniers vont pouvoir surmonter essentiellement les limites des récentes et futurs technologies CMOS à cause des qualités qu'ils présentent par rapport à la faible tension d'alimentation, leur faible consommation et la grande fréquence d'échantillonnage [4].

2.3.9 Résolution du quantificateur

Il est toujours possible de remplacer le quantificateur monobit ($B = 1$) par un quantificateur multibit ($B \geq 2$), par exemple, un convertisseur flash [37]. Ceci a deux avantages principaux : améliorer la résolution globale du modulateur sigma delta, et rendre les modulateurs d'ordre supérieur plus stables. De plus, les imperfections dans le quantificateur (par exemple l'hystérésis, l'offset, non-linéarité,...) ne dégradent pas beaucoup les performances parce que le quantificateur est précédé par plusieurs résonateurs (intégrateur dans le cas d'un passe-bas) à gain élevé. Par conséquent les erreurs ramenées à l'entrée sont négligeables [38]. Le remplacement d'un quantificateur monobit par un quantificateur multibit a deux effets positifs sur la résolution globale du modulateur, l'un direct, l'autre indirect. Premièrement, la résolution du modulateur augmente proportionnellement la résolution interne du quantificateur soit $6.02B$ en dB. Deuxièmement, le gain total de la boucle d'une structure multibit peut être plus élevé que celle d'un monobit. Grâce à ce surcroît de gain, le bruit de quantification est mis en forme plus efficacement. L'effet de ce gain supplémentaire dépend non seulement du nombre de bits mais encore de l'architecture du filtre et de l'ordre du modulateur. Par exemple, un modulateur passe-bas d'ordre 3 qui utilise un quantificateur 4-bits peut bénéficier d'un gain supplémentaire d'un facteur $23dB$ apportant une différence d'environ 3.8-bits supplémentaires soit en total $(4 - 1) + 3.8 = 6.8\text{bits}$, (41dB) par rapport à son homologue monobit, comme le montre la figure 2.20 [39]. De plus, un modulateur multibit pose moins de problèmes, e.g. des tons, vitesse de balayage de l'amplificateur opérationnel, ..., que le modulateur monobit [1].

Les trois principaux inconvénients sont d'une part l'augmentation de la complexité et la limitation de la vitesse de conversion d'un quantificateur multibit par rapport à un quantificateur monobit et d'autre part les imperfections du CNA de la boucle du retour directement ramenées à l'entrée, de sorte que les erreurs de linéarité du CNA réduisent sensiblement les performances du convertisseur. Des méthodes connues sous le nom de techniques de brassage de sources peuvent compenser les erreurs de non linéarité du CNA multibit. Celles-ci ne sont pas nécessaires dans une conception de convertisseur monobit parce que les CNA monobit sont intrinsèquement linéaires.

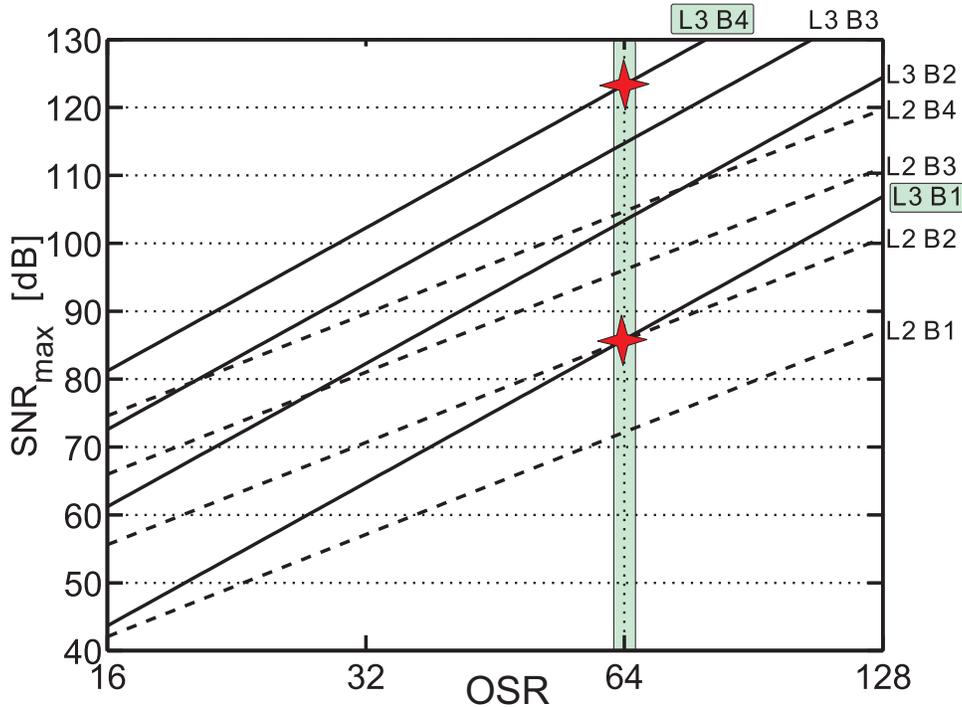


FIG. 2.20 – SNR_{max} maximal du modulateur $\Sigma\Delta$ passe-bas; d'ordre $L_i = i^{th}$ et de quantification de $B_j = j$ -bits vis-à-vis du OSR.

La technique de brassage de sources est un des objectifs de cette thèse et nous la détaillons en faisant des nouvelles propositions aux chapitres 4 et 5.

2.3.10 Calcul général du SNR_{max} d'un modulateur $\Sigma\Delta$ passe-bas et passe-bande

Dans l'architecture de base des modulateurs $\Sigma\Delta$, le chemin direct contient une série d'intégrateurs comme le montre la figure 2.21. Le nombre d'intégrateurs est égal à l'ordre du modulateur L . Normalement, chaque intégrateur devrait être accompagné d'un gain (a_i) qui permet d'ajuster le gain total de la boucle ainsi que la variation maximal de chaque etage. Or, la fonction de transfert de l'intégrateur i peut être décrite par l'équation suivante :

$$I_i(z) = \frac{a_i z^{-1}}{1 - z^{-1}} \quad (2.23)$$

Les fonctions de transfert NTF et STF pour cette architecture peuvent s'exprimer de la façon suivante :

$$NTF(z) = \frac{1}{1 + \sum_{i=1}^L \prod_{j=L+1-i}^L I_j} \quad , \quad STF(z) = \frac{\prod_{j=1}^L I_j}{1 + \sum_{i=1}^L \prod_{j=L+1-i}^L I_j} \quad (2.24)$$

Dans le domaine fréquentiel, pour une modélisation linéaire, la sortie du modulateur passe-bas (figure 2.21) correspond à la somme du signal d'entrée et du bruit de quantification mis en forme au moyen de l'approximation suivante :

$$NTF(z) = \frac{(1 - z^{-1})^L}{G} \quad (2.25)$$

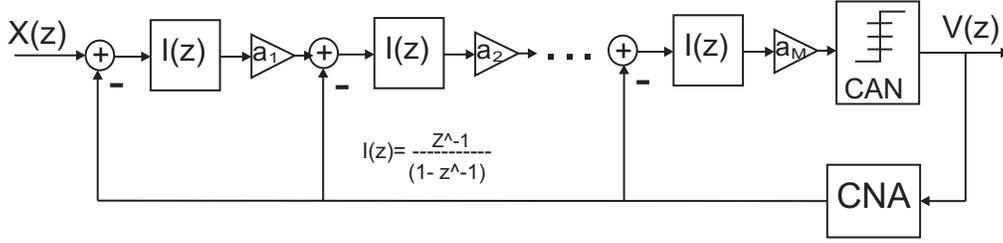


FIG. 2.21 – Architecture simple du modulateur $\Sigma\Delta$ passe-bas d'ordre L .

où, G est le gain total de la boucle, équivalant à $\prod a_i$. La conception du modulateur nécessite une STF proche de l'unité dans la bande utile. La puissance du bruit dans la bande utile $[0, f_B]$ du signal est :

$$\begin{aligned} q_B^2 &= 2 \int_{f_B} |Q(f)|^2 df = 2 \int_{f_B} |NTF(f)q_{rms}|^2 df \\ &= 2q_{rms}^2 \int_{f_B} |NTF(f)|^2 df = \frac{2q_{rms}^2}{G^2} \int_{f_B} |(1 - e^{-\frac{j2\pi f}{f_e}})^L|^2 df \end{aligned} \quad (2.26)$$

Puisque $(f/f_e) \ll 1$, nous donnons une valeur approchée de l'intégrale en utilisant le premier terme du développement limité du sinus, ce qui donne :

$$P_{q_B} \cong \frac{2q_{rms}^2}{G^2} \int_0^{f_B} \left[\frac{2\pi f}{f_e}\right]^{2L} df = \frac{\pi^{2L} \Delta^2}{12G^2(2L+1)f_e OSR^{(2L+1)}} \quad (2.27)$$

où $OSR = \frac{f_e}{2f_B}$, correspond au taux de sur-échantillonnage. Alors, le rapport signal sur bruit de quantification est calculé en prenant le rapport entre la puissance du signal d'entrée et la puissance du bruit. Pour un signal d'entrée sinusoïdal d'amplitude A et par normalisation ($\Delta \equiv 1$), le rapport signal sur bruit SNR vaut :

$$SNR_{LP} = 10 \log\left[\frac{P_{sig}}{P_{q_B}}\right] = 10 \log\left[\frac{6A^2G^2(2L+1)OSR^{(2L+1)}}{\pi^{2L}}\right] \quad (2.28)$$

Théoriquement pour un modulateur $\Sigma\Delta$ non-optimisé, le maximum de la valeur du rapport signal sur bruit, SNR est bien inférieur à celui obtenu par l'équation 2.28 en prenant $A_{max} = \frac{(2^B-1)}{2}$:

$$\begin{aligned} SNR_{LP,max} &= 10 \log\left[\frac{3(2^B-1)^2G^2(2L+1)OSR^{(2L+1)}}{2\pi^{2L}}\right] \\ &= 6.02B + 1.76 + 10 \log\left[\frac{(2L+1)OSR^{(2L+1)}}{\pi^{2L}}\right] - 20 \log\left[\frac{1}{G}\right] \end{aligned} \quad (2.29)$$

de la même manière que le cas précédent (passe-bas), la figure.2.22 montre l'architecture d'un modulateur passe-bande d'ordre $L = 2m$, simple et non optimisé, dont les fonctions de transfert de signal et de bruit de quantification sont données par les équations suivantes :

$$R_i(z) = \frac{-a_i z^{-2}}{1 + z^{-2}}, \quad i \in [1, 2, \dots, m]$$

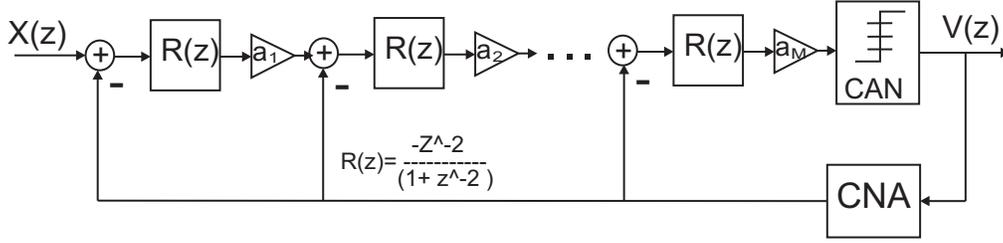


FIG. 2.22 – Architecture simple du modulateur $\Sigma\Delta$ passe-bande d'ordre L .

$$\begin{aligned}
 NTF(z) &= \frac{1}{1 + \sum_{i=1}^m \prod_{j=m+1-i}^m R_j(z)} = \frac{1}{1 + \sum_{i=1}^m (-1)^i \frac{z^{-2i}}{(1+z^{-2})^i} \prod_{j=m+1-i}^m a_j} \\
 STF(z) &= \frac{\left[\frac{-z^{-2i}}{(1+z^{-2})}\right]^m \prod_{j=1}^m a_j}{1 + \sum_{i=1}^m \prod_{j=m+1-i}^m R_j(z)} \quad (2.30)
 \end{aligned}$$

La sortie du modulateur passe-bande (figure.2.22) correspond à la somme du spectre du signal d'entrée et du spectre du bruit de quantification mis en forme par la fonction suivante :

$$NTF(z) = \frac{(1 + z^{-2})^m}{G} \quad (2.31)$$

où, G est le gain total de la boucle, équivalent à $\prod a_i$. En réalisant le même calcul pour la modulateur passe-bande que pour le modulateur passe-bas d'ordre $L = 2m$ centré au quart de la fréquence d'échantillonnage, on obtient alors :

$$\begin{aligned}
 P_{Q_B} &= \int_{f_o-f_B}^{f_o+f_B} |Q(f)|^2 df = \int_{f_o-f_B}^{f_o+f_B} |NTF(f)q_{rms}|^2 df \\
 &= q_{rms}^2 \int_{f_o-f_B}^{f_o+f_B} |NTF(f)|^2 df = \frac{q_{rms}^2}{G^2} \int_{f_o-f_B}^{f_o+f_B} |(1 + e^{-\frac{j4\pi f}{f_e}})^m|^2 df \quad (2.32)
 \end{aligned}$$

avec un changement du paramètre : $f' = f - f_o$ on simplifie les équations :

$$P_{Q_B} = \frac{2q_{rms}^2}{G^2} \int_0^{f_B} |(1 + e^{-\frac{j4\pi(f'+f_o)}{f_e}})|^{2m} df' \quad (2.33)$$

Puisque $(f'/f_e) \ll 1$, nous estimons l'intégrale en utilisant le premier terme du développement du sinus, ce qui donne un rapport du signal sur le bruit maximal SNR_{max} valant :

$$\begin{aligned}
 SNR_{BP,max} &= 10 \log \left[\frac{3(2^B - 1)^2 G^2 (2m + 1) OSR^{(2m+1)}}{2\pi^{2m}} \right] \\
 &= 6.02B + 1.76 + 10 \log \left[\frac{(2m + 1) OSR^{(2m+1)}}{\pi^{2m}} \right] - 20 \log \left[\frac{1}{G} \right] \quad (2.34)
 \end{aligned}$$

On observe bien que le résultat final du calcul du SNR_{max} est exactement le même pour le modulateur passe-bas (équation 2.29). La seule différence est l'effet de l'ordre du modulateur qui est divisé par 2, c'est-à-dire que c'est le nombre de résonateurs qui compte et non pas l'ordre du filtre passe-bande.

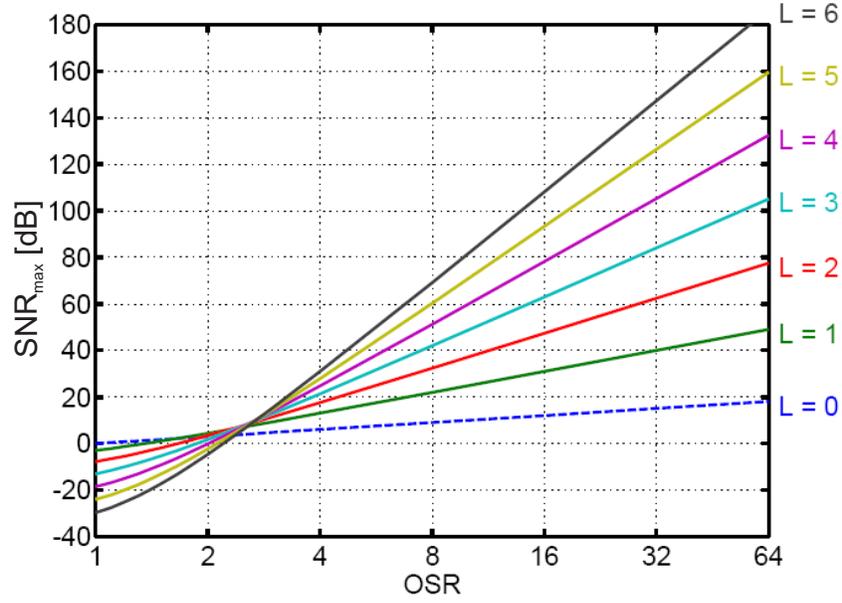


FIG. 2.23 – Performances du modulateur $\Sigma\Delta$ en fonction de l'OSR pour un NTF = $(1 - z^{-1})^L$ dans le cas d'un monobit d'ordre = $\{0 \dots 8\}$ et idéal

L'ordre du modulateur est choisi en fonction de la fréquence de travail maximale f_e et des performances souhaitées. Un réseau de courbes décrivant le rapport signal à bruit (SNR_{max}) en fonction du taux de suréchantillonnage OSR pour différentes valeurs de l'ordre est souvent utilisé (voir Figure.2.23). Ces courbes sont habituellement tracées à partir du modèle simplifié du quantificateur. Le nombre de bits effectifs ($ENOB$) du modulateur $\Sigma\Delta$ passe-bande se déduit du rapport signal à bruit SNR_{max} grâce à la formule 2.29.

2.3.11 Optimisation de fonction de transfert du bruit de quantification

Dans le cas d'un modulateur d'ordre élevé, l'utilisation d'une fréquence centrale unique pour tous les résonateurs n'est pas optimale. On peut diminuer la puissance de bruit en décalant ces fréquences. Pour déterminer les valeurs optimales des fréquences de résonance, on peut utiliser l'hypothèse de la bande étroite. Il est alors possible de développer les formules au premier ordre. Le module de chaque fonction I_i pour passe-bas et R_i pour passe-bande s'écrivent :

$$\begin{aligned} |I_i|^2 &\cong a_i^2 \left[2\pi \frac{f_i}{f_e} \right]^2 \\ |R_i|^2 &\cong a_i^2 \left[2\pi \left(\frac{f_o}{f_e} - \frac{f_i}{f_e} \right) \right]^2 \end{aligned} \quad (2.35)$$

On peut se contenter de faire le calcul pour le modèle passe-bande. Le modèle du passe-bas est déduit en prenant $f_o = 0$. En faisant varier les fréquences $f_o - f_B \leq f_i \leq f_o + f_B$, on peut minimiser la puissance de bruit P_{Q_B} qui est proportionnelle au produit P :

$$P = \int_{\frac{f_o - f_B}{f_e} - \frac{f_B}{f_e}}^{\frac{f_o + f_B}{f_e} + \frac{f_B}{f_e}} \prod \left[\frac{f_o}{f_e} - \frac{f_i}{f_e} \right]^2 df \quad (2.36)$$

TAB. 2.1 – Valeurs optimales des fréquences de résonance et amélioration de la résolution

Ordre du modulateur	Position normalisée des fréquences	Gain en dB par rapport à une fréquence centrale unique	Amélioration de la résolution en nombre de bits
2	± 0.58	3.5	0.5
3	$0, \pm 0.77$	8	1
4	$\pm 0.34, \pm 0.86$	13	2
5	$0, \pm \sqrt{\frac{5}{9}} \pm \sqrt{\frac{40}{567}}$	18	3
6	$\pm 0.24, \pm 0.66, \pm 0.93$	23	4

P est un terme proportionnel à la puissance du bruit. Dans le cas du modulateur d'ordre $L = 2$. On suppose que les fréquences sont symétriques par rapport à la fréquence centrale du filtre global, c.-à-d. $f_o = \frac{f_e}{4}$.

$$\begin{aligned}
 f_1 &= f_o + f_a \text{ et } f_2 = f_o - f_a \\
 P &= \int_{\frac{f_o - f_B}{f_e}}^{\frac{f_o + f_B}{f_e}} \prod \left(\frac{f}{f_e} - \frac{f_1}{f_e} \right)^2 \left(\frac{f}{f_e} - \frac{f_2}{f_e} \right)^2 df \\
 P &= \int_{\frac{f_o - f_B}{f_e}}^{\frac{f_o + f_B}{f_e}} \prod \left(\frac{f}{f_e} - \frac{f_o + f_a}{f_e} \right)^2 \left(\frac{f}{f_e} - \frac{f_o - f_a}{f_e} \right)^2 df
 \end{aligned} \tag{2.37}$$

Soit après calcul :

$$P = \left(\frac{f_a}{f_e} \right)^4 \left(\frac{f_B}{f_e} \right) - \frac{1}{6} \left(\frac{f_a}{f_e} \right)^2 \left(\frac{f_B}{f_e} \right)^3 + \frac{1}{80} \left(\frac{f_B}{f_e} \right)^5 \tag{2.38}$$

Cette fonction est minimale pour $f_a = \frac{f_B}{2\sqrt{3}}$ Soit :

$$f_a \cong 0.58 \frac{f_B}{2} \tag{2.39}$$

Le même principe de calcul peut être appliqué à des modulateurs d'ordre élevé. Les résultats pour ces modulateurs d'ordre $m = \{2,3,\dots,8\}$ sont regroupés dans le tableau 2.3.11. Cependant, si le facteur de qualité du filtre appliqué est faible, il vaut mieux fixer les fréquences des zéros au centre de la bande plutôt qu'utiliser les fréquences données au tableau 2.3.11.

2.3.12 Critères de performance du modulateur:

Le rapport signal sur bruit de quantification, SNR est le critère le plus utilisé pour caractériser la performance d'un modulateur. Ce rapport signal sur bruit de quantification est défini par l'équation approximative 2.28 et sa valeur maximum par l'équation 2.29. Mais il y a d'autres critères comme suit:

THD : Distorsion harmonique totale et l'IMD : inter-modulation

Un opérateur linéaire idéal ne modifie pas la composition spectrale du signal de sortie tel qu'il était à l'entrée. Toutes les non linéarités introduites par l'opérateur vont se manifester par une distorsion de la réponse dans le domaine temporel et par l'apparition de nouvelles fréquences dans le domaine

fréquentiel. Lorsqu'un CAN convertit une sinusoïde pure, la représentation spectrale du signal de sortie comporte une raie à la fréquence du signal d'entrée mais aussi aux multiples du fondamental qui composent la distorsion harmonique. La distorsion harmonique totale (*THD*) est définie par :

$$THD_{dB} = 10 \log \left[\sum_{i=2}^{\infty} (HD_i)^2 \right] \quad (2.40)$$

où la distorsion harmonique HD_i , est l'amplitude de la i^{me} harmonique ramenée sur l'amplitude de la fondamentale. La non-linéarité du CAN peut également créer de la distorsion d'inter-modulation (*IMD*).

L'IMD traduit le comportement du convertisseur en présence d'un signal d'entrée composé de deux sinusoïdes pures de fréquences différentes [1,2].

Le ***SNDR*** est le rapport de la puissance du signal à la sortie du modulateur sur la somme de la puissance du bruit et des harmoniques :

$$SNDR(dB) = 10 \log \left(\frac{P_{sig}}{P_{[bruit+harmoniques]}} \right) \quad (2.41)$$

Il faut noter que la valeur maximale du *SNDR* dépend de la fréquence et peut être utilisée pour mesurer la dégradation des performances du modulateur quand la fréquence du signal d'entrée augmente.

Le SNR et le SNDR, ainsi que la dynamique de sortie (DR) sont souvent représentés sur le graphique illustré à la figure 2.24. Comme on peut le constater, pour des faibles niveaux du signal d'entrée, la distorsion harmonique n'est pas importante ce qui implique un *SNR* et un *SNDR* approximativement égaux. Plus le niveau du signal d'entrée augmente, plus la distorsion harmonique augmente et le *SNDR* sera donc plus faible que le *SNR*.

La dynamique libre de raies parasites *SFDR* est définie comme le rapport entre la valeur maximale du signal de la sortie et la valeur maximale d'une harmonique présente dans la bande utile.

$$SFDR = 10 \log \left[\frac{\text{la puissance d'entrée}}{\text{La puissance de la plus grande harmonique}} \right] = 10 \log \left[\frac{V_1^2}{V_{HD_i}^2} \right] \quad (2.42)$$

où V_1 est le signal principal et V_{HD_i} la plus grande harmonique supposée à la sortie du modulateur. Il est important de connaître sa valeur maximum qui est définie par la relation suivante :

$$SFDR_{(dBFS)} = 10 \log \left[\frac{(2^B - 1)^2}{2 V_{HD_i}^2} \right] \quad (2.43)$$

Où, l'indice "*i*" est normalement 2 pour un modulateur simple et 3 pour un modulateur à structure différentielle.

On parle de l'importance de ce rapport (*SFDR*) essentiellement dans les systèmes de communication.

La dynamique de sortie (*DR*) est le rapport de puissance entre le niveau maximal du signal d'entrée que le modulateur peut convertir et le niveau minimal du signal d'entrée détectable. La DR sur le graphique, exprimée en *dB*, est la différence entre le niveau d'entrée quand le SNDR chute de 3 *dB* et quand la courbe du SNDR coupe l'axe des abscisses (*SNDR*=0). La DR est fortement liée

à l'ordre du modulateur (L), au facteur de suréchantillonnage (OSR) et au nombre de bit (B) à la sortie du modulateur (résolution du CAN de la boucle) par la relation suivante [2,3] :

$$\begin{aligned} DR &= 10\log\left[\frac{\text{la puissance maximal d'entrée}}{\text{La puissance minimal d'entrée détectable}}\right] \\ &\approx 10\log\left[\frac{3G^2(2m+1)OSR^{(2m+1)}}{\pi^{2m}}\right] + 3dB \end{aligned} \quad (2.44)$$

Ainsi un compromis doit être réalisé entre ces différents paramètres afin d'atteindre les performances visées.

Le nombre de bits effectifs ($ENOB$) du modulateur $\Sigma\Delta$ passe-bande se déduit du rapport signal

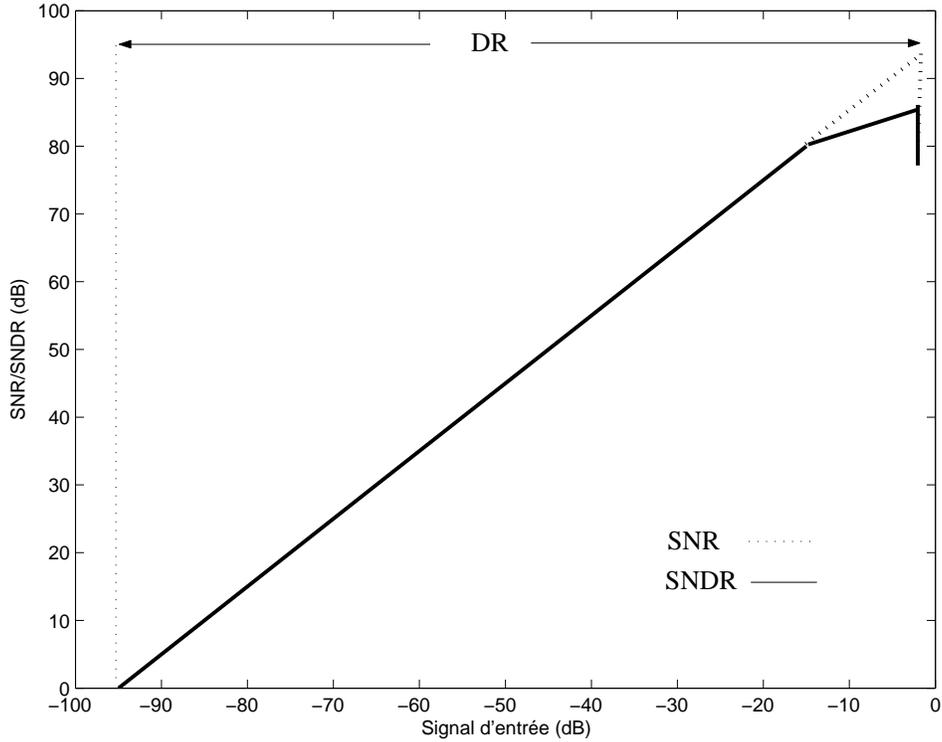


FIG. 2.24 – $SNR/SNDR$ et DR .

sur bruit SNR_{max} grâce à la formule Eq.2.29 :

$$ENOB = \frac{DR_{dB} - 1.76}{6.02} \cong B + 1.67 \log\left[\frac{(2m+1)OSR^{(2m+1)}}{\pi^{2m}}\right] - 3.3\log\left[\frac{1}{G}\right] \quad (2.45)$$

La consommation de puissance (P_{tot}) et le facteur de mérite (FOM)

Par ailleurs, la faible consommation d'un modulateur est une propriété importante en raison de l'augmentation du nombre des équipements portables utilisant des batteries dans les systèmes de communication, et l'augmentation du nombre de blocs fonctionnels intégrés sur la même puce. En correspondance des autres paramètres du modulateur, ce critère peut être résumé par *le facteur de mérite* ("Figure of merit") FOM . Dans la littérature, plusieurs définitions de ce FOM existent. La définition la plus souvent utilisée est obtenue par la relation suivante [40,41] :

$$FOM = \frac{2^{ENOB} \times 2f_B \times V_{dd} \times L}{P_{tot}} = \frac{2^{ENOB} \times f_e \times V_{dd} \times L}{P_{tot} \times OSR} \quad (2.46)$$

où P_{tot} est la consommation du modulateur (en watt ou milliwatt).

Plus le FOM est grand, plus le modulateur est performant. Actuellement, il varie entre 0.5 et quelque centaines d'après les rapports publiés [4, 41–44].

Déviati on pratique de la formule de DR calculée

Théoriquement, le rapport du signal sur bruit et la DR augmentent :

- de $9dB + (m - 1)6dB$ lorsque le taux de suréchantillonnage OSR double de valeur,
- d'environ $20\log\left[\frac{OSR}{\pi}\right]$ pour un ordre supplémentaire du modulateur,
- de $10\log[2^B - 1]$ directement pour un quantificateur B-bits par rapport au monobit (+9,5dB si $B=2$).

Eq.2.29 est une estimation très brute qui peut donner une valeur approximative et non précise car les grandeurs d'influence ne sont pas considérées. En pratique, la dynamique du signal d'entrée diminue avec l'ordre du modulateur pour des raisons de stabilité. Ce phénomène, devient plus important lorsque on veut concevoir un modulateur d'ordre élevé et monobit. Cela peut être récompensée dans le cas d'un modulateur multibit e.g. $B=3$ ou 4. En revanche l'optimisation du lieu des zéros de NTF améliore la résolution accessible du modulateur.

En considérant les effets non idéaux du modulateur, la relation 2.34 (SNDR) doit être précisée et adaptée à cause des phénomènes suivants :

- Optimisation de la NTF (effet positif et important),
- Diminution du gain de boucle G pour avoir une meilleure marge de stabilité (effet négatif important),
- Des tonalités³ (effet négatif surtout pour les modulateur monobit et ordre faible)
- Saturation possibles dans les étages intermédiaires (effet négatif),
- Erreurs de défaut d'appariement dans le CNA , (effet négatif et très important),
- Erreurs de défaut d'appariement dans le CAN , (effet négatif mais moins important),
- Non-linéarité de filtre en particulier au première étage, (effet négatif et important),
- Effets de bande limitée des circuits surtout à l'étage d'entrée, (effet négatif),
- bruit thermique et bruit en $\frac{1}{f}$, (effet négatif),
- effet de gigue d'horloge, (effet très négatif).

Certain facteurs de dégradation de la performance seront abordés plus en détail dans le chapitre suivant.

Critères pour différents choix de modulateur

Le choix d'une configuration conforme aux applications souhaitées sont dictés par un certain nombre de critères :

- DR et SNR maximum,
- largeur de la bande (BW), OSR et fréquence d'échantillonnage (f_e),
- sensibilité aux imperfections des circuits,

3. Idle tones

- sensibilité à la gigue d’horloge,
- stabilité de la boucle (conditionnelle ou non-conditionnelle),
- niveau de la consommation de puissance,
- contraintes imposées aux circuits, e.g. BW, SR⁴, dynamique d’entrée et linéarité d’ampli-op, bruit thermique de transistor, non-idéaliste de composants passifs L,C et R, tension d’alimentation, retards des différents blocs dans la boucle,
- complexité du modulateur et sa partie numérique, surface consommée, consommée
- nécessité d’un ajustement externe, étalonnage, correction numérique,
- nécessité d’un signal de bruit imposé ou non, "dither",
- disponibilité des outils de conception, formules analytiques, logiciels au niveau système.

Par ailleurs, les différentes catégories de modulateurs $\Sigma\Delta$ peuvent être classées ci-dessous, selon les choix possibles :

- type : à temps discret ou à temps continu (DT ou CT),
- application : Passe-bas ou passe-bande,
- architecture : en un seul rebouclage, en configuration cascade ou parallèle,
- quantification interne : monobit ou multibit (B),
- ordre de modulateur (L),
- d’implantation multiple : SC, SI, Gm-C, LC-intégré, TL,
- Mode d’opération : simple ou différentiel

Nous allons résumer les avantages et les désavantages de ces options en fin de chapitre.

2.4 Architecture de modulateurs $\Sigma\Delta$

Plusieurs familles de structures de modulateurs $\Sigma\Delta$ ont été proposées depuis une dizaine d’années. Elles permettent de réaliser des modulateurs d’ordre élevé afin de réduire la fréquence d’échantillonnage du modulateur et d’augmenter la résolution. Nous pouvons distinguer plusieurs grandes familles de modulateurs en fonction de leur structure interne. La complexité du modulateur va dépendre de l’ordre, du nombre de bits du quantificateur et du taux de suréchantillonnage *OSR*. Un certain nombre d’architectures de modulateurs connues à ce jour sont présentées dans cette section à fin d’éprouver leur robustesse vis-à-vis des imperfections des circuits électroniques dans un modulateur multibit haute performance.

2.4.1 Architecture de base : un seul-étage rebouclage-multiple (*CIFB/CRFB*)

L’architecture nommé un seul-étage à rebouclage-multiple *CIFB/CRFB* (*Cascade of Integrator (Resonator) Feedback Form*) a déjà été illustrée à la figure 2.21 pour un modulateur passe-bande. Le nombre d’intégrateurs (ou résonateurs) ainsi que le nombre de rétroactions, est identique à celui

4. La vitesse de balayage "Slew-Rate"

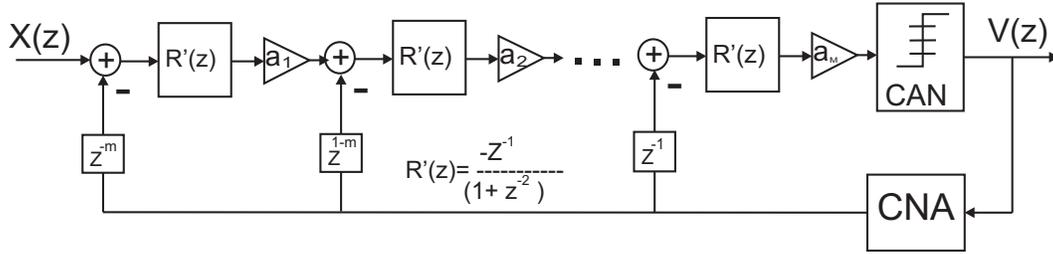


FIG. 2.25 – Architecture d'un modulateur passe-bande d'ordre L avec multiple rétroaction retardée et des numérateurs Z^{-1}

de l'ordre du modulateur sur le chemin direct. Chaque intégrateur (ou résonateur) doit posséder un gain a_i pour conférer à la boucle une marge de stabilité désirée, surtout dans le cas d'un modulateur d'ordre élevé. Les fonctions $NTF(z)$ et $STF(z)$, des passe-bas et passe-bande, ont déjà été respectivement données par les équations 2.24 et 2.30.

Les paramètres a_i ($i \in \{1, 2, \dots, m\}$) peuvent être ajustés afin d'obtenir à la fois une meilleure marge de stabilité dans la boucle et un minimum d'excursion du signal à la sortie de chaque intégrateur (ou résonateur).

Cependant, l'inconvénient important de cette topologie est que plusieurs blocs de CNA sont utilisés dans le chemin de la rétroaction. Cela rend l'architecture compliquée car elle a non seulement besoin de plus de circuits mais aussi d'ajustements ou de corrections pour au moins les 2 premiers CNA les plus proches de l'entrée, dans le cas multibit.

Un autre schéma d'architecture qui est préférable pour le cas du modulateur passe-bande est présenté à la figure.2.25.

Dans cette architecture, les retards introduits dans les chemins de rebouclage peuvent faciliter l'algorithme de brassage des sources ainsi que la réalisation du terme de résonateur.

Afin d'optimiser les lieux des zéros en NTF du modulateur, les deux architectures CIFB sont présentées aux figures 2.26 et 2.27 pour des ordres pairs et impairs du modulateur, respectivement. Sur

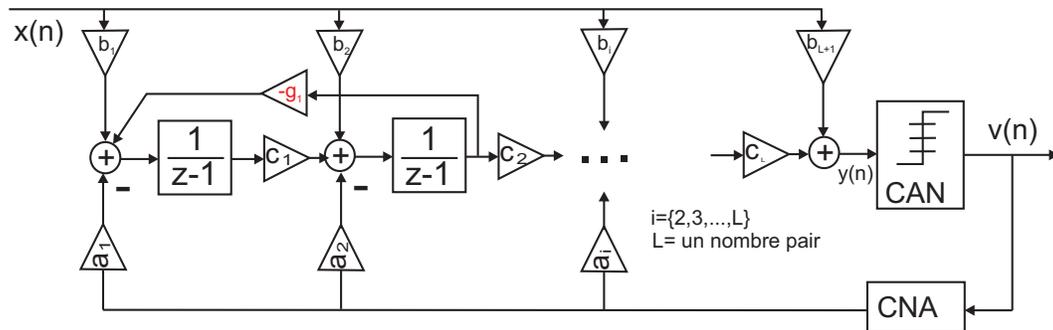


FIG. 2.26 – Architecture général CIFB d'ordre pair du modulateur

ces schémas, on voit que les branches de rebouclage " g_i " sont employées pour décaler la fréquence de résonance de chaque paire d'étages. En utilisant une méthode d'optimisation numérique basée sur

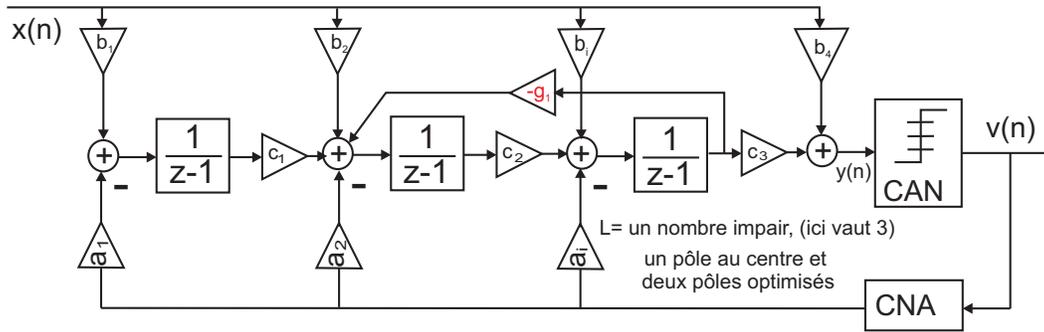


FIG. 2.27 – Architecture général CIFB d'ordre impair du modulateur

la minimisation du bruit de quantification dans la bande utile, on peut calculer les coefficients de ces architectures. Pour cela l'outil *ADOC* développé dans notre laboratoire [45] offre un environnement interactif et professionnel. Il existe un autre outil MATLAB écrit par le professeur R.Schreier, similaire mais non-interactif [46].

2.4.2 Architecture avec un seul rebouclage anticipatif CIFF/CRFF

Le prototype de l'architecture à une seule-boucle anticipatif ("*Cascade of Integrators (or Resonators) Feedforward Form*") pour un modulateur $\Sigma\Delta$ passe-bande est représenté sur la figure 2.28. Le modèle passe-bas correspondant de cette architecture peut être établi en remplaçant tous les résonateurs de type $[\frac{-z^{-2}}{(1+z^{-2})}]$ par des intégrateurs de type $[\frac{z^{-1}}{(1-z^{-1})}]$. Les équations STF et NTF de l'architecture CIFF/CRFF peuvent être déduites par le schéma 2.28. Pour un modulateur passe-bande d'ordre de $L = 2m$ centré sur un quart de la fréquence d'échantillonnage ($\frac{f_s}{4}$), on obtient alors :

$$NTF(z) = \frac{1}{1 - \sum_{i=1}^m \frac{a_i z^{-2i}}{(1+z^{-2})^i}} \quad , \quad STF(z) = \frac{\sum_{i=1}^m \frac{a_i z^{-2i}}{(1+z^{-2})^i}}{1 - \sum_{i=1}^m \frac{a_i z^{-2i}}{(1+z^{-2})^i}} \quad (2.47)$$

L'utilisation de la topologie CIFB/CRFB impose les restrictions suivants :

- premièrement, la structure n'est pas adaptée aux topologies avec un résonateur réel à retard simple au numérateur. En outre, l'arrangement libre du retard sur des chemins du signal analogique n'est pas facile avec l'architecture anticipative.
- d'autre part, la NTF optimisé causera une STF fixe, qui aura un certain gain hors bande et par conséquent un préfiltrage peut être exigé afin d'empêcher la surcharge des signaux hors bande. Il est naturellement possible de concevoir une STF plus plate mais, la remise en forme du bruit de quantification sera alors moins efficace.
- un autre inconvénient grave de cette structure apparaît lorsque les signaux de sortie analogique des résonateurs doivent être additionnés avant le CAN. Par exemple dans le cas d'une réalisation discrète, on peut utiliser un amplificateur de *SC*, qui consommera beaucoup de puissance. Dans le cas du modulateur à temps continu, on utilise un additionneur analogique qui produit du retard et consomme encore plus de puissance.

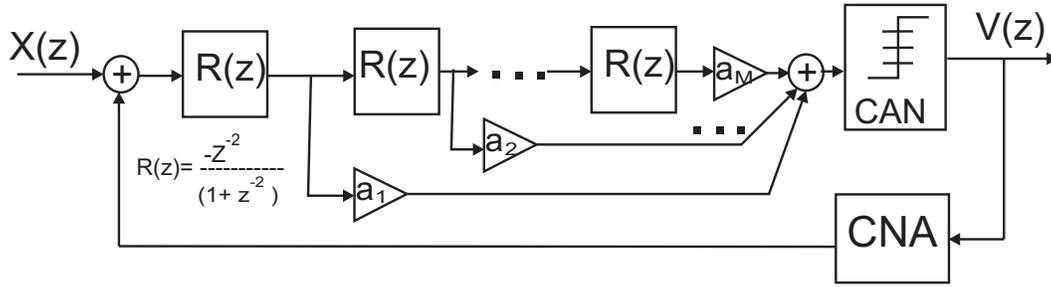


FIG. 2.28 – Architecture d'un modulateur passe-bande d'ordre L avec une seule boucle et plusieurs alimentations-*ver-l'avant* CIFA/CRFF

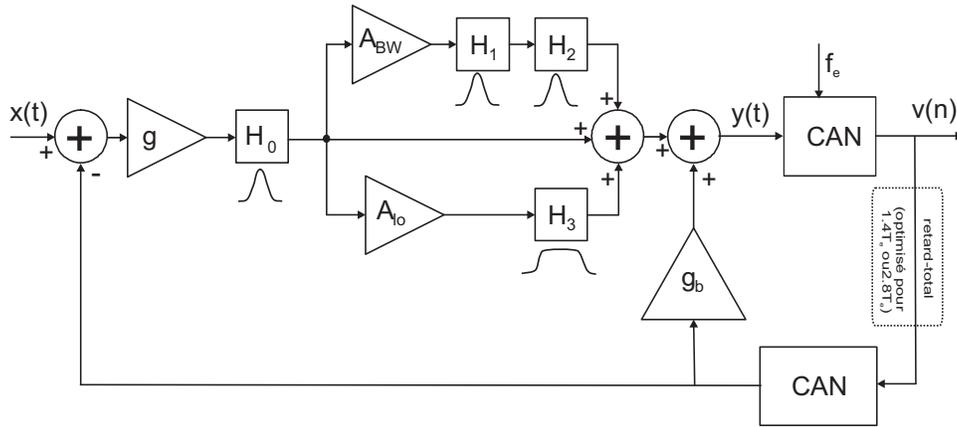


FIG. 2.29 – Architecture d'un modulateur passe-bande d'ordre 6 avec une seule boucle et plusieurs branches parallèles CIFA/CRFF, développé à Dept. SSE-Supelec

Par contre, la structure CIFA/CRFF présente deux avantages par rapport à la structure de CIFB/CRFB.

- Cette architecture peut produire une STF fixe et indépendante de la linéarité du premier étage. Cela est la raison la plus importante pour laquelle on l'utilise [47, 48].
- Elle nécessite seulement un CNA, qui simplifie le chemin de reboilage et favorise les applications multibits.

On peut également utiliser cette architecture partiellement dans le cas d'un modulateur à temps continu. Cette fois-ci, le but est d'obtenir des résonateurs réalisables par des circuits LC grâce à une architecture parallèle locale. En utilisant une architecture CIFA/CRFF modifiée, un modulateur d'ordre-six à temps continu dans lequel le problème du retard supplémentaire a été résolu par une rétroaction rapide a été récemment développée dans notre laboratoire (voir figure 2.29) [49–51]. Comme le cas précédent (c.-à-d. CIFB/CRFB), la forme générale de l'architecture CIFA/CRFF peut aussi être utilisée afin d'optimiser la fonction NTF dans la bande utile [45].

2.4.3 Architecture en cascade

Plusieurs modulateurs $\Sigma\Delta$ simples sont mis en cascade afin de réaliser un modulateur d'ordre élevé. Cette structure a l'avantage d'être stable car elle est composée de modulateurs d'ordre 1 ou 2.

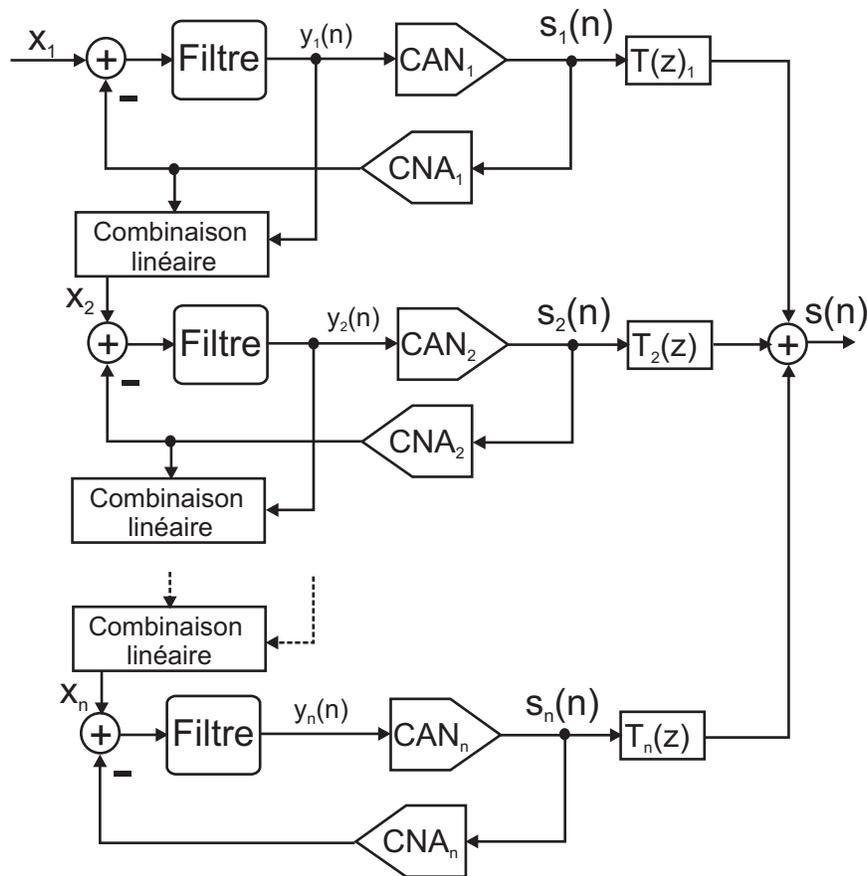


FIG. 2.30 – Les modulateurs $\Sigma\Delta$ en cascade.

Ce concept est illustré à la figure 2.30. Chaque modulateur traite une partie du bruit de quantification du modulateur précédent et un retraitement numérique ($T_i(z)$) est effectué sur les sorties (s_i). On obtient alors une meilleure précision dans la bande utile. L'entrée du $n^{\text{ième}}$ étage est définie par :

$$x_n = a y_{n-1} - b s_{n-1} \quad (2.48)$$

où, a et b sont des coefficients fixes qui servent à optimiser la performance globale du système. On distingue différentes structures en cascade :

- Structure *MASH* (*Multi stage noise SHaping*)
- Structure *nFOC*
- Structure mixtes
- Structure *MSCL* (*Multi Stage Closed Loop modulators*)

Malgré des avantages intéressants, l'architecture en cascade, présente de gros défauts pour une réalisation pratique. La majorité des différentes architectures présentées précédemment est très sensible aux imperfections des circuits analogiques. Cette faiblesse pose une difficulté non acceptable pour la réalisation de modulateurs à haute performance. Il faut bien noter que les imperfections des intégrateurs ou des résonateurs dans les différents étages ne peuvent pas être facilement réglées. Il n'existe

pas de méthode systématique comme celle du brassage de sources pour un CNA multibit. Nous décrivons brièvement les caractéristiques de chaque structure ci-après.

2.4.4 Structure *MASH*

Les structures *MASH* n se composent de n modulateurs du premier ou du deuxième ordre montés en cascade. L'entrée de chaque étage est le bruit de quantification de l'étage précédent ($b = -a$ dans l'équation 2.48). La figure 2.31 représente la structure d'un modulateur *MASH* 2. Ainsi, les signaux

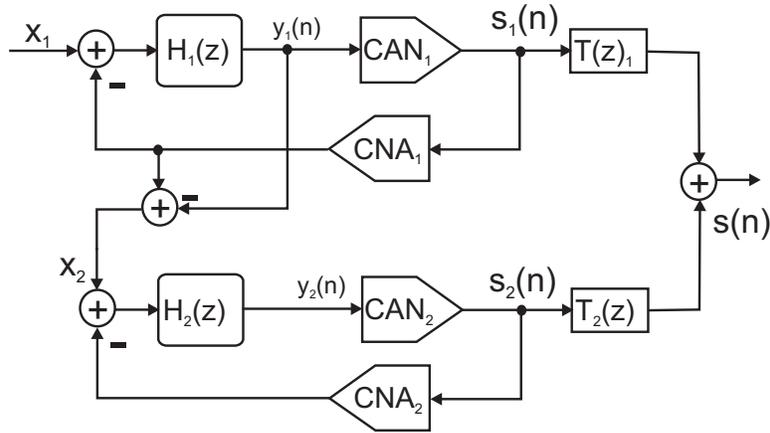


FIG. 2.31 – Modulateur $\Sigma\Delta$ *MASH* 2.

de sortie des deux étages sont donnés par :

$$S_1(z) = STF_1(z)X_1(z) + NTF_1(z)Q_1(z) \quad (2.49)$$

$$S_2(z) = STF_2(z)Q_1(z) + NTF_2(z)Q_2(z)$$

Les fonctions T_1 et T_2 constituant l'étage de préfiltrage, sont calculées de sorte que le bruit de quantification du premier étage ($Q_1(z)$) soit éliminé à la sortie globale (S). Ceci est réalisé si la condition suivante est vérifiée :

$$T_1 \cdot NTF_1 - T_2 \cdot STF_2 = 0 \quad (2.50)$$

Le choix le plus simple des fonctions T_1 et T_2 vérifiant l'équation 2.50 est $T_1 = STF_2$ et $T_2 = NTF_1$. Ce qui revient à dire :

$$T_1(z) = \frac{H_2(z)}{H_2(z) + 1} \quad , \quad T_2(z) = \frac{1}{H_1(z) + 1} \quad (2.51)$$

$H_1(z)$ et $H_2(z)$ étant respectivement les fonctions de transfert du filtre du premier et du second modulateur. Dans le cas classique d'un modulateur passe-bas du 2^{ème} ordre, la sortie globale de la structure s'écrit alors :

$$S(z) = T_1(z) V_1(z) - T_2(z) V_2(z) = z^{-4} X_1(z) - (1 - z^{-1})^4 Q_2(z) \quad (2.52)$$

On obtient alors les mêmes performances qu'un modulateur du 4^{ème} ordre à boucle unique, tandis que la stabilité est celle d'un modulateur du second ordre, d'où l'avantage de cette structure. Cependant,

les imperfections de réalisation de l'étage de prétraitement font apparaître le bruit Q_1 dans la sortie globale (condition eq.2.50 non vérifiée) et induisent alors une détérioration de la mise en forme du bruit [2].

2.4.5 Structure $nFOC$

La structure $nFOC$ réalise un modulateur $\Sigma\Delta$ d'ordre élevé. Elle est dérivée de la structure $MASH$. n modulateurs du premier ou second ordre montés en cascade la composent. Chaque étage échantillonne l'entrée du CAN de l'étage précédent ($b = 0$ dans l'équation 2.48). La figure 2.33 illustre la structure $nFOC$ du 2^{ième} ordre appelée $DFOC$.

Ainsi les interconnexions entre les étages sont plus simples à réaliser, d'où l'avantage de cette structure

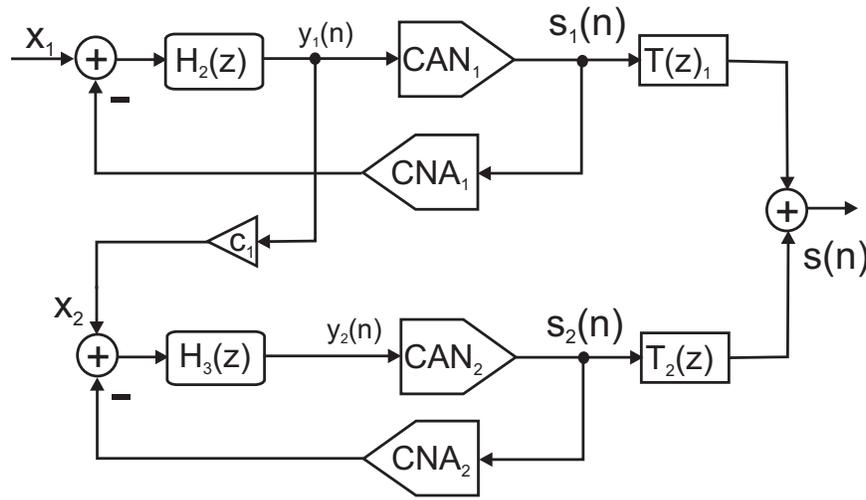


FIG. 2.32 – Modulateur $\Sigma\Delta$ $DFOC$.

par rapport à la structure $MASH$. Cependant, pour éviter la saturation de l'étage 2, un coefficient C_1 inférieur à 1 est nécessaire. Celui-ci sera compensé au niveau de l'étage de prétraitement numérique. Ceci a pour effet d'augmenter la puissance de bruit dans la bande.

2.4.6 Structures mixtes

Il s'agit de modulateurs cascades comportant des résonateurs simples ou doubles. Ces structures sont aussi appelées : structures 1-2 (simple résonateur suivi d'un double résonateur), 2-1 (double résonateur suivi d'un simple résonateur) ou 2-2 (double résonateur suivi d'un double résonateur). Elles ne présentent pas d'avantage particulier par rapport aux autres structures et présentent un risque d'instabilité. On peut utiliser également des structures mixtes passe-bas/passe-bande (double intégrateur suivi d'un simple résonateur) pour réaliser un modulateur passe-bas large bande. Le résonateur du deuxième étage est centré sur une fréquence basse, inférieure à la largeur de bande du modulateur [52].

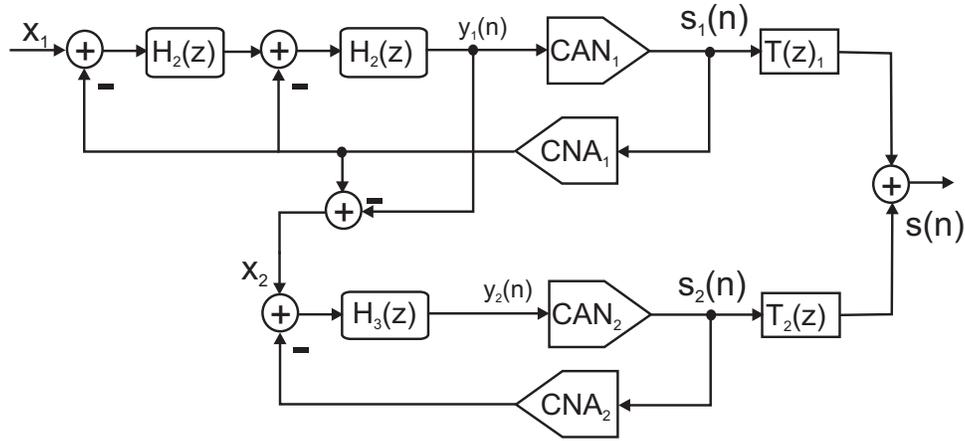


FIG. 2.33 – Modulateur $\Sigma\Delta$ mixe 2-1.

2.4.7 Structure *MSCL*

Cette structure a été proposée par P. BENABES [15] et consiste à mettre plusieurs modulateurs simples en cascade avec une rétroaction supplémentaire entre la sortie et l'entrée. Ceci permet de supprimer le préfiltrage des sorties de chaque étage mais nécessite un étalonnage à proximité de la rétroaction multibit afin d'atteindre une grande précision (chacune des rétroactions internes peut être monobit ou multibit, alors que la rétroaction globale est toujours multibit). La figure 2.34 représente une structure *MSCL* du 2^{ième} ordre. Le second modulateur échantillonne l'opposé du bruit de quan-

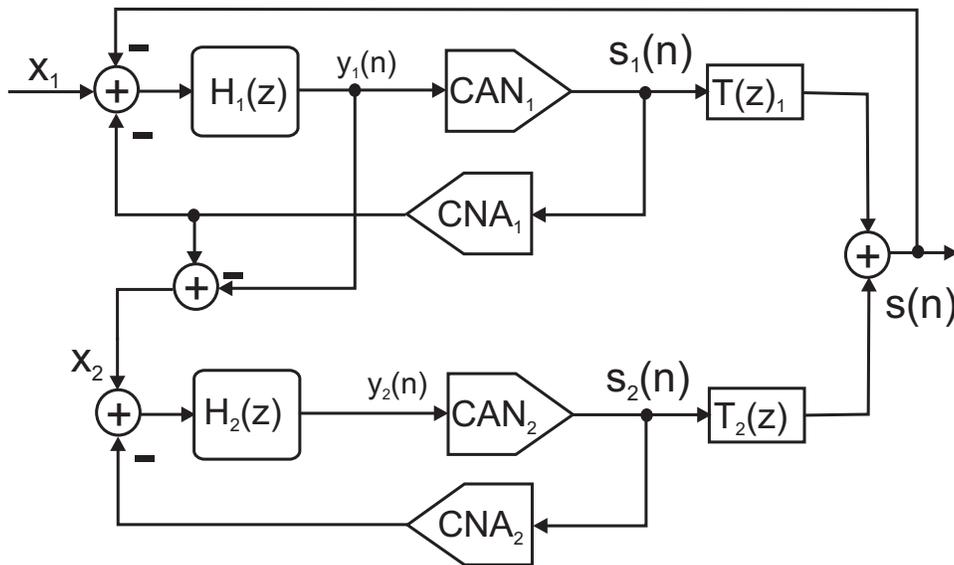


FIG. 2.34 – Modulateur $\Sigma\Delta$ *MSCL* d'ordre 2.

tification du premier modulateur ($y_1 - s_1$) et fournit une sortie s_2 visant à minimiser ce bruit. La sortie globale v est alors la somme de $s_1 + s_2$. Des études détaillées de la structure *MSCL* sur la stabilité, la sensibilité aux imperfections des composants ont été effectuées et présentées dans [15,53]. Il a été démontré que cette structure présente des performances et une stabilité proches des structures

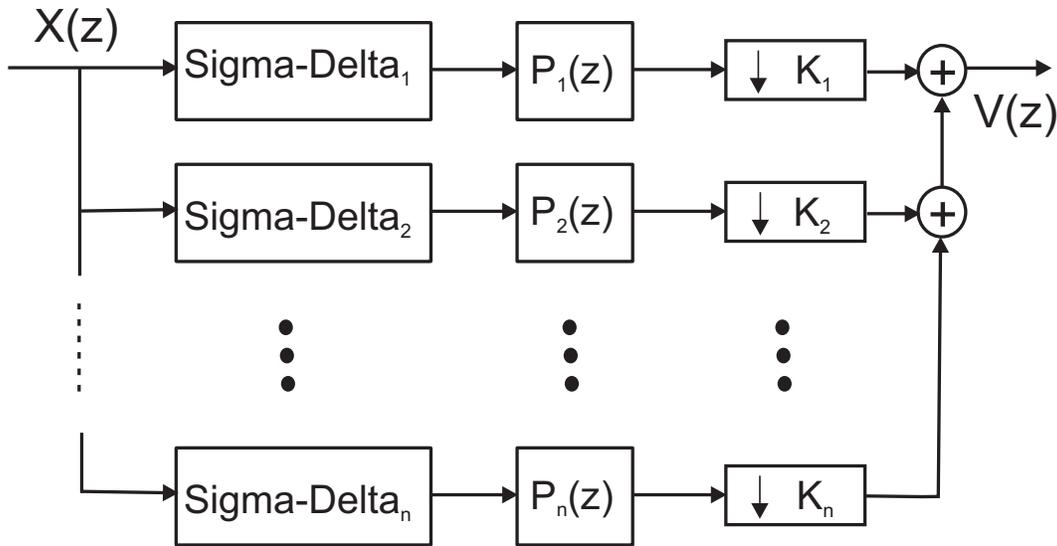


FIG. 2.35 – Architecture du convertisseur $\Sigma\Delta$ passe-bande parallèle

MASH.

De plus, il est toujours préférable que la conception d'un modulateur se base sur la structure MSCL. Cette propriété vient de la structure générale MSCL qui garde les concepts élémentaires de modulateur simple même pour obtenir une structure complexe et d'ordre élevé. $H(z)$ global d'un modulateur MSCL de J boucles cascadiées peut s'exprimer par la relation suivante :

$$\begin{aligned}
 H(z) &= \prod_{i=1}^J [1 + H_j(z)] - 1 \\
 H_j(z) &= \frac{z^{-1}}{1 + c_j z^{-1}} \quad \text{Pour le cas passe-bas,} \\
 H_j(z) &= \frac{-z^{-2}}{1 + c_j z^{-2}} \quad \text{Pour le cas passe-bande,}
 \end{aligned} \tag{2.53}$$

La structure équivalente sera ensuite obtenue à l'étape finale [45].

2.4.8 Architecture en parallèle Π

Un convertisseur $\Sigma\Delta$ passe-bande parallèle est représenté à la figure 2.35.

Un modulateur en Π résulte de la mise en parallèle de modulateurs dont les bandes utiles sont différentes mais adjacentes. Ainsi, le signal d'entrée est filtré de manière à définir autant de canaux qu'il y a de modulateurs; les sorties sont re-combinées, après filtrage, afin de reconstituer le signal. Chaque branche en parallèle est constituée par un modulateur $\Sigma\Delta$ suivi d'un filtre numérique décimateur $P_i(z)$.

Cette architecture résout un point faible de la conversion $\Sigma\Delta$: la faible largeur de bande utile. La mise en parallèle de plusieurs modulateurs $\Sigma\Delta$ permet également de réduire les contraintes que doivent respecter chacun d'entre eux. Les conditions sur les modulateurs internes sont fortes afin de garantir une reconstruction correcte du signal, tant du point de vue du gain que de celui de la phase. De ce

fait, on préfère utiliser le même type de modulateur pour la conversion de chaque sous-bande. En plus des contraintes de gain et de phase entre les différents branches, il faut veiller aux difficultés de filtrage, de séparation et de raccordement des sous bandes [54,55].

Une autre application de l'architecture parallèle est la réalisation d'un modulateur Sigma Delta en Π dont la fréquence de fonctionnement est la fréquence de Nyquist de la bande de fréquence du signal d'entrée (sans suréchantillonnage) [56]. Le suréchantillonnage de chaque sous bande est compensé par le nombre de modulateurs mis en parallèle. Cependant le coût de la réalisation de ce type de structure semble non négligeable.

De plus, les convertisseurs $\Sigma\Delta$ parallèles tels que ceux qui sont basés sur la modulation d'Hadamard nécessitent un système de correction supplémentaire (par exemple étalonnage) afin de compenser les problèmes de décalage (offset) et de gain [57].

2.4.9 Réduction des tons en utilisant un "dither" ou un pôle "Chaotique"

L'erreur de quantification d'un modulateur $\Sigma\Delta$ n'est typiquement pas de nature bruit blanc. Les entrées proches du DC ou du $fs/2$ pour lesquelles les erreurs de quantification sont périodiques, produisent des tons parasites connus généralement sous le nom de "*idle tone*", tonalité, ou simplement ton. Le spectre du modulateur contient des tonalités discrètes dont des fréquences et les amplitudes varient en fonction de la valeur et de la fréquence d'entrée. L'investigation prouve que les tons sont présents dans plupart des applications $\Sigma\Delta$ y compris les modulateurs simple ou multi-étapes (*MASH*), monobit ou multibit, du premier ordre ou supérieur. Cependant, les effets des tonalités diminuent quand le nombre de niveaux de quantification ou l'ordre du modulateur augmente. Afin d'observer le comportement tonal d'un modulateur $\Sigma\Delta$, il est plus pratique d'évaluer à court terme l'autocorrélation ou la densité spectrale *DSP* de la sortie. La nature périodique des tonalités peut poser des problèmes si le modulateur est destiné aux applications audio. L'oreille est très sensible à certaines types de corrélations et de périodicité dans le bruit.

On peut montrer que l'utilisation d'un bruit imposé appelée "*dither*" élimine les tonalités et blanchit le bruit de quantification du modulateur. Les techniques de dither peuvent employer un signal additif simple juste avant le quantificateur. Un signal mis en forme peut être aussi ajouté à l'entrée de la boucle. Il peut être de nature déterministe ou dynamique. La technique de dither et en particulier la technique de dither dynamique peuvent améliorer la résolution du modulateur $\Sigma\Delta$ pour des entrées de faible niveau. La figure 2.36 montre le spectre de la sortie d'un modulateur $\Sigma\Delta$ passe-bas de 2^{ime} ordre avec et sans dither et une structure typique du modulateur ayant un signal de dither.

Une seconde méthode qui permet de réduire ces tonalités est l'emploi d'un pôle *chaotique* [2]. Le concept de base cette technique consiste à déstabiliser le modulateur en déplacent un ou plusieurs zéros de la fonction de transfert du bruit (ou pôle du filtre) en dehors de la zone de stabilité ($|z| = 1$). Le résultat obtenu au moyen de la méthode *chaotique* est comparable à celui obtenu avec la technique de dither, mais au prix de circuits beaucoup plus complexes.

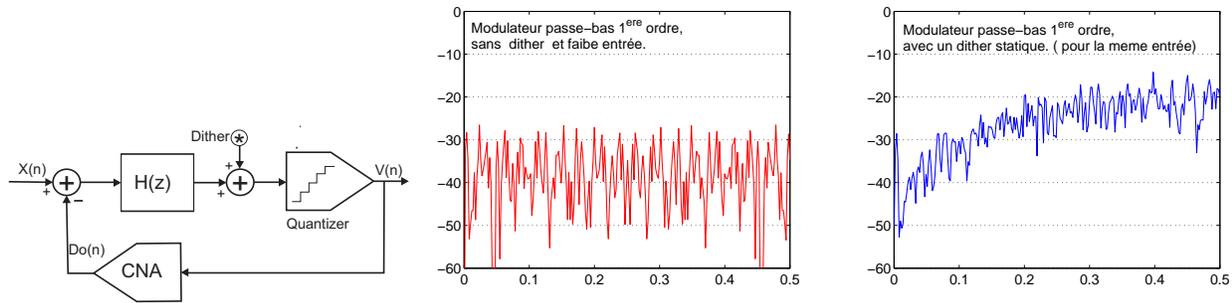


FIG. 2.36 – a) Structure d'un modulateur avec dither, b) sortie du modulateur sans aucun dither, c) avec dither.

2.4.10 Réalisation intégrée d'un modulateur $\Sigma\Delta$

Un modulateur $\Sigma\Delta$ peut être implémenté au moyen de circuits intégrés selon trois modes :

- **SC**; capacités commutées (*switched capacitance*) est le mode de réalisation essentiellement pour des modulateurs discrets à basses fréquences ($< 100\text{MHz}$). Actuellement, la majorité des modulateurs $\Sigma\Delta$ sont réalisés en mode SC.

Avantages :

- il est robuste et précis, car la fonction de transfert du filtre (comme intégrateur et résonateur) est déterminée par la valeur relative des capacités et non pas par leurs valeurs absolues
- il est très peu sensible à la gigue d'horloge
- il est compatible avec les technologies CMOS à très large échelle (VLSI),
- il est facile à simuler.

Désavantages :

- le signal d'entrée doit être filtré et échantillonné avant d'entrer dans le modulateur. Par conséquent, la performance du système peut être limitée par les erreurs d'échantillonnage.
- il a besoin de grande valeurs de capacités,
- il est sensible au bruit thermique (bruit KT/c)
- la vitesse peut être limitée par celle des amplis-op ($f_e < \frac{1}{5}f_{3dB-AOP}$),
- il a besoin de commutateurs ("switches") très performants avec une faible résistance.

La figure 2.37 montre le schéma d'un intégrateur SC dans le cas de structures simples et différentielles.

- **SI**; la méthode de courant commuté (*switched current*) est aussi une technique qui convient aux traitements discrets de signaux en particulier pour des modulateur $\Sigma\Delta$ à temps discret. Cette technique est prometteuse mais elle a eu peu de succès jusqu'à maintenant car elle est peu précise:

Avantages :

- elle nécessite seulement des transistors (et pas de grandes capacités),

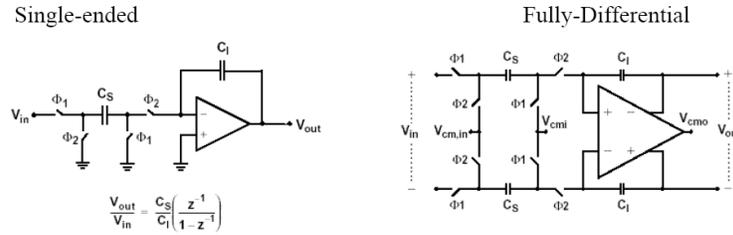


FIG. 2.37 – Le circuit d'intégrateur aux capacités commutées a) simple b) différentiel

- la consommation et la tension d'alimentation sont faible
- la surface de la puce est moindre,
- l'impédance d'entrée est faible,
- elle est compatible avec des processus totalement numériques

Inconvénient :

- très sensible aux injections de charges,
- très sensible aux défauts d'appariement des transistors,
- sensible à la valeur d'impédance (finie) à la sortie des cellules,
- moins rapide que SC,

Ses performances peuvent être améliorées en utilisant plusieurs méthodes d'étalonnage. Malgré tout, le maximum de résolution et les fréquences atteintes sont très loin derrière celles des SC (12-bit et quelques MHz). La figure 2.38 montre le schéma d'un intégrateur en SI dans le cas d'une structure simple.

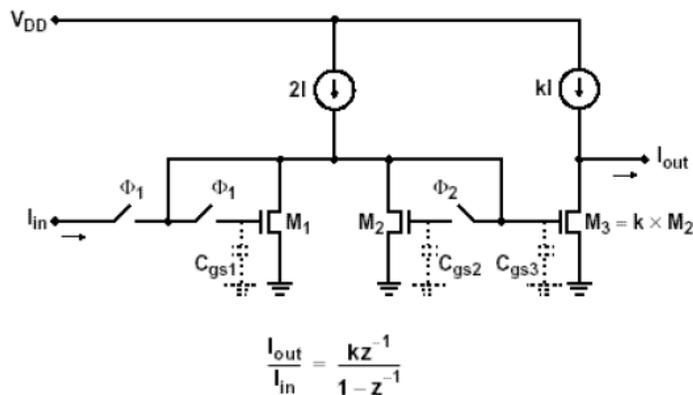


FIG. 2.38 – Circuit d'intégrateur aux courants commutés simple

– **RC-active, Gm-C, Mosfet-C, Gm-MC:**

Afin de réaliser un modulateur à temps continu, on utilise souvent des intégrateurs de type actif comme le montre la figure 2.39. Toutes ces techniques (RC-active, Gm-C, Mosfet-C, Gm-MC) sont conformes à la technologie d'intégration CMOS [58]. L'élément clé de ces intégrateurs est

toujours l'ampli-op à transconductance, mais, la grande difficulté est la variation aléatoire des valeurs absolues des éléments. Les erreurs des valeurs absolues des éléments (G_m , C et R_{ch}) déterminent l'erreur des caractéristiques d'intégrateur comme exprimée à la figure 2.39. On remarque que cela est un inconvénient important de l'intégrateur continu par rapport au cas précédent (SC et SI) où les erreurs relatives des éléments sont seulement déterminant. En plus, la linéarité de circuits G_m-C est faible et ils consomment beaucoup. Cette technique ne semble donc pas prometteuse pour l'avenir des convertisseurs précis de fréquence élevée.

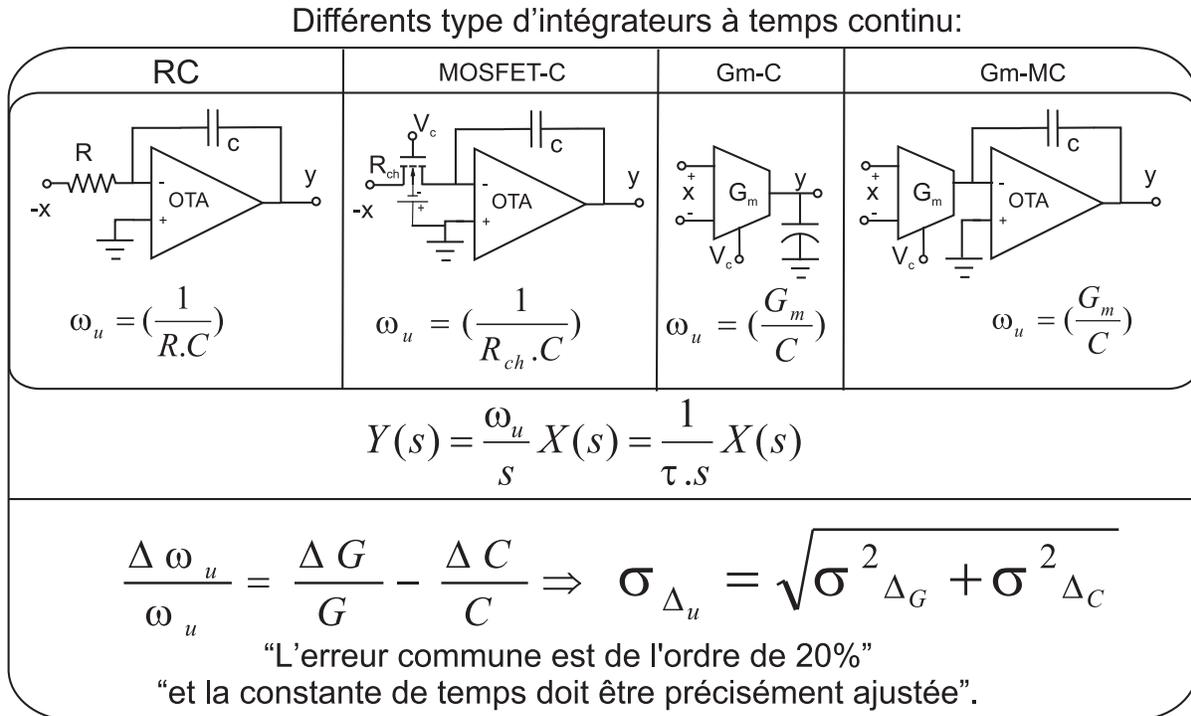


FIG. 2.39 – Différents type de réalisation d'un intégrateur à temps continu

– LC-intégré

Afin de construire un modulateur passe-bande à temps continu et de fréquence élevée (fréquence du signal d'entrée supérieure à 300MHz), un filtre LC-intégré peut potentiellement être l'un des candidats favoris. En utilisant une technologie CMOS standard, les modulateurs parviennent difficilement jusqu'à ces fréquences, alors qu'une self intégrée (L) ne peut pas être réalisable en basses fréquences. La technologie avancée comme celle de 90nm-CMOS est un des candidats, mais le meilleur candidat reste le BICMOS. Le développement actuel de la technologie CMOS qui est à la fois peu coûteuse et fiable, et de consommation raisonnable, pourra être la première à être choisie à l'avenir afin de réaliser des circuits radio-fréquences pouvant atteindre 3GHz, fréquences des récepteurs portables, GSM, ADSL, VDSL. Un autre travail de thèse met en œuvre différents aspects du modulateur passe-bande en utilisant des filtres à la base de LC-intégré dans notre laboratoire [24].

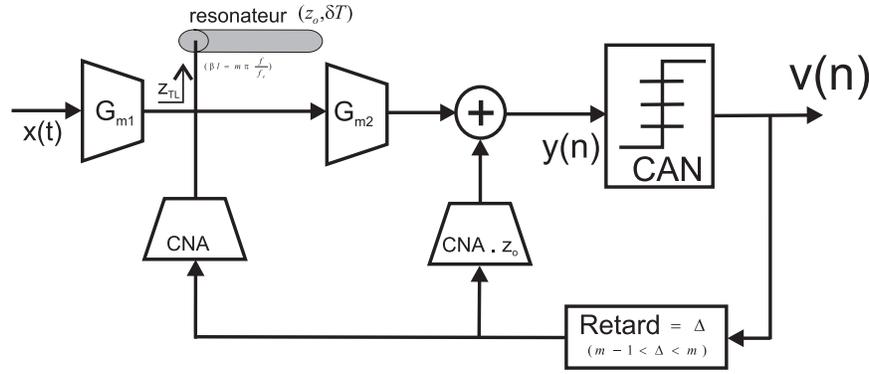


FIG. 2.40 – Modulateur $\Sigma\Delta$ en utilisant ligne de transmission TL

– TL, ou câble-coaxial

Cette méthode est utilisée dans les modulateurs à temps continu fonctionnant à des fréquences beaucoup plus élevées (au-delà de 1GHz). C'est une des méthodes efficaces contre l'effet de la gigue d'horloge dans les modulateurs à temps continu. Ces méthodes sont pour l'instant en phase de recherche et de développement. La limite principale est l'implantation de n'importe quelle sorte de guide d'onde (par exemple celle de ligne de transmission TL) sur des puces électroniques de technologie CMOS. La figure 2.40 montre un exemple de cette architecture récemment utilisée dans la réalisation de modulateurs à haute performance [25, 26, 59]. Le principe sera mieux compris si on rappelle l'équation de correspondance de l'impédance d'un résonateur de TL avec une longueur $\beta l = m\pi \frac{c}{f_e}$ illustrée sur la figure 2.40 :

$$\begin{aligned}
 Z_{TL}(s) &= \frac{V_{TL}(s)}{I_{TL}(s)} = Z_o \frac{1 \pm e^{-2s\delta T}}{1 \mp e^{-2s\delta T}} && \text{connecté à la masse / ouvert} \\
 Z_{TL}(z) &= Z_o \frac{1 \pm z^{-m}}{1 \mp z^{-m}}, && \text{; en mettant } z \equiv e^{sT} \text{ pour } \approx \frac{f_e}{4} \\
 NTF_{TL}(z) &= \frac{1 \pm z^{-m}}{1 \mp (1 \pm 2 |CNA(z)| z_o) z^{-m}}, && \text{équivalant en } z \text{ à } \frac{f_e}{4}
 \end{aligned} \tag{2.54}$$

En résumé, un modulateur réalisé par TL se comporte :

- comme un modulateur à temps continu car la partie d'échantillonnage se trouve juste avant le quantificateur.
- comme un modulateur à temps discret car des résonateurs et des intégrateurs de base sur les circuits TL ont des fonctions de transfert similaires à celles à temps discret.

2.5 Historique et évolution des modulateurs $\Sigma\Delta$

Bien que l'idée d'utiliser une rétroaction pour améliorer la résolution des convertisseurs analogique numérique (modulation delta) a été inventée durant les années 50, le principe de la modulation $\Sigma\Delta$ a été proposé en 1962 [60]. En 1963 le même auteur présentait une implantation réelle d'un modulateur

$\Sigma\Delta$ passe-bas.

Douze ans plus tard, Ritchie proposait les modulateurs d'ordre élevé [61]. Le problème de stabilité s'est alors posé et plusieurs recherches ont été consacrées autour de ce sujet [62]. Basés sur ses résultats, plusieurs modulateurs d'ordre élevé avec une cascade d'intégrateurs ont été réalisés en temps discret en utilisant des circuits à capacités commutées et des sources de courant commutées. Le principe de la structure MASH appliqué au départ sur les CNA $\Sigma\Delta$ [63], a été introduit pour la première fois sur les CAN $\Sigma\Delta$ en [64].

La modulation $\Sigma\Delta$ multibit a été proposée par Larson [13], et l'utilisation d'un étage d'appariement dynamique des cellules pour corriger les défauts d'appariement (*dynamic element matching*) a été introduit dans la structure interne du CNA de la boucle du modulateur par Plassche [6] et plus tard par Carley [7, 65]. De nombreuses publications suggéraient des algorithmes de correction des erreurs de non linéarité du CNA de la boucle du modulateur [8–12, 66–69].

Avec le développement des communications sans fils, les modulateurs $\Sigma\Delta$ passe-bande dérivés des modulateurs passe-bas sont apparus à la fin des années 80 [22]. Juste après, les modulateurs à temps continu ont attiré beaucoup l'attention des chercheurs [70, 71] en raison de la grande vitesse des filtres à temps continu comparés aux filtres à capacités commutées. Motivées par le besoin d'étendre encore plus l'intervalle de fréquence sans réduire les performances du modulateur, de nombreuses recherches ont été effectuées sur les modulateurs $\Sigma\Delta$ passe-bande à temps continu [4, 72].

Les performances (SNR , $SNDR$) et les spécifications telles que la fréquence du signal d'entrée (f_o), la fréquence d'échantillonnage (f_e), la largeur de la bande (BW), varient avec l'ensemble des performances technologiques. La limite de la fréquence maximale d'échantillonnage imposée par la technologie, peut être compensée en augmentant le nombre de niveaux de quantification interne si le défaut d'appariement de CNA interne est bien corrigé. En revanche, l'augmentation du nombre de niveaux de quantification interne améliore la résolution du modulateur, mais, compense peu la limite de la fréquence d'échantillonnage. Avec les technologies courantes, il est difficile pour les modulateurs $\Sigma\Delta$ passe-bande conventionnels d'opérer à des fréquences IF ou RF dans l'intervalle de $70MHz$ à quelques centaines de MHz . Bien qu'il y ait eu plusieurs publications sur la théorie et la conception des modulateurs $\Sigma\Delta$ passe-bande, la fréquence centrale des modulateurs réalisés en technologie CMOS restent dans la gamme de $455KHz$ à $20MHz$, avec les avancées les plus récentes de 40 à $100 MHz$. Les recherches actuelles sont plutôt consacrées au développement des modulateurs multibit à haute résolution pour qu'ils soient plus robuste aux imperfections des circuits, ainsi que le développement des modulateurs à temps continu pour qu'ils résistent à l'effet de la gigue d'horloge. En plus afin d'obtenir un modulateur large bande, on essaie de développer des architectures parallèles et des techniques à banc-filtres hybrides. Le tableau 2.2 rassemble les performances et les caractéristiques des modulateurs $\Sigma\Delta$ réalisés récemment.

TAB. 2.2 – Performances et caractéristiques des modulateurs $\Sigma\Delta$ multibits publiés ces dernières années.

Réf. date.	ordre, type	Band	OSR	f_e (MHz)	CNA (B-bit)	correction méthode	DR [dB]	technologie
[48] 04	4 th -CT	LP	10	300	4	DEM (RDEM)	67	0.13u-CMOS
[73] 04	3 th -CT	LP	32	35	5	calibration	88	0.5u-CMOS
[74] 03	2 th -CT	LP	24	48	4	Rotation(BS)	63	0.18u-CMOS
[75] 00	3 th -DT	LP	24	60	4	DWA	89	0.65u-CMOS
[76] 01	5 th _{MASH-DT}	LP	16	64	5	P-DWA	95	0.5u-CMOS
[77] 02	2 th _{MASH-DT}	LP	4	1	1+10	digital-estim	85	0.5u-CMOS
[78] 04	4 th -DT	BP	64	20	4	SDEM-1 st ord.	x+30	0.35u-CMOS
[79] 98	3 th -DT	LP	100	5	3	SDEM-2 th ord.	80	0.6u-CMOS
[80] 00	3 th -DT	LP	128	10.7	5	SDEM-2 th ord.	120	0.35u-CMOS
[81] 01	3 th -DT	LP	64	x	5	TDEM-2 th or.	98	0.35u-CMOS
[82] 99	8 th -DT	BP	x	0.5	4	BF.Sh.-2 th ord.	90	test-FPGA
[83] 02	6 th -DT	BP	64	x	3	SDEM-2 th ord.	79	0.25u-CMOS
[84] 00	3 th -DT	LP	192	0.192	5	DWA-partial	120	0.5u-CMOS
[85] 04	5 th DT	LP	128	6.14	4	SDEM-2 th ord.	120	0.35u-CMOS
[86] 99	4 th -CT	LP	128	13	1	DEM (RDEM)	86	0.35u-CMOS
[44] 05	3 th -CT	LP	48	4.8	1	SCR	80	0.5u-CMOS
[87] 04	2 th -CT	LP	812	2000	1	non	79	0.18u-CMOS

2.6 Conclusion

Les notions fondamentales de la modulation sigma-delta ont été présentées dans ce chapitre. Après, la description de sa structure et de la théorie associée (suréchantillonnage et la mise en forme du bruit de quantification), les différents paramètres définissant sa structure (ordre, facteur de suréchantillonnage, résolution de quantificateur interne), les propriétés évaluant ses performances (SNR , $SNDR$, $SFDR$, DR , $ENOB$, THD , $SFDR$ et FOM) ainsi que les critères de stabilité ont été donnés. Les performances d'un modulateur $\Sigma\Delta$ et la stabilité de la boucle sont directement liées à l'architecture du modulateur et au signal d'entrée. Nous avons décrit les différentes architectures à une seule boucle ou en cascade. Nous avons également décrit les modulateurs $\Sigma\Delta$ passe-bande et à temps continu et l'intérêt qu'ils présentent en haute fréquence comparés à leurs équivalents à temps discret. Ensuite les différentes techniques et technologies de réalisation ont été brièvement rappelées. Dans le reste de ce document, nous nous limitons sur l'histoire et l'état de l'art des modulateurs sigma-delta multibit publiés ces dernières années. Celui-ci rassemble les réalisations réalisées en technologies CMOS.

Le tableau-2.3 rassemble les différentes catégories de modulateurs sigma delta et leurs avantages et inconvénients. Pour conclure, on constate que la plupart des inconvénients des différentes sortes du modulateur peuvent être réduits en utilisant la quantification multibit (CAN et CNA interne) si le défaut d'appariement des cellules du CNA interne est corrigé. Au cours des prochains chapitres, nous vérifierons les avantages et inconvénients des différentes solutions actuelles et nous proposerons quatre nouvelles méthodes de correction des défauts d'appariement du CNA.

TAB. 2.3 – Différents catégories de $\Sigma\Delta$ modulateur et leurs limites

Modulateur	avantages	inconvenients	solutions
1) mode de filtre • temps discret (DT) • temps continu (CT)	-réalisable par SC et SI -peu sensible à la gigue - f_e élevé, (<2GHz) -faible consommation -faible tension	f_e limitée, <100MHz BW très bas sensible au bruit $\frac{KT}{C}$ erreur de SAH, CFT,.. -très sensible à la gigue - et au retard	passer au CT *multibit *multibit rebouclage court ou multiple
2)CAN et CNA • monobit -CT: • multibit	peu sensible aux défauts du CNA *moins sensible à la gigue *DR élevé, modeste OSR , *meilleure stabilité, *gain de la boucle élevé, *moins contraint à l'AOP, *faible consommation *post-filtration facile	-faible SNR -tons parasites très sensible à la gigue -et au dynamique d'AOP -besoin fort OSR -non-linéarité de CNA	*multibit *multibit *NRZ-multibit *multibit SCR ou Sin. !!! Correction, par ex. : DEM, calibrage, etc.
3)seule boucle -en cascade	+moins sensible aux +imperfection de circuit +souhaitable avec SC, et pour DR moyenne	-problème de stabilité pour $m>2$ - sensible aux imperfections de circuit	concept. stricte *multibit calibrage
4)réalisation • SC • SI • Gm-C,Gm-MC • LC-intégré • TL	+précise(0.1%), faible coût +faible consommation +faible tension,faible surface +compatible avec CMOS +très faible consommation +souhaitable $1.2 < f_e < 4\text{GHz}$ +bonne linéarité +moins sensible à la gigue	- f_e limitée -imprécise et f_e limitée -CFT injection de charge -imprécise(>10%), -linéarité limitée -grande consommation f_e limitée -difficile à réaliser -très sensible aux valeurs L et de C -impraticable en CMOS	*multibit BICMOS ou CMOS-avancée utiliser d'autre technologies
5) ordre de Mod. • m=1 • m=2 • m=3-5	+stable + stable +peu de tons, bon SNR	-tons, mauvais SNR -tons, faible SNR -instabilité, -complexité	choisir $m>1$ *multibit *multibit et concept. stricte

Chapitre 3

Sources d'erreur dans un convertisseur $\Sigma\Delta$ multibit

3.1 Introduction

La fonction essentielle d'un modulateur de type delta sigma est de diminuer le bruit de la quantification dans la bande utile. La relation concernée du SNR a été calculée dans le chapitre précédent. Le SNR maximal est donné par la relation suivante, qui reprend l'équation 2.29 :

$$SNR_{max} = 10 \text{Log} \left[\frac{3(2^B)^2 G^2 (2L + 1) OSR^{(2L+1)}}{2\pi^{2L}} \right] \quad (3.1)$$

Théoriquement, pour des applications qui exigent une haute résolution, le rapport SNR_{max} peut être amélioré en faisant varier correctement les paramètres B (le nombre de bits du quantificateur), L (l'ordre de modulateur)¹, OSR (le facteur de suréchantillonnage) et G (le gain de la boucle de modulateur).

Mais en pratique, le bruit de quantification n'est pas le seul facteur qui limite la performance. D'autres phénomènes interviennent, comme :

1. le bruit thermique et le bruit en 1/f :
 - dans l'échantillonneur/bloqueur ($\frac{KT}{C}$),
 - dans l'intégrateur ou dans le résonateur SC, ($\frac{KT}{C}$ ainsi que $\frac{KT}{g_m}$),
 - dans le filtre à temps continu ($\frac{KT}{g_m}$)
2. la non-linéarité associée aux filtres (de l'intégrateur, du résonateur, du sommateur et des éléments passifs R,C,L),
3. la non-linéarité du quantificateur (CAN),
4. la non-linéarité du convertisseur numérique-analogique interne (CNA),
5. la gigue d'horloge,
6. la non-idéalité de la partie du traitement numérique (décimation et filtration),
7. le retard de la boucle,

1. L est en fait l'ordre du modulateur passe-bas. Mais pour les applications passe-bandes, il faut utiliser le terme m au lieu du L quand $L = 2m$ est l'ordre du modulateur.

8. les tons parasite ("idle-tones").

Tous ces phénomènes sont des sources d'erreurs qui sont potentiellement capables de dégrader les performances théoriques du modulateur. La figure 3.1 représente les emplacements présumés des erreurs citées ci-dessus dans un modulateur $\Sigma\Delta$ multibit. Leurs effets peuvent être comparés avec

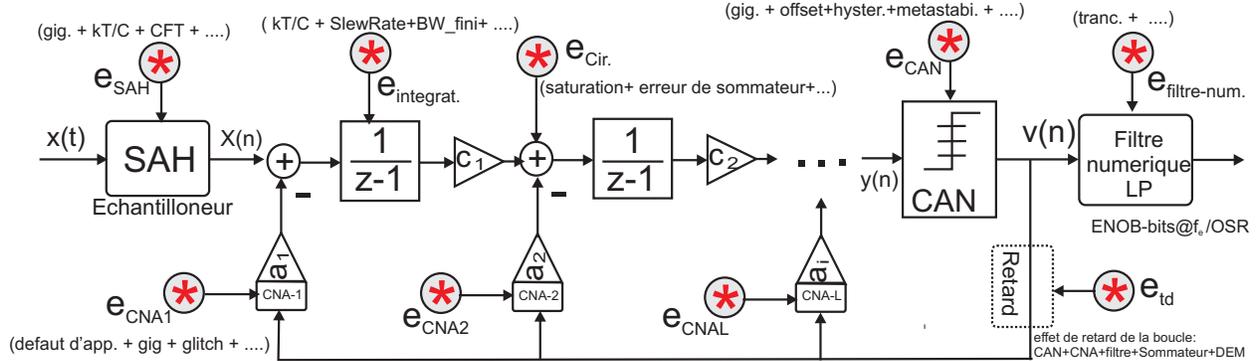


FIG. 3.1 – lieux présumés des erreurs du modulateur $\Sigma\Delta$

celui du bruit de quantification dans la bande utile (équation 2.29) :

$$\sigma_{q_{rms}} = \sqrt{\frac{\Delta^2}{12} \cdot \frac{G^2 \cdot \pi^2 L}{(2L+1)OSR^{(2L+1)}}} \quad (3.2)$$

Par conséquent, la performance d'un modulateur $\Sigma\Delta$ ayant une entrée sinusoïdale, une sortie normalisée $0 \leq v(n) \leq 2^B \Delta$, et $\Delta = 1$, peut être exprimée par la relation suivante :

$$SNR_{max} = \frac{(\frac{2^B}{2\sqrt{2}})^2}{\sigma_{N_{total}}^2} |_{dB} = 20 \log \left[\frac{2^B}{\sqrt{8} \cdot \sigma_{N_{total}}} \right] \quad (3.3)$$

où le bruit total du système N_{total} doit compter toutes les sources d'erreurs, par exemple, le bruit de quantification σ_q , le bruit d'échantillonnage $\sigma_{kt/c}$, le bruit d'AOP σ_{op} , l'erreur du CNA σ_{CNA} , l'erreur de gigue d'horloge σ_g , et autres selon l'architecture utilisée. Cela peut s'exprimer selon la formule :

$$\sigma_{N_{total}} = \sqrt{\sigma_q^2 + \sigma_{kt/c}^2 + \sigma_{op}^2 + \sigma_{CNA}^2 + \sigma_g^2 + \sigma_{autre}^2} \quad (3.4)$$

La résolution finale du modulateur $ENOB = \frac{DR_{dB} - 1.76}{6.02}$ est alors donnée par² :

$$ENOB = \text{Log}_2 \left[\frac{2^B}{2\sqrt{3} \cdot \sigma_{N_{total}}} \right], \text{ en bit} \quad (3.5)$$

Pour une meilleure analyse, on préfère exprimer séparément la résolution du système pour chacune des erreurs quand les autres erreurs sont négligées, à condition que les sources d'erreurs sont décorréelées. En outre, la valeur du bruit dans la bande d'une source de non idéalité du système doit être estimée puis remplacée dans l'expression de $\sigma_{N_{total}}$ dans l'équation 3.5.

2. En fait, cette formule a été définie pour un convertisseur analogique numérique ayant une erreur de quantification $\sigma_q^2 = \frac{\Delta^2}{12}$, où $\Delta \equiv \frac{F.S.}{ENOB}$. Cela touche son SNDR maximum à $x(t) = \frac{2^B}{2} \sin(2\pi ft)$, c.-à-d. la puissance plein d'échelle vaut $(\frac{2^B}{2\sqrt{2}})^2$. Cependant afin de donner une expression comparable, elle est couramment utilisée pour le CNA [2,88-90].

Chacune des sources d'erreur peut avoir plusieurs causes différentes ainsi que plusieurs corrections possibles. D'autre part, une source d'erreur peut avoir des effets différents selon l'application (passe-bas, passe-bande ou passe-haut), l'architecture (monobit, multibit, seul étage ou cascade), le type de modulateur (à temps discret DT ou à temps continu CT) et la technologie (CMOS, BICMOS, AsGa).

Dans ce chapitre, nous allons étudier brièvement les causes d'erreurs importantes dans les modulateurs $\Sigma\Delta$ multibits, ainsi que les corrections envisageables. Nous allons aussi discuter de leur différence selon l'application et les types de modulateur, séparément. Évidemment, cela n'est qu'une vue globale et brève. Plus de détails seront donnés dans les chapitres suivants ou dans les références indiquées.

3.2 Le bruit thermique et le bruit en $1/f$

Une des limitations principales de la performance des circuits électroniques est le bruit intrinsèque. Il y a deux bruits intrinsèques importants dans les circuits électroniques, en particulier dans les transistors MOSFET : le bruit thermique et le bruit en $1/f$ (*flicker*). Le bruit thermique est provoqué par le mouvement thermique des porteurs de charge (électrons ou trous) dans le canal du dispositif (canal entre le drain et la source). Ceci entraîne une fluctuation aléatoire dans le courant de drain. Si le transistor fonctionne dans sa région ohmique (*triode*), comme il le fait pour un commutateur, le bruit peut être représenté par une source de tension en série avec le transistor. La densité spectral de puissance (PSD) du bruit thermique de commutateur est blanche et sa valeur estimée est indiquée comme suit [91] :

$$\begin{aligned} S_{vt}(f) &= 4k_B T R_{on}, \left(\frac{V^2}{Hz} \right) \\ R_{on} &= \frac{1}{\eta C_{ox} W L} (V_{GS} - V_{Th}) \end{aligned} \quad (3.6)$$

où, k_B est la constante de Boltzmann, $k_B = 1.38 \times 10^{-23} \text{ J/k}$, T est la température absolue du canal en degrés Kelvin, R_{on} est la résistance du canal en ohms, C_{ox} est la capacité de la surface unité de la grille, η est un paramètre de fabrication, et W et L sont les dimensions du canal. La valeur moyenne du bruit thermique est toujours égale à zéro, sinon ce bruit peut servir de source d'énergie électrique. Pour un transistor MOSFET fonctionnant dans sa région active, le bruit thermique peut être modélisé par une source de courant parallèle au canal. La PSD du courant de bruit est approchée par :

$$S_{it}(f) = \frac{8 g_m k_B T}{3} \left(\frac{A^2}{Hz} \right) \quad (3.7)$$

où, $g_m(\propto \sqrt{I_D} \cdot \sqrt{\frac{W}{L}})$ est la transconductance du canal.

Le bruit en $1/f$ est généré par des porteurs de charge capturés par des atomes cristallins du canal et plus tard libérés. Il est habituellement modélisé par une source de tension de bruit série reliée à la grille. La PSD de cette tension est approximativement donnée comme suit :

$$S_{vf}(f) = \frac{K_f}{W L f}, \left(\frac{V^2}{Hz} \right) \quad (3.8)$$

où, K_f est un paramètre de technologie ($\approx 10^{-28} (F/m)^2$). On note que le bruit en $1/f$ n'est pas blanc; la majeure partie de sa puissance est concentrée autour du continu ($f \rightarrow 0$). Dans beaucoup de cas, l'effet du bruit en $1/f$ peut être réduit en utilisant un dispositif d'entrée de grande taille (c.-à-d. de grandes longueurs et largeurs W et L), et en choisissant des transistors PMOS plutôt que des transistors NMOS à l'entrée du système [1,91]. Cependant afin d'obtenir plus de précision pour les applications à fréquence très basse, l'effet du bruit en $1/f$ peut être éliminé par la méthode de *chopper* ou double échantillonnage corrélé. [1,91,92].

Dans cette section nous étudions seulement l'effet du bruit thermique dans les circuits d'un échantillonneur et d'un intégrateur (ou résonateur) de type SC. En effet, ce dernier peut représenter à lui seul la majorité du bruit dans un modulateur à capacités commutées. La figure 3.2 montre un exemple de la contribution de bruit thermique dans le cas pratique, où le seul bruit de l'intégrateur SC ($\sigma_{kt/c}^2$) occupe trois quart de la puissance dans la bande ($\sigma_{N_{total}}$) de l'expression 3.4.

De plus, mêmes si il n'est pas un obstacle sérieux pour un filtre à temps continu, l'effet du bruit

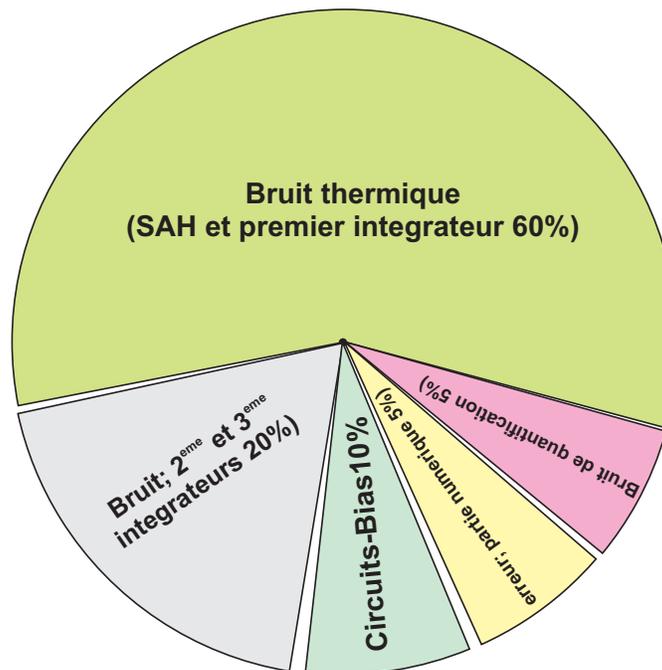


FIG. 3.2 – Exemple de la contribution des différents bruits dans un modulateur $\Sigma\Delta$, à la réalisation SC

thermique sera aussi étudié dans le cas d'un amplificateur (*AOP*) employé dans un modulateur CT.

3.2.1 Effet de bruit thermique au niveau de l'échantillonneur

La conversion d'un signal analogique en un signal numérique se fait en général en deux étapes: l'échantillonnage puis la quantification numérique, comme cela est représenté sur la figure 3.3. Dans un premier temps, un bloc, nommé échantillonneur-bloqueur ("*Sample-And-Hold*", *SAH*), suit régulièrement le signal d'entrée aux instants d'échantillonnage (fréquence f_e). Ensuite, la sortie du bloqueur $x_o(nT)$ est convertie en une valeur numérique dans un processus de quantification. La quantification

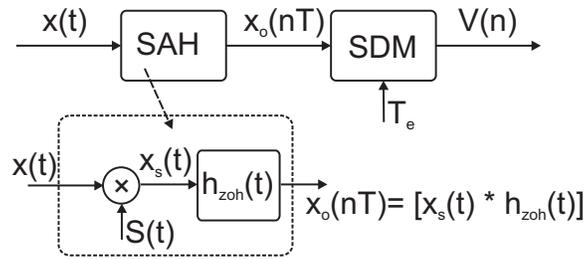


FIG. 3.3 – Chaîne de conversion analogique-numérique (sans décimation)

peut s'effectuer de plusieurs manières, par exemple par la modulation $\Sigma\Delta$. L'échantillonneur-bloqueur est indispensable dans la plupart des convertisseurs analogique-numérique (A/N). Il doit conserver le signal d'entrée analogique à un niveau constant pendant le temps de conversion.

En effet pour des raisons de précision et de simplicité, l'échantillonnage-bloquage est souvent réalisé à l'entrée du premier étage du filtre où une sortie de la branche de contre-réaction est aussi l'entrée. Dans le cas d'un modulateur à temps continu, l'échantillonneur-bloqueur est réalisé à l'intérieur du modulateur juste en amont du quantificateur. Cela rend les effets dégradants de l'échantillonneur-bloqueur beaucoup moins sévère dans le cas continu que dans le cas discret.

Dans cette section, le comportement d'un échantillonneur-bloqueur est décrit; une description plus détaillée est donnée dans [91, 93].

Un échantillonneur/bloqueur idéal devrait conserver le signal d'entrée $x(t)$ jusqu'à ce que la conversion soit finie. La sortie du SAH, nommée $x_o(nT_e)$, peut être alors définie par l'équation ci-dessous :

$$x_o(nT) = x(t) |_{t=nT_e}, \quad \text{quand : } nT_e \leq t \leq (n+1)T_e \quad (3.9)$$

où n est le nombre d'échantillons et $T_e = \frac{1}{f_e}$ est la période d'échantillonnage. Le terme nT , qui indique l'instant d'échantillonnage, est souvent simplement remplacé par n . Afin de mettre en équation le fonctionnement d'un SAH, on peut supposer que l'entrée $x(t)$ est en fait multipliée par une série d'impulsions de Dirac. Puis, la valeur obtenue est maintenue jusqu'à la fin de la période comme cela est expliqué par la suite :

$$\begin{aligned} s(t) &= \sum_{-\infty}^{\infty} \delta(t - nT_e), \\ x_s(t) &= x(t) \sum_{-\infty}^{\infty} \delta(t - nT_e) \\ x_o(n) &= x_s(t) * h_{zoh}(t), \\ h_{zoh}(t) &= \begin{cases} 1 & \text{quand } 0 < t < T_e \\ 0 & \text{quand } t \notin]0, T_e[\end{cases} \end{aligned} \quad (3.10)$$

où, $\delta(t - nT_e)$ dénote une série d'impulsions et $h_{zoh}(t)$ est une fenêtre qui marque la valeur échantillonnée dans chaque période. Dans le cas d'une entrée sinusoïdale, c.-à-d. $x(t) = A_{in} \cdot \sin(2\pi ft)$, la sortie peut être représentée comme sur la figure 3.4.

Le signal échantillonné $x_s(t)$ présente un spectre périodique selon un pas fréquentiel de f_e dont le

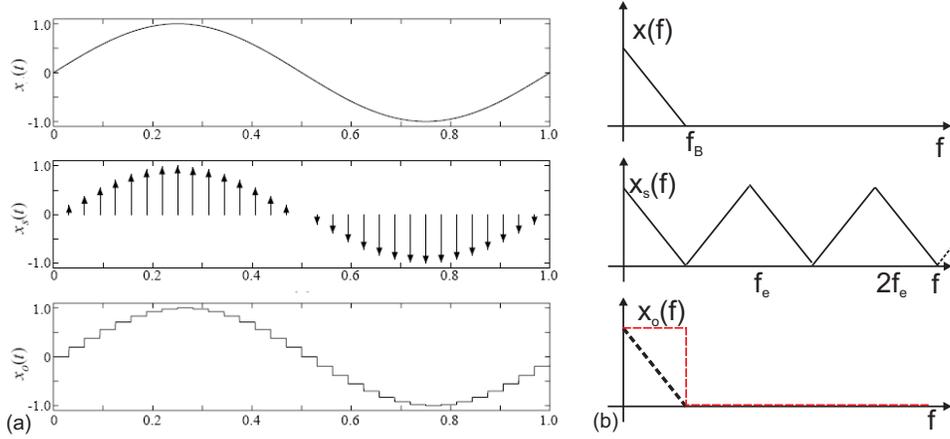


FIG. 3.4 – a) échantillonnage, b) spectre du signal; suivant les différentes étapes d'échantillonnage avec un bloquer idéal

motif de base est le spectre du signal analogique. $x(t)$ peut être reconstitué à l'aide d'un filtre passe-bas s'il n'y a pas de recouvrement des spectres, c'est à dire, si la condition $f_e > 2f_B$ est vérifiée. Un repliement de spectre peut aussi avoir lieu si un bruit haute fréquence non porteur d'information se superpose au signal d'entrée. C'est la raison pour laquelle, d'après le théorème d'échantillonnage développé par *Shannon* en 1949 [94], un filtre anti-repliement en tête du chemin de conversion est nécessaire afin d'éviter ce risque. La transformée de *Fourier* de $x_s(t)$, de $h_{zoh}(t)$ et l'ensemble de $x(n)$ s'écrivent :

$$\begin{aligned}
 X_s(f) &= f_e \sum_{i=-\infty}^{\infty} X(f - if_e) \\
 H_{zoh}(f) &= \frac{2\sin(\pi fT)}{2\pi fT} e^{-j\pi fT} \\
 X_o(f) &= X_s(f) \cdot H_{zoh}(f), \\
 |X_o(f)| &= |X_s(f)| \cdot \frac{|\text{sinc}(\frac{\pi f}{f_e})|}{f_e}
 \end{aligned} \tag{3.11}$$

Le blocage de $x_o(n)$ durant la période T_e se traduit mathématiquement par un sinus cardinal dans le domaine fréquentiel. Par conséquent, le spectre en bande de base n'est pas filtré, comme le montre la figure 3.4-b, mais il subit tout de même une erreur de gain si sa fréquence maximale f_B est juste inférieure à $\frac{f_e}{2}$. Ce phénomène nécessite un suréchantillonnage afin d'obtenir de bonnes performances, ce qui en pratique se produit dans un modulateur $\Sigma\Delta$.

Un échantillonneur-bloqueur introduit naturellement des erreurs. Notamment, pour les signaux haute fréquence, les limites résident dans le suivi du signal et dans les erreurs de non-linéarité. Un circuit simple de SAH, en technologie CMOS, peut être conçu avec des transistors et des capacités de transistor MOSFET comme cela est présenté sur la figure 3.5. Pendant la phase d'échantillonnage, le transistor M1 est activé ("on", région linéaire). Alors, la capacité C est chargée par $x(t)$. Dans la phase de blocage, c.-à-d. $\overline{\Phi_1} = 1$, la tension de capacité est isolée du signal d'entrée et reste constante à la sortie. Généralement pour établir un bon SAH, les transistors NMOS sont souvent utilisés pour

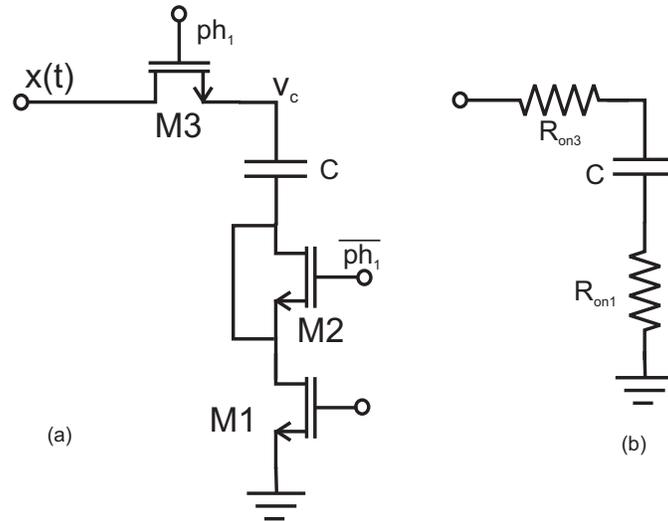


FIG. 3.5 – a) La structure d'un échantillonneur/bloqueur, b) le modèle électrique du SAH en phase d'échantillonnage

atteindre une vitesse plus élevée, les résistances R_{on} sont plus faibles que celles des transistors PMOS. Un transistor fantôme³ M2 peut être utilisé afin de compenser les charges accumulées dans le canal aux instants de commutation. Pour réaliser la modélisation, nous nous intéressons au modèle simple du transistor en régime linéaire (triode), ainsi qu'en régime bloqué ("off"). Le transistor agit en tant que commutateur ($R_{on} \simeq 0$). Cependant, le transistor réel conduit quand $V_{GS} > V_{Th}$, où V_{Th} est la tension de seuil, puis présente une résistance (R_{on}) (expression 3.6). Les performances du modulateur peuvent être détériorées par la non-linéarité du circuit de SAH. De nombreuses études sur des méthodes de compensation de ces erreurs ont été menées [86, 88, 89]. Les trois sources d'erreur importantes dans un SAH sont les suivantes :

- le bruit thermique
- l'injection de charge et CFT ("*Clock Feedthrough*"),
- la gigue d'horloge

Le modèle électrique du SAH en phase d'échantillonnage est illustré à la figure 3.5-b. Afin de calculer la puissance du bruit thermique, on ajoute une source de tension en série à chaque résistance dans les deux phases d'échantillonnage. Quand ϕ est haut, le commutateur M1 est fermé et la densité spectrale du bruit est :

$$v_{n,in}^2(f) = 4k_B T R_{on} \quad (3.12)$$

La variance du bruit échantillonné dans la capacité C_H est déterminée par l'intégrale de la densité spectrale de $v_{n,in}^2(f)$ multipliée par la fonction de transfert du filtre passe-bas⁴ de réseau $R_{on}C_H$,

3. *Dummy transistor*

4. Comme la largeur de la bande de bruit blanc n'est pas équivalente de celle de $-3dB$ du filtre, cette intégration ne peut pas être simplifiée par une multiplication simple de $f_{-3db} = \frac{1}{2\pi RC}$. Cependant, cette intégration peut être remplacée, ici, par une multiplication par la largeur effective de bruit blanc de filtre passe-bas qui est $\frac{\pi}{2} \cdot f_{-3db} = \frac{1}{4RC}$.

$\left| \frac{1}{(1+jR_{on}C_H\omega)} \right|^2$:

$$\begin{aligned} \overline{v_{n,out}}^2 &= \int_0^\infty 4k_B T R_{on} \frac{df}{1 + (2\pi f R_{on} C_H)^2} = \frac{k_B T}{C_H} \\ v_{n,out,rms} &= \sqrt{\frac{k_B T}{C_H}} \end{aligned} \quad (3.13)$$

où, l'effet du bruit qui est induit par la résistance du commutateur ne dépend plus de la valeur de sa résistance R_{on} , et dépend seulement de la valeur de la capacité C_H . Cela signifie que l'effet du bruit thermique dans un échantillonneur-bloqueur peut être diminué quand on augmente la valeur de capacité. Par exemple dans un convertisseur avec une résolution finale de ENOB, on peut estimer la capacité minimum à utiliser dans la chaîne d'échantillonnage (cas simple), en comparant le bruit thermique avec l'erreur de quantification supposée $\sigma_q^2 = \frac{LSB^2}{12}$, $LSB = \frac{V_{FS}}{2^{ENOB}}$, quant V_{FS} est la tension plein échelle de sortie :

$$C_{Hmin} > \frac{12 \cdot K_B T \cdot 2^{2 \cdot ENOB}}{V_{FS}^2} \quad (3.14)$$

Si on limite la puissance de bruit thermique à un quart de celle du bruit quantification (ce qui correspond à une perte d'approximativement 1dB en SNDR final) on en déduit une valeur minimale de la capacité d'échantillonnage: $C_{Hmin} \geq \frac{48 \cdot K_B T \cdot 2^{2 \cdot ENOB}}{V_{FS}^2}$. On obtient par exemple, pour $ENOB = 16$ bit et $V_{FS} = 1 V$, une capacité minimum de $853pF$.

Pendant la deuxième phase ($\phi = 0$), la charge accumulée dans la capacité reste indépendante de la variation d'entrée, et peut être distribuée aux circuits suivants. Un éventuel courant de fuite peut introduire une erreur supplémentaire sur la valeur échantillonnée via le transistor M_1 qui est dans l'état fermé, mais elle est heureusement négligeable dans la plupart des cas en technologie CMOS. Dans le cas de suréchantillonnage du signal avec un facteur OSR , la puissance résultante après filtrage numérique est réduite de ce même facteur, ainsi la capacité minimum diminue du fait de l'OSR.

D'autre part afin de minimiser l'injection de charge CFT, le couplage capacitif entre la commande ϕ et la sortie, et la dimension de transistor doit être diminué. L'effet d'injection de charge dans un commutateur peut être mieux minimisé si on y ajoute un transistor fantôme similaire en série avec un court circuit entre son drain et sa source qui est excité par le commande $\bar{\phi}$ [91]. De plus le maximum et le minimum d'amplitude du signal de commande ϕ doivent se rapprocher de sorte que le courant de fuite reste assez faible.

Une autre erreur qui peut influencer le signal échantillonné est induite par l'incertitude du moment d'échantillonnage (gigue d'horloge). L'effet de la gigue d'horloge peut être estimé par la vitesse de variation de la sortie $\Delta v_{out}(n) = v_{out}(n) - v_{out}(n-1)$ et le taux de la variation du temps d'échantillonnage $\frac{\Delta T_e}{T_e}$ comme suit :

$$P_{\Delta T_e} = \Delta^2 v_{out}(n) \frac{\delta T_e}{T_e} \quad (3.15)$$

Pour un signal sinusoïdal, le pas de sortie est proportionnel à la fréquence du signal ainsi qu'à son amplitude, comme le montre la figure 3.6. La puissance maximale de bruit introduite par la gigue

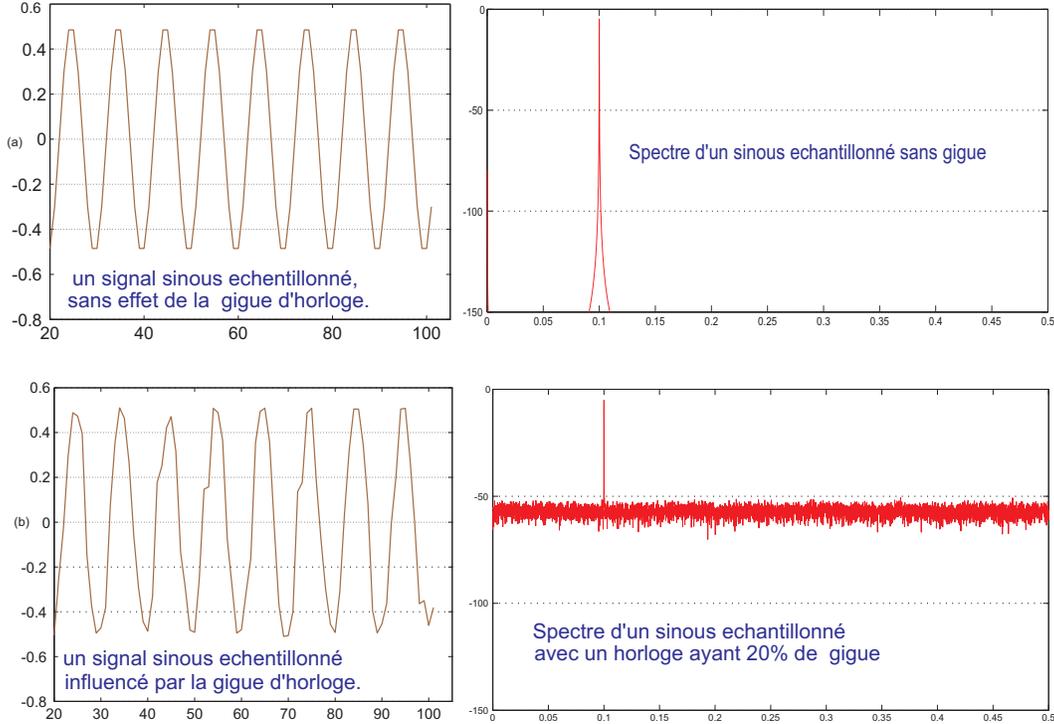


FIG. 3.6 – Effet de gigue d'horloge sur un signal sinusoïdal échantillonné

d'horloge (à l'équation 3.15) doit être inférieure à l'erreur de quantification du système (σ_q^2). On suppose que le pas entre deux échantillons peut augmenter jusqu'à la pleine échelle (pire cas). En l'absence de suréchantillonnage ($OSR=1$), on obtient alors :

$$2\pi^2 V_{FS}^2 \left(\frac{\delta T_e}{T_e} \right) < \frac{LSB^2}{12} \quad (3.16)$$

A partir de cette expression, la valeur maximale de l'erreur due à la période d'échantillonnage s'exprime par :

$$\delta T_e < \frac{T_e}{3 \times 2^{ENOB+2}} \quad (3.17)$$

On estime que la précision de phase de l'horloge d'échantillonnage peut avoir un impact majeur sur la performance du système. Il est facile de perdre des décibels de plage dynamique (résolution) pour quelques *ppm* ("part per million") de gigue. Afin d'estimer l'erreur autorisée, prenons un exemple. Une résolution de 16-bits, à une fréquence d'échantillonnage de l'ordre de 300-MHz avec un $OSR=1$, exige une précision de phase d'horloge meilleure que 4×10^{-15} (4 femto-secondes), soit 1 ppm. Cette énorme précision semble inatteignable même pour une résolution moyenne de 16-bits. On recherche, donc, des solutions notamment le suréchantillonnage en limitant la bande utile ainsi que la quantification multibit.

Une autre remarque importante est que l'effet de la gigue d'horloge dépend de la valeur absolue de la fréquence d'échantillonnage. Par conséquent, les performances d'un système passe-bande peuvent subir plus de dégradations que celles d'un système passe-bas.

3.2.2 Bruit thermique dans un amplificateur opérationnel (AOP)

Dans cette sous-section, l'évaluation du bruit thermique d'un AOP sera discutée. La figure 3.7-a montre une structure simple d'un AOP qui est composé d'un seul étage différentiel. Le bruit thermique

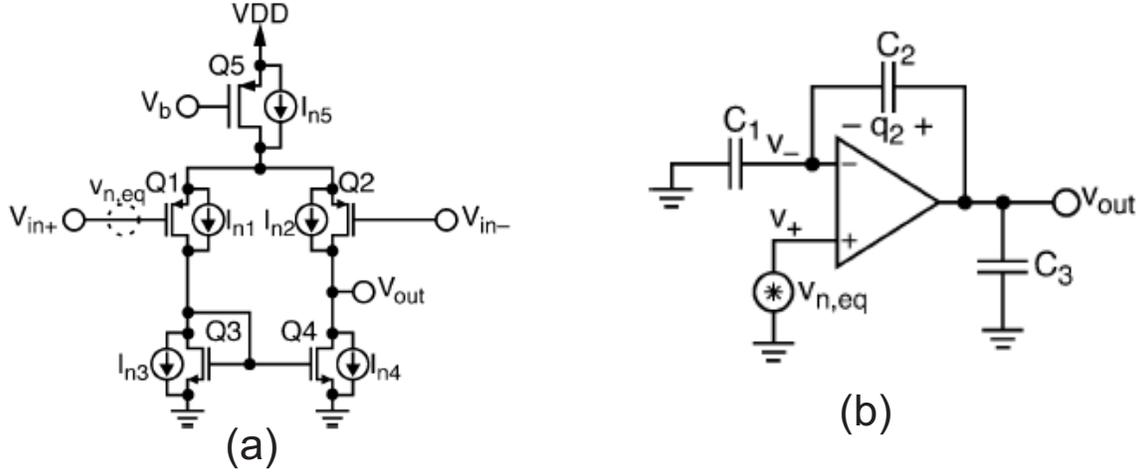


FIG. 3.7 – Structure d'un AOP simple avec ses sources du bruit thermique, b) circuit équivalent de bruit d'un AOP utilisé comme un amplificateur en combinaison rétroaction capacitif

est lié au courant du canal et est modélisé comme une source de courant en parallèle à la sortie de chaque transistor. Le bruit du transistor M5 ne contribue pas au courant de la sortie, puisqu'il est présent dans les courants de drain de M2 et de M4, et par conséquent, il est annulé dans le courant de sortie. Les sources de bruit des transistors peuvent être représentés par une source équivalente sur l'une des grille qui reçoit l'entrée principale. En utilisant l'équation 3.7, la densité spectrale du bruit équivalent ramené à l'entrée d'AOP peut être estimée comme suit [93] :

$$\begin{aligned}
 S_{vt}(f) &= \frac{16k_B T}{3g_{m1}} \left(1 + \frac{g_{m3}}{g_{m1}}\right) = \frac{16k_B T}{3g_{m1}}, \quad \left(\frac{V^2}{Hz}\right) \\
 v_{eq,rms,in} &= \frac{\pi}{2} \cdot f_B \sqrt{\frac{16k_B T}{3g_{m1}}}, \quad \left(\frac{V}{\sqrt{Hz}}\right)
 \end{aligned} \tag{3.18}$$

On remarque que la tension équivalente vis à vis de l'entrée est inversement proportionnelle à la racine carrée de g_m . Cela signifie que dans ce cas pour obtenir un AOP plus propre en terme de bruit thermique, g_m doit être augmenté. g_m à son tour peut être augmenté en augmentant la consommation ce qui n'est pas souhaitable. Alors, il faut faire des compromis entre la diminution du bruit thermique et l'augmentation de la consommation.

Cependant, l'effet exact du bruit d'un AOP doit être examiné dans le contexte du circuit qui l'emploie. Afin de donner un exemple, considérons un AOP avec une rétroaction négative (figure 3.7-b). Notons que le bruit de l'AOP est représenté par la source équivalente $v_{eq,rms,in}$, et qu'aucun signal d'entrée n'est présent puisque seul l'amplification du bruit est le sujet de l'analyse. Nous supposons que l'AOP est correctement compensé, de sorte que sa fonction de transfert peut être exprimée avec un seul pôle

comme suit :

$$G(s) = \frac{V_{out}(s)}{V_+(s)} = \frac{G_o}{1 + s\tau} \quad (3.19)$$

où, G_o est le gain d'AOP, et τ est sa constante de temps. G_o est déterminé par le facteur de rétroaction $\beta = \frac{C_2}{C_1+C_2}$ et par le gain DC de l'amplificateur à sortie en courant A_o . G_o peut être simplifié si $A_o \gg \frac{1}{\beta}$:

$$G_o = \frac{1}{\beta + \frac{1}{A_o}} \simeq \frac{1}{\beta} = 1 + \frac{C_1}{C_2} \quad (3.20)$$

τ dépend de la structure de l'AOP. Dans le cas simple de la figure 3.7-a, c.-à-d. un seul étage et sans compensation interne, τ vaut $\frac{C_o}{\beta g_{m1}}$ quand $C_o = C_3 + \frac{C_1 C_2}{C_1 + C_2}$. Ainsi la sortie peut s'exprimer par la relation suivante :

$$\overline{v_{out}^2} = \frac{4k_B T}{3\beta C_o} \quad (3.21)$$

Notez que bien que la PSD de la source de bruit de l'AOP soit déterminée par g_m , la valeur de bruit sortant d'un circuit amplificateur ne contient pas g_m . En effet, la PSD est inversement proportionnelle à g_m , alors que la largeur de bande de l'étage de compensation interne est directement proportionnelle à celle-ci. Par conséquent, elle est annulée à la sortie. Cette indépendance de g_m est une exception inattendue en regard de l'équation 3.18, car la charge et les circuits de rétroaction sont purement capacitifs. En pratique, l'effet de g_m se situe entre ces deux extrêmes, (3.18 et 3.21).

3.3 Bruit des circuits de filtre, (intégrateur)

Le filtre de boucle d'un modulateur $\Sigma\Delta$ d'ordre L est en général composé de L étages. Chaque étage peut être un intégrateur ou un résonateur, (respectivement pour un modulateur passe-bas ou passe-bande). La performance globale du modulateur peut être sensible aux non-idéalités de chacun des étages selon l'architecture et le rôle des différentes étages. Mais en général, l'influence des non-idéalités du premier étage est beaucoup plus forte que celle des suivants. Cela est particulièrement vrai pour les effets de la non linéarité et des bruits intrinsèques, parce que celles des étages suivants sont divisées par les gains des étages précédents. En revanche, certaines non idéalités des circuits doivent être séparément considérées pour tout les étages, par exemple le retard, la bande passante et la saturation.

Pour des raisons de simplicité, nous analysons les conséquences générales d'imperfection du premier étage dans le cas d'un modulateur passe-bas d'ordre L , qui a une seule boucle, représenté à la figure 3.8. Tout d'abord, considérons un intégrateur SC classique et non différentiel utilisé dans un modulateur DT, comme cela est représenté sur la figure 3.9. Pour des composants idéaux, C_1 échantillonne la tension d'entrée V_{in} quand Φ_1 est haut ("on"), puis se charge proportionnellement à l'entrée avant que cette phase soit terminée, $q_1(n) = C_1 v_{in}(n)$. Pendant la deuxième phase ($\Phi_2 = 1$), la capacité C_1 se décharge complètement sur C_2 à travers la masse de l'AOP, donc :

$$q_2(n + \frac{1}{2}) = q_2(n) + C_1 V_{in}(n)$$

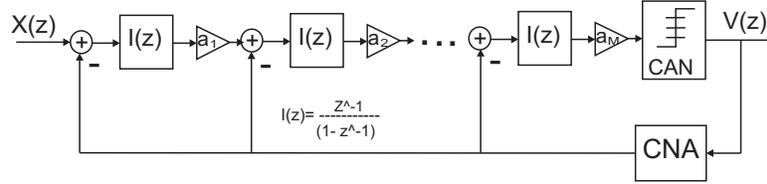


FIG. 3.8 – Structure d'un modulateur $\Sigma\Delta$ passe-bas

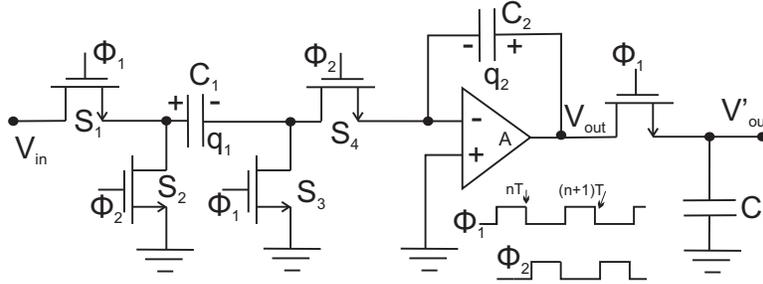


FIG. 3.9 – Un intégrateur SC, qui est moins sensible aux imperfections des éléments

$$V_{out}(n + 1) = V_{out}(n) + \frac{C_1}{C_2} V_{in}(n) \quad (3.22)$$

La fonction de transfert de l'intégrateur, décrit dans la figure 3.9, est donnée par l'équation suivante :

$$H_{int}(z) = \frac{C_1}{C_2} \frac{z^{-1}}{(1 - z^{-1})} \quad (3.23)$$

Quand $C_1 = C_2$, ce filtre réalise la fonction de transfert d'un intégrateur idéal avec un gain unité.

En tenant compte des différentes capacités parasites, la fonction $H_{int}(z)$ peut être présentée comme dans l'équation 3.24 :

$$\begin{aligned} H_{int}(z) &= \frac{C_1}{C_2} \frac{(1 - \alpha)z^{-1}}{(1 - \beta z^{-1})} \\ \alpha &= \frac{1}{A} \frac{C_1 + C_2 + C_{in}}{C_2 + \frac{C_1 + C_2 + C_{in}}{A}} = \frac{1}{A} \left(1 + \frac{C_1}{C_2} + \frac{C_{in}}{C_2} \right), \text{ et} \\ \beta &= \frac{-(1 + \frac{C_2 + C_{in}}{AC_2})}{1 + \frac{C_1 + C_2 + C_{in}}{AC_2}} \end{aligned} \quad (3.24)$$

où C_{in} représente non seulement les capacités parasites à l'entrée de l'amplificateur opérationnel mais aussi différentes capacités parasites dues aux commutateurs connectés à l'entrée négative de l'amplificateur. β est un paramètre très critique car il détermine la stabilité du filtre [95, 96]. En effet, si il est inférieur à 1, alors le filtre est stable. Idéalement, β vaut 1. Pourtant, en présence de différentes capacités parasites en particulier la capacité d'entrée de l'amplificateur, ce paramètre peut être supérieur à 1. Dans ce cas, le pôle du filtre se trouvera en dehors du cercle d'unité, $|z| = 1$, ce qui signifie que le filtre est intrinsèquement instable. Dépendant de la sévérité de l'erreur sur le pôle du filtre et de l'amplitude de l'état initial du modulateur, l'instabilité peut être introduite dans le modulateur même si l'entrée reste nulle. Cette instabilité est caractérisée par le fait que les signaux

internes continuent à croître sans être bornés ou divergents même si l'entrée du modulateur est nulle. Mais il est primordial de remarquer ici que l'instabilité du filtre n'entraîne pas directement celle du modulateur. En effet, le filtre peut être instable alors que le modulateur reste stable si le pôle du filtre est suffisamment proche de 1. Il est donc intéressant de déterminer l'intervalle de valeurs du pôle du filtre pour que le modulateur reste toujours stable [29]. Ces valeurs dépendant du modulateur utilisé peuvent être déterminées seulement par simulations. A partir de l'équation 3.24, on trouve la condition :

$$\beta < \left(\frac{C_3 + C_2}{C_2} - \frac{C_2 + C_{in}}{AC_2} \right) \quad (3.25)$$

où A est le gain statique de l'amplificateur. Cette relation montre que l'erreur sur le pôle du filtre qui est idéalement égal à 1 lorsque $C_{in} \ll C_1 + C_2$ ($\beta = 1$), est inférieure à l'erreur maximale du rapport capacitif.

Le bruit thermique d'un intégrateur SC est induit par la résistance des commutateurs (R_{on} comme dans le cas de l'échantillonneur-bloqueur). La figure 3.10-a montre le modèle d'un intégrateur SC dans la phase d'acquisition ($\phi_1 = 1$), où la résistance de commutateur combinée des transistors S_1 et S_3 , induit un bruit thermique $v_n^2(f) = 8R_{on}k_B T$ comme simplifié à la figure 3.10-b. Alors, la

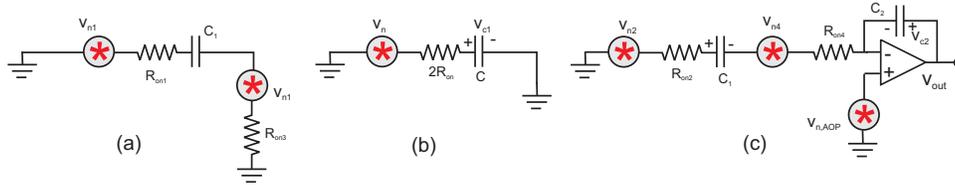


FIG. 3.10 – *Modèle d'un intégrateur SC avec sources de bruit thermique a) quand $\phi_1 = 1$ b) quand $\phi_2 = 1$, c) l'équivalent*

densité spectrale du bruit à la sortie du filtre ($2R_{on} - C_1$) est donnée par :

$$\overline{v_{c_1}^2} = \frac{k_B T}{C_1} \quad (3.26)$$

Par conséquent, la charge de bruit thermique échantillonné $q_{1,rms} = C_1 v_{1,rms} = k_B T C_1$ sera stockée sur la capacité C_1 à la fin de phase ϕ_1 . Dans la deuxième phase, quand $\phi_2 = 1$, deux autres sources de bruits doivent être considérées : le bruit de l'AOP et le bruit des commutateurs $R_{on2} + R_{on4}$ comme cela est décrit à la figure 3.10-c, où l'équivalent du bruit de l'AOP ($v_{n,in} = \sqrt{\frac{16k_B T}{3g_{m1}} \left(1 + \frac{g_{m3}}{g_{m1}}\right)}$) est ramené à l'entrée. A la phase suivante, $\phi_1 = 1$ le bruit sortant de l'intégrateur sera transféré à la capacité C_L par une autre commutation entre l'AOP et la sortie. En tenant compte de tout ces bruits thermiques échantillonnés, la densité spectrale équivalente (vis à vis l'entrée) peut alors s'exprimer par [1, 3, 93, 95] :

$$\overline{v_{n-in,total}^2} \simeq \frac{k_B T}{C_2} \left[1 + \frac{4}{3} \cdot \frac{1 + n_t}{\frac{C_L}{C_2} \left(1 + \frac{C_1}{C_2} + \frac{C_{in}}{C_2}\right) + \frac{C_1}{C_2} + \frac{C_{in}}{C_2}} \right] \quad (3.27)$$

où n_t est le facteur de contribution de bruit d'AOP si il y a plus d'un étage simple. Dans le cas montré à la figure 3.10, n_t vaut 1. On constate que le comportement de l'intégrateur à capacités commutées

est essentiellement dominé par le facteur $\frac{k_B T}{C_2}$ qui dépend fortement de la dimension de la capacité. Cela met en évidence la faible performance d'un intégrateur *SC*. Semblablement, cette dépendance est aussi forte dans un intégrateur de type *SI* du fait d'une relation similaire, $\frac{2k_B T}{3C}(1+n_t)$, [3,97]. Cependant, l'effet de bruit thermique peut être beaucoup plus faible si on peut utiliser une structure à temps continu (passe-bande ou passe-bas) du type présenté à la figure 2.35 [24], où le bruit est dominé par celui de l'AOP [$\frac{4k_B T(1+n_t)}{g_m}$]. Autrement dit, avoir un faible bruit thermique est un des avantages du modulateur à temps continu.

D'autres non-idéalités des intégrateurs, qui ne sont pas traitées dans ce travail mais peuvent limiter les performances du modulateur sont :

- le défaut d'appariement des capacités,
- la non-linéarité des capacités,
- la non-linéarité des résistances des commutateurs,
- La vitesse de balayage limitée de l'AOP,
- La saturation de sortie de l'AOP,
- Le gain fini de l'AOP.

En général, ces erreurs sont assez petites et n'influencent pas trop les performances globales, mais, on doit quand même les examiner [3–5,41,91]. Probablement, le pire cas est celui du gain fini d'AOP qui limite le gain du filtre (équation 3.24) ainsi que celle de la vitesse de balayage limitée.

3.4 Non-linéarité du quantificateur (CAN)

La figure 3.11-a montre le schéma global d'un convertisseur rapide (CAN) à faible nombre de bits ($B=3$ bit). Pour un CAN à 2^B niveaux, un diviseur potentiométrique ("ladder") composé de $2^B + 1$ résistances (ou capacités) détermine les différents niveaux de comparaison auxquels le signal d'entrée doit être comparé en traversant 2^B comparateurs. Dans une telle structure, la précision des tensions de référence dépend de l'erreur sur les valeurs de résistance qui constituent le diviseur potentiométrique. Cette erreur peut être de nature déterministe ou de nature aléatoire. Dans le premier cas, des dispositions particulières dans le dessin des masques peuvent être efficaces pour en réduire l'effet [4,91]. En ce qui concerne les sources aléatoires d'erreur, la dispersion est en général réduite par un accroissement de la surface des composants. Dans le cas du diviseur potentiométrique, si chaque résistance R a un écart type σ_R la référence numéro j , $j \in \{1,2,\dots,2^B\}$ est exprimée comme suit [98] :

$$V_j = \frac{j}{2^B} V_{p-p}, \quad \sigma_{V_j} = \sqrt{\frac{(\frac{j}{2^B})(1 - \frac{j}{2^B})}{2^B}} (\frac{j}{2^B}) V_{p-p} \quad (3.28)$$

où V_{p-p} est la différence de tension aux bornes de l'échelle qui indique aussi le plage d'entrée du CAN. Cette erreur est maximale pour le potentiel central ($j = 2^{(B-1)} - 1$) avec un écart type :

$$\sigma_{V_j, max} = \frac{1}{2\sqrt{2^B}} (\frac{\sigma_R}{R}) V_{p-p} \quad (3.29)$$

Heureusement, dans un modulateur de type $\Sigma\Delta$ qui utilise un CAN à faible nombre de bits, une linéarité modérée, meilleure que $\frac{\sigma_R}{R} < \frac{1}{\sqrt{2B}}$, qui est tout à fait possible avec les technologies CMOS actuelles, est suffisante.

L'élément clé du CAN est le comparateur qui a besoin à la fois d'un fort gain et d'une rapidité suffisante. Un quantificateur compare la tension d'entrée avec une tension de seuil, ou une tension de référence. Le comparateur de tension doit amplifier une petite différence entre ses entrées analogiques, qui est souvent de quelques milli-volts, jusqu'à un niveau logique qui est de l'ordre quelques volts ($> \frac{2V_{dd}}{3}$). Dans un modulateur à hautes performances, les trois grandeurs suivantes doivent être examinées pour chaque comparateur⁵ :

- la tension de décalage "l'offset",
- l'hystérésis
- le temps d'établissement.

Dans le contexte des modulateurs sigma delta, toutes les erreurs additionnelles qui se trouvent en fin de boucle, par exemple celles du quantificateur, sont divisées par le gain des étages précédents. L'effet de l'offset du comparateur est une erreur additionnelle et peut être considéré de la même façon que l'erreur d'appariement des potentiomètres que l'équation 3.29 traduit comme suit :

$$\sigma_{V_j, max} = \frac{1}{2\sqrt{2B}} \left(\frac{\sigma_R}{R} \right) V_{p-p} + \sigma_{V_{offset}} \quad (3.30)$$

Notons que la tension d'offset de chaque comparateur est ajoutée directement à chaque entrée, c.-à-d. elle n'est pas divisée par la racine carrée du nombre de comparateur. $\sigma_{V_{offset}}$ est en général inférieur à quelques milli-volt, ce qui ne pose pas de problème quand on applique un bon *OSR* et une bonne NTF [2, 4]. Une autre remarque est que si on utilise un seul diviseur potentiométrique mais avec plusieurs quantificateurs, par exemple avec une architecture *ping-pong* ou entrelacé dans le temps [51], l'effet de l'offset reste inchangé. Cette remarque intéressante sera utilisée dans notre proposition qui sera expliquée au chapitre 6.

Une autre caractéristique importante à étudier est l'hystérésis du comparateur (voir la figure 3.11). L'hystérésis du comparateur rend chaque décision dépendante des décisions précédentes, mettant une mémoire dans les comparaisons. Cette mémoire peut créer des pôles non désirés dans le système, qui peuvent causer des erreurs dans les fonctions des transfert du signal (STF) et du bruit (NTF). Afin d'empêcher ce phénomène et d'améliorer la vitesse du comparateur, on utilise souvent une technique de remise à zéro avant chaque comparaison (voir l'exemple à la figure 3.12). Dans la phase de remise à zéro, la sortie de la précédente comparaison doit être maintenue par une bascule RS jusqu'à une nouvelle décision valide.

En revanche le temps d'établissement joue un rôle beaucoup plus sérieux sur la précision que les problèmes étudiés précédemment. Il limite en quelques sorte σ la fréquence maximale de comparaison, surtout dans le cas des modulateurs à temps continu. Il est dû essentiellement à la bande passante

5. Le bruit thermique peut aussi contribuer à la performance d'un comparateur si son entrée est limitée dans un très petit intervalle, mais, ce n'est pas toujours le cas. De plus, la gigue d'horloge peut aussi ajouter des erreur et dégrader la performance d'un quantificateur. Cette non idéalité sera évoquée dans la prochaine section

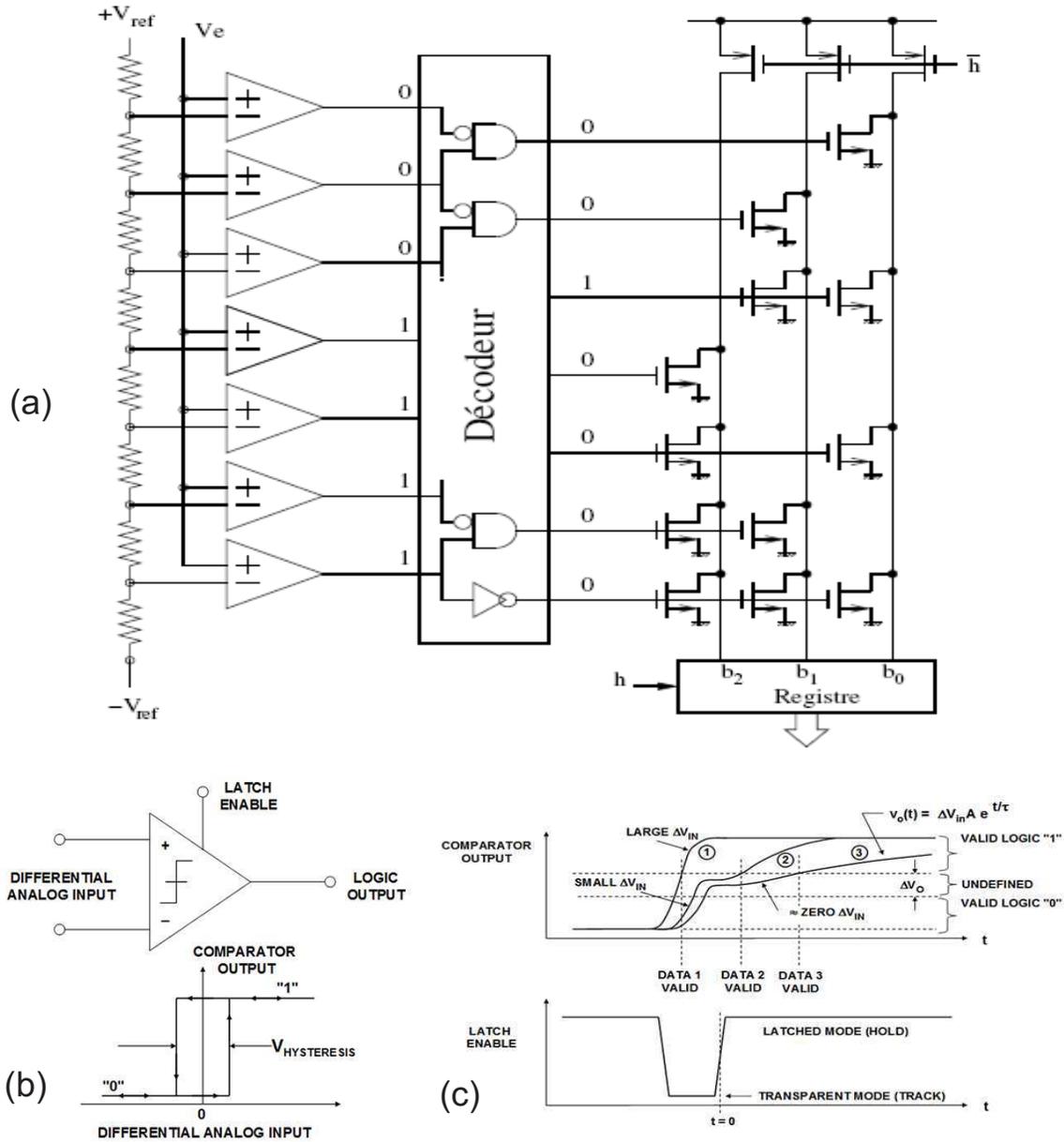


FIG. 3.11 – Quantificateur : a) le schéma bloc global, b) le schéma du comparateur c) la réponse d'un comparateur réel aux trois différents niveaux d'entrées

finie des amplificateurs et à une caractéristique propre du comparateur : la métastabilité. Le temps nécessaire pour atteindre un niveau logique bien défini en sortie du comparateur dépend fortement de la tension d'entrée, et peut même être supérieur à une période d'horloge. Ceci affecte localement le code thermométrique et peut conduire à des erreurs grossières en sortie du convertisseur, comme l'illustre la figure 3.11-c.

L'idée d'utiliser une structure d'amplificateur simple au d'amplificateur cascode comme un comparateur donne de mauvais résultats avec un temps d'établissement très lent [1, 3]. On utilise souvent une structure dégénérée avec une rétroaction positive dont un exemple a déjà été décrit sur la figure 3.12. Afin de mieux expliquer l'effet du niveau d'entrée dans un comparateur de type dynamique

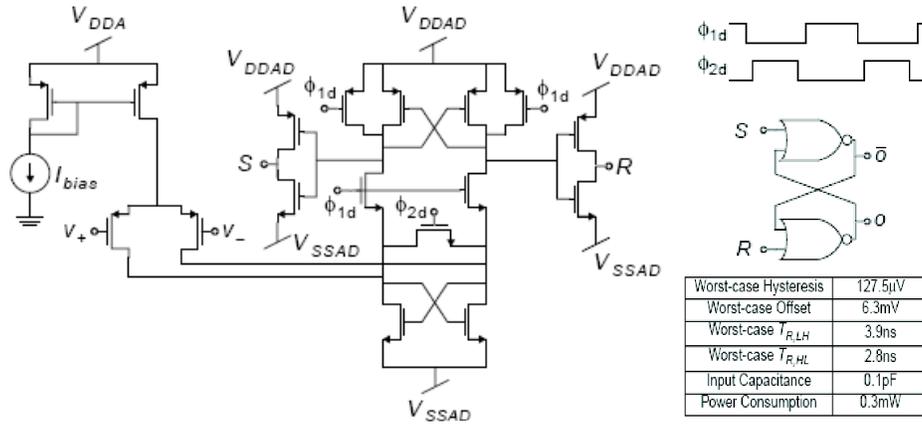


FIG. 3.12 – Un comparateur CMOS dynamique, il marque la sortie au front montant d'horloge

("dynamic latched-comparateur"), analysons d'abord un schéma symbolique d'une combinaison des deux amplificateurs dans une structure de rétroaction positive montrée sur la figure 3.13, où on sup-

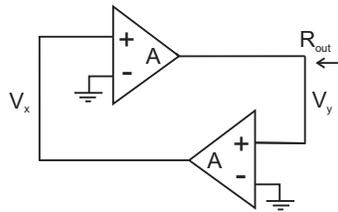


FIG. 3.13 – Une combinaison des deux amplificateurs d'une structure de rétroaction positive dans un comparateur

pose que chacun des amplificateurs possède un pôle dominant, une transconductance directe g_m , une résistance de sortie R_{out} , et une capacité de charge C_L . Par le modèle linéaire, on obtient alors [3] :

$$\begin{aligned} g_m V_x(t) + \frac{V_y(t)}{R_{out}} &= -C_L \cdot \frac{dV_y(t)}{dt} \\ g_m V_y(t) + \frac{V_x(t)}{R_{out}} &= -C_L \cdot \frac{dV_x(t)}{dt} \end{aligned} \quad (3.31)$$

Après simplification, nous avons les relations suivantes, où " $A = g_m \cdot R_{out}$ " est le gain de l'AOP, et $\tau = C_L \cdot R_{out} = \frac{A}{\omega_u} = \frac{1}{\omega_{-3dB}}$ est la constante de temps de l'AOP quand ω_u est la fréquence de passage au gain unité et ω_{-3dB} est la largeur de la bande de l'AOP.

$$\begin{aligned} AV_x(t) + V_y(t) &= -\tau \cdot \frac{dV_y(t)}{dt} \\ AV_y(t) + V_x(t) &= -\tau \cdot \frac{dV_x(t)}{dt} \\ \Delta V = V_x(t) - V_y(t) &= \frac{-\tau}{A-1} \cdot \frac{d}{dt} \Delta V \simeq \frac{\tau}{A} \cdot \frac{d}{dt} \Delta V = \omega_u \cdot \frac{d}{dt} \Delta V \end{aligned} \quad (3.32)$$

qui peut avoir la solution suivante :

$$\Delta V = \Delta V_0 \cdot e^{\omega_u \cdot t} = \Delta V_0 \cdot e^{\frac{t}{\tau_l}} \quad , \quad \tau_l = \frac{1}{\omega_u} = \frac{\tau}{A} = \frac{C_L}{g_m} \quad (3.33)$$

où ΔV_0 est la tension initiale, par exemple, la valeur de l'entrée. On remarque que la constante de temps d'une combinaison de rétroaction positive est celle d'un AOP, mais, divisée par le gain de l'AOP. De plus en diminuant la capacité de charge ou par l'augmentation de g_m , on peut construire un comparateur plus rapide. Afin de générer une différence de tension à la sortie du comparateur $\Delta V_{logique}$, le temps nécessaire est exprimé par :

$$T_{latch} = \tau_l \cdot \text{Ln} \left[V_0 \cdot \frac{\Delta_{logique}}{\Delta V_0} \right] \quad (3.34)$$

Si à la fin de la période, la sortie du comparateur n'atteint pas d'une différence détectable par la logique suivante, c.-à-d. $T_{com} < T_{latch}$, une erreur de métastabilité se produit. Autrement dit, la sortie du comparateur aura une valeur entre 0 et 1, ce qui n'est pas logiquement défini. La valeur de l'entrée différentielle minimum qui est nécessaire pour avoir une comparaison correcte à chaque période est obtenue par la relation suivante :

$$v_{min} = \Delta V_{logique} \cdot e^{-\frac{T_{com}}{\tau_l}} \quad (3.35)$$

Cette tension minimale correspond à la résolution du comparateur soit la valeur de la tension de LSB. L'erreur de métastabilité du comparateur peut être ensuite définie comme :

$$P_e = \frac{v_{min}}{V_{LSB}} = \frac{\Delta V_{logique}}{V_{LSB}} \cdot e^{-\frac{T_{com}}{\tau_l}} = 2^B \cdot e^{-\frac{T_{com}}{\tau_l}} \quad (3.36)$$

Si le comparateur utilise une horloge avec deux phases complémentaires le temps de comparaison peut être considéré comme égal à la moitié de la période d'horloge. Par conséquent, la densité d'erreur de métastabilité s'exprime comme suit :

$$P_e = 2^B \cdot e^{-\pi \cdot \frac{BW}{f_e}} \quad (3.37)$$

où, $f_e = \frac{1}{T_e}$ est la fréquence d'échantillonnage $BW = f_{-3dB,comparateur} = \frac{g_m}{2 \cdot \pi \cdot C_L}$ est la largeur de bande du comparateur en phase de comparaison. En plus, si le comparateur suit un étage pré-amplificateur de gain A_p la valeur de v_{min} à l'équation 3.35 ainsi que la probabilité d'erreur à l'équation 3.37 sont divisées par le gain A_p .

Pour conclure, l'erreur de métastabilité diminue d'une façon linéaire en augmentant le gain du comparateur, alors qu'elle est atténuée exceptionnellement quand on augmente le temps de comparaison. Cette conclusion peut être intéressante quand un modulateur à temps continu rencontre la limite de fréquence d'échantillonnage imposée par le temps de comparaison T_{com} [24, 51]. Par exemple, l'utilisation d'un temps de comparaison équivalant à $\frac{3T_e}{4}$ au lieu de $\frac{T_e}{2}$, ou l'utilisation de deux séries de comparateurs entrelacés peuvent augmenter la limite de la fréquence d'échantillonnage jusqu'à la faire doubler.

En tous cas, le temps d'établissement du comparateur doit être considéré comme une partie du retard de boucle du modulateur à l'étape de conception au niveau système même si il ne limite pas de la fréquence d'échantillonnage. Ce retard, qui représente une bonne partie du retard de la boucle, avec l'ensemble des retards des autres éléments peut aussi limiter la performance ainsi que la fréquence de fonctionnement du système global [28, 33, 35, 36, 99].

3.5 Convertisseur numérique-analogique interne (CNA)

Un convertisseur numérique-analogique (CNA) est nécessaire pour reboucler la sortie de quantificateur à l'entrée du filtre du modulateur $\Sigma\Delta$. Le signal généré par le CNA interne (nommé $Do(n)$, "DAC-out") doit opposer au signal d'entrée. En outre, la sortie du CNA interne doit être aussi propre que l'entrée. Il est indispensable que l'ensemble des erreurs dues au CNA soit plus petite que la variation minimale d'entrée que l'on souhaite détecter.

Un convertisseur numérique-analogique réel transforme une entrée binaire B-bits $v(n)$ en une sortie analogique qui est mise à l'échelle par un gain Δ , accompagnée d'une erreur $e(n)$ comme illustré sur le schéma 3.14.

$$Do(n) = \Delta.v(n) + e(n) \quad (3.38)$$

Pour des raisons de simplicité d'analyse, le coefficient Δ peut être normalisé à l'unité, $\Delta = 1$. Dans la

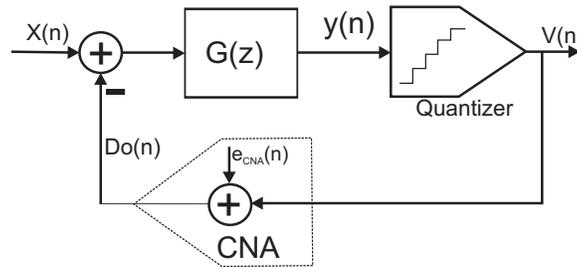


FIG. 3.14 – Modèle des erreurs du CNA interne dans un modulateur multibit

pratique, ce coefficient est dénormalisé par rapport aux valeurs nominales des entrées et des sorties du modulateur. La sortie du modulateur s'écrit alors⁶ :

$$\begin{aligned} V(z) &= [X(z) + E(z)] \frac{G(z)}{1 + G(z)} + Q(z) \frac{1}{1 + G(z)} \\ &= [X(z) + E(z)].STF(z) + Q(z).NTF(z) \end{aligned} \quad (3.39)$$

On voit bien que l'erreur due au CNA se retrouve en sortie via la fonction de transfert $STF(z)$, de la même façon que l'entrée principale de modulateur. Cela signifie que cette erreur doit être toujours inférieure à la valeur minimale d'entrée détectable :

$$\frac{e_{rms}}{Do_{max}} < \frac{1}{\sqrt{2}.2^{(ENOB+1)}} \quad (3.40)$$

Par exemple, pour une résolution globale de 18-bits, l'erreur maximale du CNA doit être au moins de 2^{-19} , soit 0.0002%. Cette précision est exigée, mais inatteignable quelle que soit la technologie. Cependant, les concepteurs s'en approchent par différentes méthodes de correction dont certaines sont étudiées dans cette section.

6. L'erreur " $E(z)$ " est un signal de nature stochastique pour lequel la signe peut être choisi librement "plus" ou "moins" dans les équations.

3.5.1 Architecture générale du CNA interne

La majorité des CNA situés dans le rétroaction d'un modulateur $\Sigma\Delta$, emploie des cellules unitaires appariées ("matched units"). Le CNA peut avoir différentes structures, par exemple, thermométrique, segmentée ou binaire. Parmi ces architectures, nous nous intéressons plutôt au cas thermométrique car elle est le meilleur choix pour le CNA faible bits mais haute résolution, dans un modulateur $\Sigma\Delta$. Un CNA thermométrique peut être réalisé au moyen de plusieurs types de circuits électroniques tels que des sources de courant pilotées "current-steering-DAC", des capacités pilotées "capacitor-steering-DAC", et des résistances pilotées, "resistor-steering-DAC". Un CNA en courant est souvent préférable dans le cas des modulateurs à temps continu pour sa plus grande rapidité, cependant, la structure avec capacités pilotées est plus adaptée aux convertisseurs à temps discret de type "SC". Dans la pratique, des composantes parfaitement identiques ou appariées ne sont pas possibles à réaliser en raison des limites imposées par le processus de fabrication, le gradient de température des dispositifs et le vieillissement des composants. La différence entre la valeur idéale et la valeur réelle des composants est souvent appelée défaut d'appariement (*Mismatch*). Les erreurs de non-idéalité du CNA peuvent être considérées comme une erreur additionnelle stochastique, comme le suppose l'équation 3.38.

La figure 3.15 montre un exemple de fonction de transfert d'un CNA réel et idéal dans laquelle les

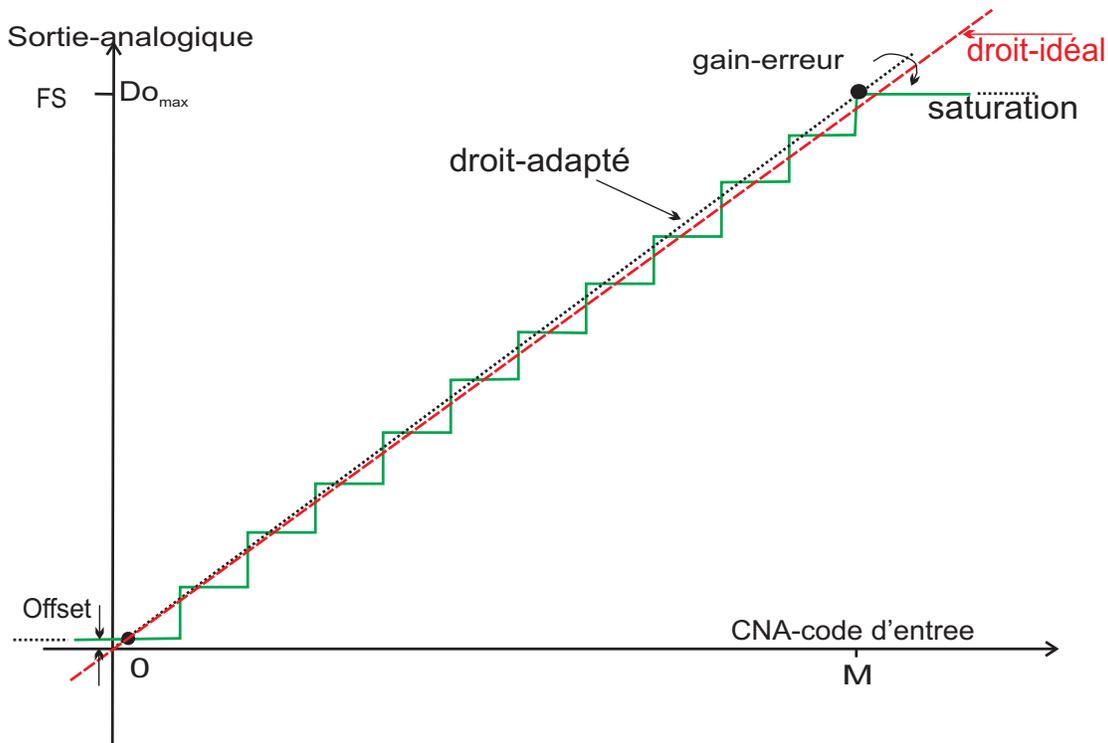


FIG. 3.15 – Exemple de fonction de transfert d'un convertisseur numérique-analogique multibit

deux extrêmes de la courbe sont souvent superposés. Cela n'est pas vrai, mais ne pose pas beaucoup de problème, car, il ne se produit qu'une petite erreur linéaire de gain (dérivation de Δ à l'équation

3.38). Cette erreur peut être négligée car elle ne produit pas de distorsion harmonique si le modulateur ne sature pas. En revanche, l'erreur de non-linéarité liée au terme $e(n)$ a une grande influence sur la résolution finale du modulateur puisqu'elle produit des distorsion harmoniques, qui selon l'équation 3.39, peuvent dégrader la performance globale du système.

Il existe des CNA qui réduisent des erreurs de défaut d'appariement des composantes en employant des processus spéciaux en phase de fabrication, par exemple, un ajustement physique par rayon laser ("*laser-trimming*"). Cela coûte cher, cependant, et on ne peut pas complètement éliminer tous les types d'erreur induites par le défaut d'appariement, comme par exemple celles qui viennent après l'étape de la fabrication de la puce. D'autres méthodes peuvent alors être utilisées, telles que la correction numérique [77], l'étalonnage ("*calibration*") [100] et le brassage de sources par une méthode d'appariement dynamique des composants (DEM). Pour cette thèse, nous nous limitons au dernier cas pour lequel différents algorithmes vont être proposés (voir les deux chapitres suivants). Les techniques de DEM réarrangent dynamiquement les interconnexions des composants du CNA thermométrique de sorte que la moyenne d'erreur des composants soit minimale dans la bande du signal. Par conséquent, l'erreur due aux défauts d'appariement peut être réduite ou décalée le long de l'axe fréquentiel, et ainsi, le SFDR, le SNDR et le ENOB d'un CNA peuvent être améliorés. Dans cette section nous allons donner certains éléments nécessaires pour l'analyse et le développement de l'idée de DEM, qui sera présenté dans les chapitres suivants.

La structure d'un CNA thermométrique simple est représenté sur la figure 3.16. où une entrée B-bits

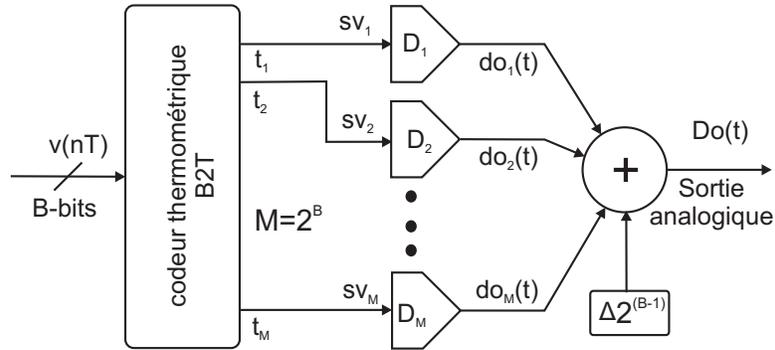


FIG. 3.16 – Le schéma simple d'un CNA thermométrique sans DEM

normalisée (c.-à-d. $0 \leq v(nT) \leq M$, $M = 2^B$) est d'abord transformée en code thermoélectrique $t(n)$. En l'absence d'un bloc intermédiaire de DEM, le code thermométrique est directement lié aux cellules du CNA, $sv(n) = t(n)$. Chacune des cellules D_i , a une sortie analogique donnée idéalement par l'équation 3.41. Afin de compléter cette structure, un terme égal à $\Delta \cdot 2^{(B-1)}$ est ajouté, qui ramène la sortie d'échelle normalisée à l'échelle pratique. Pour des raisons de simplicité d'analyse, ce terme de normalisation n'est pas considéré dans les équations ci-dessous. De plus, le gain Δ est supposé unitaire.

$$do_i(t) = \begin{cases} \Delta & \text{si } sv_i(nT) = 1 \\ 0 & \text{si } sv_i(nT) = 0 \end{cases} \quad (3.41)$$

Le traitement numérique du premier bloc, décrit à la figure 3.16, doit sauvegarder la valeur du signal d'entrée du CNA à chaque instant, comme les équivalences suivantes l'explicitent :

$$v(n) = t(n) \ t^T(n) = \sum_{i=1}^M t_i(n) = \sum_{i=1}^M sv_i(n) \quad (3.42)$$

où, T est l'opérateur de transposition de matrice, $t(n)$ est un vecteur qui se compose de M éléments 1 bit : $[t_1(n) \ t_2(n) \ \dots \ t_M]$, $i \in \{1, 2, \dots, M\}$.

3.5.2 Erreurs statiques du CNA

Le modèle statique d'une cellule (normalisée avec $\Delta = 1$) de CNA réel peut être dérivé de l'équation 3.41 comme suit :

$$do_i(t) = \begin{cases} 1 + eh_i & \text{si } sv_i(nT) = 1 \\ el_i & \text{si } sv_i(nT) = 0 \end{cases} \quad (3.43)$$

où, eh_i est l'erreur additionnelle à la sortie de la cellule numéro i quant elle reçoit une entrée $sv_i(n)$ au niveau haut, et el_i est l'erreur pour une sortie au niveau bas. Cette définition est représentée sur la figure 3.17. On peut expliquer ce modèle par la relation 3.44 :

$$do_i(t) = sv_i(n)[1 + eh_i - el_i] + el_i = sv_i(n)[1 + \alpha_i] + \epsilon_i \quad (3.44)$$

où, les nouveaux paramètres α_i et ϵ_i sont des erreurs de type statique nommées erreur de gain et erreur d'offset de la cellule i et définies comme : $\alpha_i = (eh_i - el_i)$ et $\epsilon_i = el_i$.

La forme du transitoire de sortie d'une cellule du CNA peut varier par rapport à la commutation

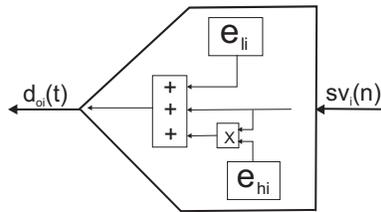


FIG. 3.17 – Modèle statique d'une cellule du CNA réel située au chemin de retour

appliquée ($sv_i(n)$). Les erreurs dynamiques qui concernent la forme transitoire de la sortie du CNA sont analysées à la section suivante.

En prenant en compte la non idéalité donnée par l'équation 3.44, la sortie normalisée d'un CNA s'exprime par :

$$Do(t) = \sum_{i=1}^M do_i(t) = \sum_{i=1}^M sv_i(n)[1 + \alpha_i] + \sum_{i=1}^M \epsilon_i \quad (3.45)$$

Cela peut se simplifier en utilisant l'équation 3.42 :

$$Do(t) = v(n) + e(n) + \epsilon, \quad \text{pour : } nT \leq t < nT$$

$$\text{quand : } e(n) = \sum_{i=1}^M \alpha_i sv_i(n), \text{ et } \epsilon = \sum_{i=1}^M \epsilon_i \quad (3.46)$$

où $e(n)$ est appelé l'erreur de non linéarité du CNA, qui dépend des caractéristiques des entrées $sv(n)$, donc, qui peut potentiellement produire des distorsions harmoniques. Le terme ϵ est appelée l'erreur d'offset, qui est indépendante de l'entrée.

Le terme $e(n)$ est l'erreur la plus importante du CNA multibit que nous allons analyser par la suite. En revanche, l'erreur d'offset a peu d'influence sur la performance d'un CNA utilisé dans un modulateur $\Sigma\Delta$, surtout dans le cas passe-bande. De plus, la réalisation du modulateur en mode différentiel peut diminuer l'effet des erreurs d'offset. La seule inquiétude majeure due à l'erreur d'offset apparaît lorsqu'on utilise deux systèmes identiques en alternance ou en parallèle. Cette architecture peut être employée aux certains cas de modulateurs à temps continu [23, 101].

Un autre résultat provenant de l'équation 3.46 est que le terme $e(n)$ dépend plus du décalage de la différence entre la sortie haut et la sortie bas ($\alpha_i = e_{h_i} - e_{l_i}$) de chaque cellule que de leurs valeurs absolues. Ce raisonnement peut être intéressant quant on veut optimiser le comportement dynamique des circuit [31].

La cause d'erreur de α_i est détaillée dans la référence [91]. Pour l'instant, nous allons calculer l'effet des erreurs statiques dues au CNA multibit. Notons que, dans le cas monobit, le terme $e(n)$ devient une simple erreur de gain, (c.-à-d. $do(t) = sv(n).(\Delta + \alpha) = sv(n).\Delta'$). Il n'y a donc pas d'effet de non linéarité, ce qui constitue le grand avantage d'un système monobit.

Critères du CNA

Après avoir présenté le modèle du CNA, nous rappelons certains critères qui peuvent être utilisés pour l'analyse d'un CNA hors du modulateur.

Droite idéale:

La réponse analogique d'un CNA idéal est une fonction identité ayant pour entrée un signal numérique. Dans le cas normalisé, $\Delta = 1$, cette droite s'étend de 0 à $M = 2^B$ comme décrit à la figure 3.15. Si "x" est la valeur numérique d'entrée, l'expression de la fonction de transfert s'exprime par l'équation $Do(x) = x$, pour $0 \leq x \leq M$.

Droite adaptée:

La caractéristique réelle du CNA est obtenue en additionnant une à une les sorties des cellules qui présentent des erreurs aléatoires, $Do(x) = \sum_{i=1}^M [(1 + eh_i - el_i) sv_i(n)) + el_i]$. La sommation de ces variables aléatoires donne une réponse $Do(x)$ qui s'écarte de la droite idéale. Il y a deux manières d'exprimer cet écart :

- définir la droite des moindres-carrés qui fait apparaître une erreur de gain et une erreur d'offset supplémentaires.
- ou définir une droite reliant l'origine à l'extrémité de la caractéristique⁷ : seule une erreur de non linéarité de gain apparaît. Cette droite adaptée que nous utiliserons par la suite, peut

7. c.-à-d. $[Do_{adapte}(0) = \epsilon = \sum_{i=1}^M el_i \text{ pour } x=0]$, et $[Do_{adapte}(M) = \sum_{i=1}^M (1 + eh_i) \text{ pour } x=M]$

s'exprimer par l'expression suivante :

$$Do_{adapte}(x) = x \frac{\sum_{i=1}^M (1 + eh_i) - \sum_{i=1}^M el_i}{M} + \sum_{i=1}^M el_i = x(1 + \bar{\alpha}) + \epsilon \quad (3.47)$$

où, $\bar{\alpha} = \frac{\sum_{i=1}^M eh_i}{M}$ est le moyenne d'erreur de gain du CNA.

Comme précisé auparavant, l'erreur ϵ peut être négligée⁸ car, en pratique, elle ne produit pas d'harmoniques. Alors, l'équation de la droite adaptée se simplifie comme suit $Do_{adapte}(x) \cong x(1 + \bar{\alpha})$.

Erreur de linéarité intégrale, INL:

La courbe réelle du CNA s'écarte de la droite utilisée dans l'équation Eq.3.47 d'une quantité que l'on appellera "l'Écart". L'écart peut s'exprimer en "LSB" comme suit :

$$Ecart(x) = \frac{Do(x) - Do_{adapte}(x)}{\frac{Do_{max}}{M}} \quad (3.48)$$

Pour faciliter la démonstration, on peut considérer que :

$$\epsilon = \sum_{i=1}^M el_i \cong 0, \quad Do_{max} \simeq M \quad (3.49)$$

On définit alors l'erreur de linéarité intégrale ("*Integral Non Linearity error*", *INL*) du CNA par la valeur maximale de l'écart. Par conséquent :

$$\begin{aligned} INL(x) = Ecart(x) &\simeq \sum_{i=1}^x (1 + \alpha_i) - x(1 + \frac{\sum_{i=1}^M \alpha_i}{M}) \\ &= \sum_{i=1}^x (1 - \frac{x}{M}) \alpha_i - \sum_{j=x+1}^M \frac{x}{M} \alpha_j \end{aligned} \quad (3.50)$$

Si chacune des cellules du CNA présente une erreur aléatoire indépendante et uniformément distribuée⁹, la variance d'erreur $\sigma_{e(x)}^2$, peut avoir une forme simple donnée ci-dessous, si les erreurs α_i et α_j sont supposées indépendantes pour $i \neq j$:

$$\begin{aligned} \sigma_{e(x)}^2 &= var[\sum_{i=1}^x (1 - \frac{x}{M}) \alpha_i - \sum_{j=x+1}^M \frac{x}{M} \alpha_j] = var[\sum_{i=1}^x (1 - \frac{x}{M}) \alpha_i] + var[\sum_{j=x+1}^M \frac{x}{M} \alpha_j] \\ &= var[\sum_{i=1}^x (1 - \frac{x}{M}) \alpha_i] + var[\sum_{j=x+1}^M \frac{x}{M} \alpha_j] = x(1 - \frac{x}{M}) \sigma_{\alpha}^2 \end{aligned} \quad (3.51)$$

De forme parabolique, cette fonction passe par un maximum quelque part entre $x = 0$ et $x = M$. Si on prend $\frac{\partial(\sigma_{e(x)}^2)}{\partial(x)}$ puis on trouve des point 0, on obtient alors le maximum d'écart qui est en milieu d'échelle :

$$INL_{max} = \sigma_{e(x),max}^2 = \frac{M}{4} \cdot \sigma_{\alpha}^2 \quad (3.52)$$

Afin d'avoir un CNA monotone vis-à-vis de la variation d'entrée, l'INL doit être toujours inférieure à 0,5LSB.

8. On peut analyser ses effets séparément de ceux des α_i

9. $var(\alpha_i) = var(\alpha) = \sigma_{\alpha}^2$, la distribution uniforme pour toutes les cellules du CNA

Remarque: L'expression 3.52 est obtenue en utilisant les hypothèses prises dans les équations précédentes, c'est à dire $Do_{max}(x) \cong M$. Connaissant l'écart maximal par rapport à la droite adaptée 3.47, la valeur du LSB change un peu, ainsi que INL_{max} :

$$1 \text{ LSB} = 1 \pm \sigma_\alpha \sqrt{M} \quad , \quad INL_{max} = \frac{M}{2(1 + \frac{\sigma_\alpha}{\sqrt{M}})} \cdot \sigma_\alpha \quad (3.53)$$

Par exemple, l'erreur relative sur l'estimation de l'INL pour un défaut d'appariement à l'ordre de 1% est de $\frac{1}{100\sqrt{M}}$, ce qui est négligeable.

Erreur de linéarité différentielle, DNL:

Il s'agit de l'erreur d'incrément élémentaire (1 LSB) par rapport à l'incrément théorique. La caractéristique de transfert sera monotone si $DNL < 1 \text{ LSB}$; on s'efforcera de la limiter à 0.5 LSB. Dans le cas d'un CNA thermométrique cette condition est déjà remplie par la limite de l'INL dans l'équation 3.52, donc le DNL n'y a pas un rôle important.

3.5.3 Performances d'un modulateur avec CNA multibit réel sans correction

Afin de vérifier la performance globale du modulateur, les critères de INL ou de DNL du CNA interne ne sont pas suffisamment précis et utiles. La raison en est que le modulateur $\Sigma\Delta$ est sensible à l'ensemble des paramètres dont une partie est liée au CNA interne. L'influence de la non-idéalité du CNA est le plus souvent comparée à celle de l'erreur de quantification (Eq.2.27); elle se retrouve dans les expressions de SNDR, SFDR, DR et ENOB.

Afin de calculer l'effet de non idéalité du CNA-rebouclé, nous allons chercher une relation analytique pour la puissance d'erreur dans la bande utile du modulateur. L'erreur peut être réécrite : $e(n) = \sum_{i=1}^M \alpha_i sv_i(n)$, (Eq.3.46). Elle peut s'exprimer par sa variance temporelle :

$$\begin{aligned} P_e &= Var[e(n)] = Var\left[\sum_{i=1}^M e_i(n)\right] \\ &= \sum_{i=1}^M Var[e_i(n)] + \sum_{i \neq j} Covar[e_i(n), e_j(n)] \end{aligned} \quad (3.54)$$

Si nous supposons que la probabilité des signaux d'entrées aux entrées des cellules du CNA ont une distribution uniforme, l'équation ci-dessus peut se simplifier en admettant¹⁰ $P[sv_i(n) = 1] = P[sv_i(n) = 0] = \frac{1}{2}$. On obtient alors :

$$Var[sv_i] = E[sv_i^2] - E^2[sv_i] = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4} \quad (3.55)$$

Durant la période d'observation, le terme α_i est supposé constant, et invariant dans le temps. Par conséquent :

$$Var[e_i] = Var[\alpha_i sv_i] = \sigma_{\alpha_i}^2 Var[sv_i] = \frac{\sigma_{\alpha_i}^2}{4} \quad (3.56)$$

10. Cela n'est effectivement vrai que pour des entrées proches de moitié de la pleine échelle.

Les termes de covariances à l'équation 3.54 dépendent des caractéristique stochastiques des entrées :

$$Covar[e_i, e_j] = \sigma_{\alpha_i} \sigma_{\alpha_j} Covar[sv_i, sv_j] \quad (3.57)$$

Cela ne peut malheureusement pas s'expliquer sous une forme analytique pour un signal arbitraire. Cette dépendance est normalement faible quand les index "i" et "j" sont distants, mais peut être plus fort pour les signaux consécutivement appliqués aux cellules voisines. Par ailleurs, ces dépendances varient dans deux sens opposées, c.-à-d. une partie négative et une partie positive qui limitent ses effets dégradants dans le cas d'absence de DEM. Une méthode de DEM en fait utilise cette dépendance afin de diminuer l'erreur dans la bande utile. Au meilleur cas, une bonne manipulation des interconnexions des cellules peut éliminer la premier partie d'équation 3.54 par celle de deuxième, ce qui semble une intervention assez délicate en pratique qui est à la charge de l'algorithme de DEM.

En tout cas, on essaie de les simplifier afin de trouver une expression approchée. Pour l'instant, on néglige leurs dépendances ($Covar[sv_i(n), sv_j(n)] \simeq 0$), mais on devra examiner les résultat par la suite. Alors, la puissance totale des erreurs statiques du CNA se simplifie comme suit :

$$P_e \simeq \sum_{i=1}^M Var[e_i(n)] = \sum_{i=1}^M \frac{\sigma_{\alpha_i}^2}{4} \quad (3.58)$$

Si les défauts d'appariements sont uniformément distribués, on obtient alors :

$$P_e \simeq \frac{\sigma_{\alpha}^2}{4M} \quad (3.59)$$

ce qui est conforme à l'équation d'INL 3.52. De plus, son écart-type est : $e_{rms} = \frac{\sigma_{\alpha}}{2\sqrt{M}}$. L'autre difficulté à la suite de cette analyse, apparaît lorsque l'on veut calculer la contribution d'erreur du CNA de rebouclage à la sortie du modulateur. Cela provient du fait que l'erreur $e(n)$ n'a pas un spectre défini, car il n'est pas blanc comme le bruit thermique. Sa caractéristique est fortement liée à la fois à la caractéristique stochastique de l'entrée du modulateur et à la caractéristique du modulateur, ainsi qu'aux valeurs absolues de α_i s.

On peut supposer que la STF du modulateur fait passer la quasi totalité de ces erreurs en bande utile comme il le fait pour le signal d'entrée. Par contre, le phénomène de suréchantillonnage et le filtrage suivant le modulateur limitent la puissance correspondant à ces erreurs dans la sortie du filtre. Cependant, dans le pire cas, il est probable que toutes ces erreurs apparaissent en bande utile. Donc, la contribution maximale d'erreur de défaut d'appariement du CNA est la suivante :

$$\sigma_e^2 = \frac{\sigma_{\alpha}^2}{4M} \quad (3.60)$$

En tenant en compte de ces erreurs, le SNDR est calculé à partir des équations 2.27 et 2.41 :

$$SNDR = \frac{P_{entre}}{P_{Qinband} + P_{eCNA}} = \frac{\left(\frac{M}{2\sqrt{2}}\right)^2}{\frac{\pi^{2L}}{12G^2(2L+1)OSR^{(2L+1)}} + \frac{\sigma_{\alpha}^2}{4M}} \quad (3.61)$$

D'autre part, afin d'examiner un CNA composé des M cellules en dehors du modulateur, la résolution propre à ce CNA sans aucune DEM peut s'exprimer par la définition 3.5 :

$$ENOB_{CNA} = \text{Log}_2\left[\frac{\sqrt{M}}{\sqrt{3}\sigma_{\alpha}}\right], \text{ en bit} \quad (3.62)$$

En pratique un autre phénomène dû à cette erreur va dominer. Il s'agit de la modulation de l'erreur de quantification qui s'étend tout au long de l'axe des fréquences. Autrement dit, la puissance de bruit de la non linéarité produite par la modulation entre l'erreur du CNA et l'erreur de la quantification fait augmenter le niveau du bruit dans la bande, même si la contribution des erreurs donnée par l'équation 3.60 n'est pas blanc. Dans la plupart des cas, la propre résolution de CNA (ENOB), peut être encore acceptée. En revanche, la présence de grands tons produit par la modulation avec des signaux d'entrée n'est pas acceptable dans les systèmes de communication, car ces tons limitent fortement le SFDR du système.

Pour donner des chiffres, considérons un CNA 4-bits avec $\sigma_\alpha = 0.01$. La résolution propre du CNA ENOB vaut 8 bits. Le SNDR maximal d'un modulateur, d'ordre 3 avec un $OSR=64$ et $G=0.2$, qui emploie un tel CNA est alors limité à 8 bits, contre une résolution idéale de l'ordre de 20 bits. Cependant, si le même modulateur emploie un quantificateur (et un CNA) d'un bit, la résolution finale est aussi 8 bit de même qu'avec 16 cellules désappariées ayant 1% d'erreur. Cet exemple montre qu'une quantification multibit ne peut pas donner une résolution supérieure à celle d'une quantification monobit si les cellules du CNA sont désappariées même d'un faible pourcentage. De plus, le SFDR d'un tel système multibit peut être plus mauvais que celui d'un modulateur monobit similaire. La figure 3.18 monte le spectre de la sortie d'un modulateur passe-bande multibit pour une entrée sinusoïdale proche de moitié de la pleine échelle ayant des paramètres : $\{L=3, OSR=64, B=4, M=16\}$. Le niveau du bruit de quantification du cas idéal $-125dB$ revient à la hausse vers $-80dB$ pour le cas réel avec une erreur de défaut d'appariement supposé $\sigma_\alpha = 0.01$. La figure suivante (Fig.3.19) montre la

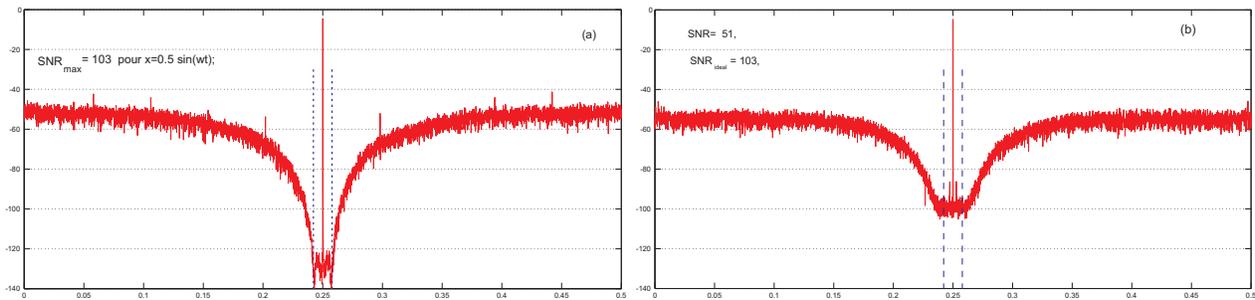


FIG. 3.18 – Le spectre de la sortie du modulateur passe-bande multibit a) idéal, b) avec un CNA réel ayant 1% d'erreur d'appariement sans aucune correction.

performance (SNDR) de ce modulateur vis-à-vis des différentes entrées. Comme on l'a théoriquement prévu, la dégradation due aux défauts d'appariement est fortement liée aux caractéristiques des entrées, mais il manque une relation analytique adaptée. L'analyse simple ci-dessus qui se confirme par des simulations, montre bien qu'un modulateur multibit, ayant quelques dixièmes de pourcentages d'erreur d'appariement entre ses différentes cellules, peut avoir une performance beaucoup plus mauvaise qu'un modulateur monobit. Ainsi, pour bénéficier des avantages de la structure multibit, il est indispensable d'utiliser au moins une méthode de correction, même si cela coûte quelques dizaines de pourcentage de la surface et de la consommation globale du modulateur. Certaines de ces méthodes seront analysées au chapitre suivant, et de nouvelles méthodes seront proposées au chapitre d'après.

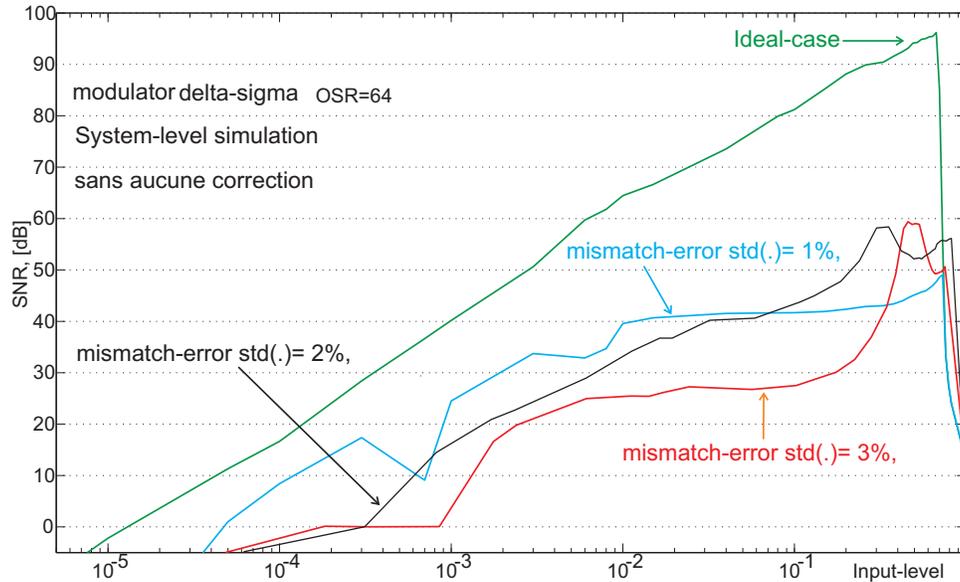


FIG. 3.19 – SNDR du modulateur multibit a) idéal, b) avec un CNA réel ayant 1%, 2% et 3% d'erreur d'appariement.

3.5.4 Erreurs de non linéarité de la tension de sortie du CNA

Un CNA construit avec les condensateurs (C_i , $i \in \{1, \dots, M\}$) peut avoir un autre type d'erreur que celles liées au défaut d'appariement des différentes cellules, lié cette fois à la dérivation de leur valeur absolue vis-à-vis du niveau de tension de sortie ($V_o(t) = V_0 + v_0$) [3,91].

$$C_i(v_o) = C_0 + c_1 v_o(t) + c_2 v_o^2(t) + \dots \quad (3.63)$$

Cela produit une erreur de non linéarité (harmoniques) qui n'est pas négligeable si la variation de tension de sortie est élevée, comme dans le cas d'un convertisseur numérique analogique classique (sans modulation sigma delta) avec une résolution de 10 bit ou plus. Cependant, dans le cas d'un CNA interne qui se compose, par exemple, de 16 cellules, cette non linéarité n'est pas vraiment dégradante.

On peut également constater ce type d'erreur dans le cas d'un CNA composé de sources de courant pilotées, mais, par la variation de courant fuite dans l'impédance de sortie des cellules, si la résistance ou la capacité de sortie de la cellule dépend de la tension. Pour un CNA constitué de sources de courant piloté de faible bit cette erreur est négligeable. Par contre une résistance limitée de la cellule influence la performance de ce type de CNA comme nous allons le montrer par la suite.

3.5.5 Influence de l'impédance finie de sortie du CNA

La figure 3.20 montre un exemple d'une structure simple de CNA thermométrique composée de sources de courant pilotées et sa cellule de base. Dans le cas simple, chacune des cellules peut être représentée par une source de courant I_u et une résistance de sortie $R_u = \frac{1}{G_u}$. Si la résistance de

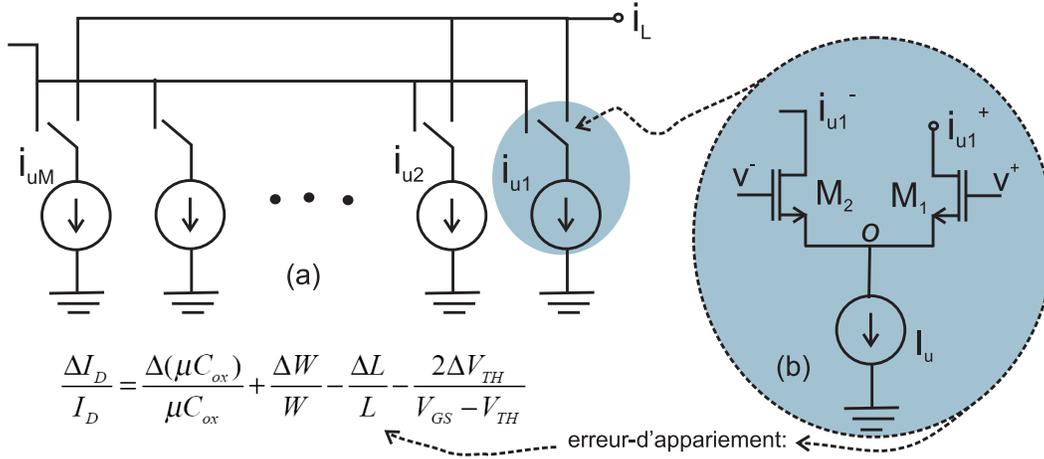


FIG. 3.20 – CNA de type source de courant pilotée a) schéma global de M cellules b) une source de courant et présentation d'erreur d'appariement

charge est supposée $R_L = \frac{1}{G_L}$, le courant qui passe à travers la charge peut s'exprimer par :

$$I_L(x) = \frac{x \cdot I_u}{1 + x \cdot G_u \cdot R_L} \quad (3.64)$$

où, x est l'entrée numérique du CNA. En général, R_u est beaucoup plus élevée que R_L , et $x \cdot G_u \cdot R_L$, donc, $x \cdot G_u \cdot R_L \ll 1$. Par conséquent, la sortie du CNA peut se simplifier :

$$I_L(x) \simeq x \cdot I_u (1 - x \cdot G_u \cdot R_L) = x \cdot I_u - x^2 G_u \cdot R_L \cdot I_u \quad (3.65)$$

Ce qui peut produire une harmonique d'ordre 2. Afin de limiter cet effet, $[x \cdot G_u \cdot R_L]$ doit être assez petit pour l'entrée maximale de x . En supposant une entrée sinusoïdale pleine échelle $x(t) = \frac{M}{2} + \frac{M}{2} \sin(\omega t)$, dans laquelle le premier terme correspond au niveau équilibre d'un modulateur avec une source de courant fixe, la puissance de cette harmonique peut s'estimer comme suit :

$$P_{R_u} \simeq \left[\frac{M^2}{4} G_u \cdot R_L \cdot I_u \right]^2, \quad \sigma_{P_{R_u}} = \frac{M^2}{4} G_u \cdot R_L \cdot I_u \quad (3.66)$$

Quand, I_u est supposé égal à 1 pour le pas LSB, la résolution finale du CNA s'exprime comme suit :

$$ENOB = \text{Log}_2 \left[\frac{M}{2\sqrt{3}\sigma_{R_u}} \right] = \text{Log}_2 \left[\frac{2R_u}{\sqrt{3}M \cdot R_L} \right] \quad (3.67)$$

Par exemple, une résolution globale de 16 bit avec $M=8$, $R_L = 50 \text{ Ohm}$, exige une résistance de cellule supérieur à 22-Mega Ohm. Si cela n'est pas disponible, alors, on utilise plutôt une structure en cascade ou même plus complexe [102–105].

En pratique, trois autres éléments peuvent intervenir, la résistance et la capacité finies de la cellule à l'état "off", le courant qui passe au chemin de " $V_{DC} - R_L - R_{out}$ ", et le capacité parasite à la sortie $C_{eq} = x \cdot C_{on} + (M - x) \cdot C_{off} + C_{par}$. Une cellule de CNA de mode courant a aussi une résistance finie $R_{u,off}$ et une capacité $C_{u,off}$. $R_{u,off}$ est en général négligeable devant celle de l'état "on", mais par contre, la capacité $C_{u,off}$ est de l'ordre de la capacité de l'état "on". Un nouveau modèle de CNA,

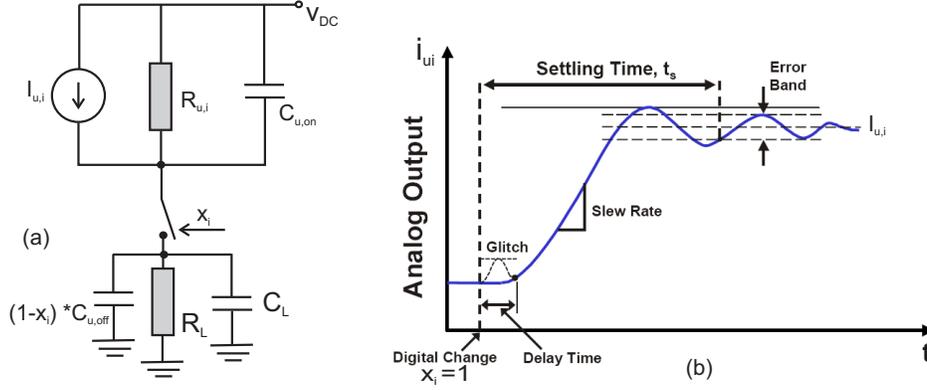


FIG. 3.21 – Modèle d'un CNA de type sources de courant piloté avec ses éléments parasites ainsi que sa réponse transitoire.

qui contient ces éléments parasites est représenté sur la figure 3.21. C_{par} est la capacité parasite de sortie (pour le charge et les connexions). Le courant de fuite de V_{DC} ajoute une partie pour $I_L(x)$ d'environ $\frac{V_{DC}}{R_L + \frac{R_u}{x}}$ ce qui compense un peu l'influence de la résistance finie des cellules. Par conséquent, l'équation 3.64 doit se réécrire :

$$I_L(x) \simeq \frac{x \cdot (I_u + G_u \cdot V_{DC})}{(1 + \rho x)} \quad (3.68)$$

où, on simplifie par : $\rho = \frac{R_L}{R_u} = G_u R_L$. Il est possible de définir une droite linéaire réelle pour la sortie de CNA qui passe par les deux extrémités des valeurs données par l'équation 3.68, c.-à-d. $I_L(0) = 0$ et $I_L(M) = \frac{M \cdot (I_u + G_u \cdot V_{DC})}{(1 + \rho M)}$. Cette ligne appelée droite adaptée $I'_L(x)$, s'exprime par la relation suivante :

$$I'_L(x) = \frac{x \cdot (I_u + G_u \cdot V_{DC})}{(1 + \rho M)} \quad (3.69)$$

Par conséquent, l'erreur de linéarité du CNA peut s'exprimer comme suit :

$$\begin{aligned} \Delta_{I_L}(x) &= I_L(x) - I'_L(x) = x \cdot (I_u + G_u \cdot V_{DC}) \left[\frac{1}{(1 + \rho x)} - \frac{1}{(1 + \rho M)} \right] \\ &= \rho x (M - x) \cdot (I_u + G_u \cdot V_{DC}) \end{aligned} \quad (3.70)$$

Cette erreur a une forme parabolique qui passe par une valeur maximale pour $x = \frac{M}{2}$, tout comme celle de l'erreur d'appariement (en équation 3.53). L'erreur maximale est estimée selon $\Delta_{I_L}(x) = \rho \frac{M^2}{4} \cdot (I_u + G_u \cdot V_{DC})$. On constate que par cette définition, V_{DC} ajoute un gain linéaire, donc, il n'a pas d'influence sur la linéarité du système. Ce raisonnement n'est pas conforme avec les définitions et les discussion présentées par [24, 102] en raison de sa définition qui révèle seulement la non linéarité. Alors, on peut définir un nouveau courant de la cellule unitaire comme $I'_u = I_u + G_u \cdot V_{DC}$. De plus, dans le cas de réalisation en mode différentiel, V_{DC} est idéalement fixé à zéro. La puissance d'erreur statique de la résistance finie des cellules vaut $\sigma_{R_u}^2 = \rho \frac{M^2}{4} \cdot I'_u{}^2$. En prenant une valeur normalisée pour la cellule unitaire avec $I'_u \equiv 1$ la résolution finale s'exprime alors par :

$$ENOB = \text{Log}_2 \left[\frac{2}{\sqrt{3} \rho M} \right] = \text{Log}_2 \left[\frac{2R_u}{\sqrt{3} M \cdot R_L} \right] \quad (3.71)$$

La puissance des harmoniques peut se calculer à partir de l'expression 3.69, qui permet d'estimer le SFDR. Dans le cas non différentiel la seconde harmonique est dominante, alors que dans le cas différentiel la troisième harmonique limite le SFDR, comme cela est donné par les relations suivantes [24, 102] :

$$\begin{aligned} SFDR_{non-diff} &= 20 \text{Log}_{10} \left[1 + \frac{2}{\rho M} (1 + \sqrt{1 + \rho M}) \right] \\ SFDR_{diff} &\cong 40 \text{Log}_{10} \left[2 \left(1 + \frac{2}{\rho M} \right) \right] \end{aligned} \quad (3.72)$$

En ce qui concerne l'effet dynamique de l'impédance de sortie des cellules, l'analyse peut se faire comme dans le cas statique en posant :

$$Z(x) = \frac{R_u}{x} \parallel (xC_{u,on}) \parallel [(M-x)C_{u,off} \parallel C_L] = \frac{R_u}{x} \parallel x(C_{u,on} - C_{u,off}) \parallel (C_L + MC_{u,off}) \quad (3.73)$$

au lieu de $\frac{R_u}{x}$. Cependant, on peut noter que $(C_L + MC_{u,off})$ ne dépend pas de l'entrée x , il ne produit donc pas d'effet non linéaire; il peut être considéré une partie du filtre de système. Le terme imaginaire $x(C_{u,on} - C_{u,off})$, qui dépend de l'entrée, peut être aussi moins gênant parce que la différence de la capacité de cellules entre deux état "on" et "off" est en général plus petite que leurs valeurs absolues, surtout quand on utilise une structure de commutation avec *SRD* "*Swing Reduced Deriver*" ainsi que les transistors fantômes comme nous l'avons proposé dans [51] (voir chapitre 6).

En conclusion, l'effet d'impédance limitée des cellules du CNA est dominé par la résistance, mais pour un CNA interne CMOS, elle ne pose pas de problème sérieux tant que le nombre des cellules reste faible. En revanche, d'autres effets comme l'impulsion transitoire "*Glitch*" et le temps d'établissement doivent être examinés.

3.5.6 Erreurs dynamiques des CNA

Certaines imperfections ont une influence considérable sur les performances dynamiques du CNA. Ces erreurs dynamiques sont liées à la forme du signal de sortie du CNA et aux caractéristiques des commutations des cellules. Ces erreurs dynamiques sont classées comme suit :

- déséquilibre entre le temps de montée et le temps de descente,
- phénomène de "*Glitch*" au instants de commutation,
- injection de charge d'horloge "*CFT*"
- gigue d'horloge,

Effet des impulsions transitoires *Glitches*

Il s'agit d'impulsions transitoires du signal de sortie dues aux commutations à grande vitesse des sources de courant. Pendant une transition entre deux niveaux de sortie, un niveau de sortie erroné peut apparaître à cause de la différence des temps de propagation de chaque bit. Par exemple dans un CNA binaire, pour une entrée digitale qui varie entre $\overline{0111111}$ et $\overline{1000000}$, si le bit de poids le plus fort *MSB* change d'état plus rapidement que le bit de poids le plus faible *LSB*, la sortie va être

$\overline{1111111}$ pendant un court instant. De la même façon, la sortie peut correspondre au code $\overline{0000000}$ si la *LSB* varie plus rapidement que le *MSB*. La figure ci-dessous illustre ce phénomène. La puissance

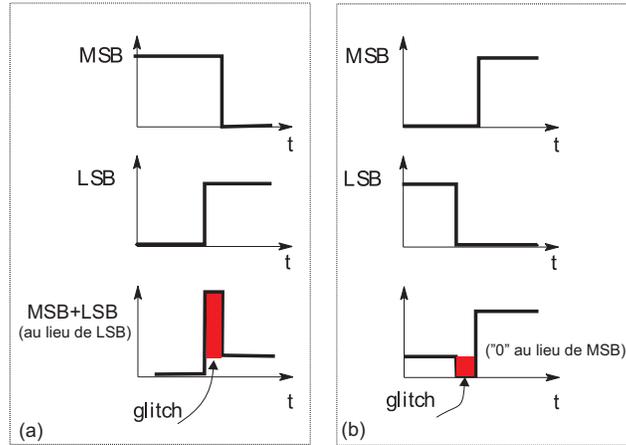


FIG. 3.22 – Les impulsions transitoires, a) glitch positif, b) glitch négatif.

d'erreur du glitch dépend du nombre de bits commutés, exprimé par $[N_x(n)]$, et du décalage entre le moment de la commutation "on" vers "off" et le moment de la commutation "off" vers "on", exprimé par τ_i pour la cellule numéro "*i*" avec une sortie de valeur nominale I_i . Dans le cas d'un CNA en courant, l'erreur de glitch peut s'exprimer par :

$$\sigma_{glitch} = \sum_{i \in N_x(n)} I_i \frac{\tau_i}{T_e} \quad (3.74)$$

La valeur ΔI_{glitch} peut être négative ou positive. En pratique, le signal de glitch n'a pas une forme rectangulaire comme représentée à la figure 3.22, par conséquent, τ_i doit exprimer une valeur effective de sorte que la surface finale $\tau_i I_i$ reste équivalente à celle de la surface réelle. De plus, on peut simplifier l'équation en prenant une valeur moyenne pour toutes les τ_i dénoté par τ_{im} , ce qui est en fait la différence moyenne d'entre le temps de propagation "0" et celle de propagation "1" dans les cellules. L'erreur maximale d'un CNA B-bit binaire correspond à l'entrée de moitié d'échelle selon la formule [102] :

$$\sigma_{glitch,max} = \sqrt{2B-1} \cdot I_{LSB} \frac{\tau_m}{T_e} \quad (3.75)$$

Du point de vue statique, cette erreur doit être inférieure à $0,5LSB$ comme celle de l'INL.

L'erreur de glitch est une grande faiblesse des convertisseurs binaires. Cependant, dans le cas d'un CNA thermométrique sans méthode de DEM, le problème est moins grave que dans le cas précédent, car l'erreur maximale est liée à la différence entre deux codes consécutifs $N_x(n)$. Dans un CNA sur échantillonné, une telle différence est en général assez petite (par rapport au terme $(2^B - 1)$ dans l'équation ci-dessus). Par contre, en utilisant une méthode de DEM (n'importe quelle méthode sauf la "Restricted-DEM" [106]) la probabilité d'erreur de glitch augmente ce qui limite la performance de DEM.

Plusieurs techniques ont été développées afin de résoudre le problème du glitch. Parmi ces techniques,

l'utilisation d'un CNA de type remise à zéro RZ ("Return-to-Zero") peut complètement éliminer le glitch, mais, la sortie rencontre d'autres problèmes comme l'effet de la gigue d'horloge et la non linéarité imposée par l'amplitude supplémentaire des sorties de la cellule RZ. Dans le chapitre 6 nous présenterons un nouveau type de CNA-RZ qui peut limiter l'effet de glitch sans autre effet dégradant supplémentaire.

Une autre méthode, qui est préférable, consiste à utiliser un bloc d'anti-glitch "deglitcher" suivi d'un étage de synchronisation. Un circuit deglitcher équilibre le temps de propagation de l'état "1" à l'état "0" et le temps de propagation inverse. Dans le chapitre 6, nous présenterons aussi un nouveau procédé [51].

Injection de charge d'horloge (CFT, clock feedthrough)

Cette erreur constitue une limitation majeure des sources de courants commutés. Elle se produit lorsque les transistors interrupteurs passent de l'état *on* à l'état *off*. Il y a tout d'abord la fuite des électrons stockés dans le canal quand l'interrupteur était passant. La grande partie des électrons s'évacue dans la source et le drain. Il y a aussi un transfert d'énergie entre la grille, la source et le drain, à chaque variation de la tension V_{gs} en raison des capacités parasites C_{gs} et C_{gd} (voire la figure 3.23), ce qui peut se limiter en employant des transistors fantômes. Une technique complexe, qui peut

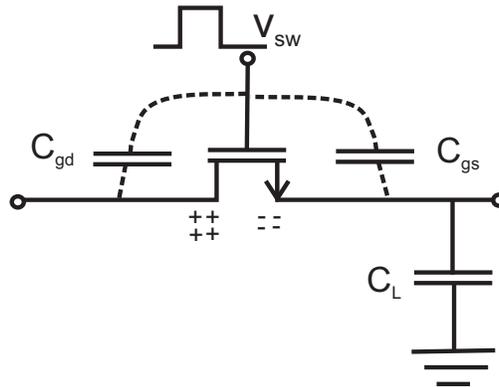


FIG. 3.23 – Illustration de l'injection de charge d'horloge dans une cellule de CNA de type courant commuté.

limiter ces effets, sera présentée au chapitre 6.

Temps de transition, t_{on} et t_{off}

Le processus de commutation des sources de courant peut se répartir en deux phases distinctes. Durant la première phase, les signaux de commande font basculer le transistor interrupteur de l'état bloqué (*off*) à la saturation (*on*), tandis que l'interrupteur complémentaire va basculer dans le sens opposé. Cette phase s'achève quand les signaux de commandes atteignent leur valeur finale.

Durant la deuxième phase, le potentiel de drain du transistor interrupteur V_o (nœud *o* dans la figure 3.20) ainsi que celui de la source commune V_X varient jusqu'à atteindre leur état permanent. Les variations du potentiel V_o dépendent de la capacité parasite au nœud *o* et de la charge à la sortie (Z_l), et le potentiel V_X dépend de la capacité parasite au nœud *X* et de la transconductance du

commutateur (S_W).

Ainsi, une longueur et une largeur minimale des transistors commutateurs S_W et une largeur minimale du transistor C_S permet de réduire le temps d'établissement. Cependant, les largeurs des transistors S_W et C_S sont directement liées aux capacités parasites. Dans ce cas, augmenter la largeur des commutateurs S_W peut compromettre le transfert de charge (*clock-feedthrough*).

Malgré tous les moyens d'optimisation au niveau du circuit, le temps de transition des cellules du CNA à courant commuté peut encore influencer la performance du système par deux effets. Comme cela a été présenté dans la section précédente, une différence non nulle entre le temps de transition de "0" à "1" $t_{on} = t_d + t_r$ et le temps de transition au sens opposé $t_{off} = t'_d + t_f$, fait apparaître une erreur de glitch. Même si on ajuste ces paramètres pour que leur différence soit nulle ($t_{on} = t_{off} \neq 0$), leurs valeurs absolues ajoutent une erreur importante dans le cas de CNA de type NRZ "Non-Return-to-Zero". Comme le montre la figure 3.24, la surface (équivalent de charge transférée) pour deux séries de données équivalentes (par exemple, $\{0,1,1,0,0\}$ et $\{0,1,0,1,0\}$) n'est pas la même, ce qui produit une erreur dépendant du signal d'entrée. Ainsi, on préfère utiliser un CNA de type RZ en diminuant ses effets secondaires. Dans le cas d'un CNA-RZ, il suffit de considérer la surface réelle d'une impulsion équivalent à celle de sortie de cellule au niveau système [99, 107]. En tous cas, l'ensemble des temps de transition doit être plus petit que la période de commutation, $t_{on} + t_{off} \ll T_e$.

3.6 Effet de la gigue d'horloge

Un signal d'horloge réel possède en général un peu d'incertitude sur la valeur de sa période appelée la gigue d'horloge. Ainsi, chaque fois qu'une opération doit se réaliser à un instant précis, par exemple l'échantillonnage, la gigue d'horloge peut dégrader les performances du système (voir la figure 3.25-a). Si l'erreur de la période d'horloge T_e est notée σ_τ , l'expression de l'erreur sur la sortie d'un échantillonneur est donnée à l'équation 3.17, où elle dépend de l'amplitude et de la fréquence de signal d'entrée. Ce principe, c.-à-d. la dépendance par rapport à la variation d'amplitude et de la fréquence de signal d'entrée est toujours la même, mais l'influence est différente selon le type de modulateur.

Dans un modulateur de type discret, l'échantillonnage s'effectue avant que le signal n'entre dans la boucle du modulateur. En pratique, l'erreur de SAH incluant l'effet de gigue limite la fréquence et la performance du convertisseur DT de sorte que leurs applications actuelles restent inférieures à quelques dizaines de MHz. Et puis, l'autre signal d'entrée de la boucle, c'est à dire la sortie de CNA-SC a une forme de transition exponentielle. Par conséquent, l'effet de la gigue d'horloge a une influence limitée sur la performance d'un modulateur discret, sauf si la bande passante de l'AOP de l'intégrateur est limitée [1].

En revanche, l'effet de la gigue d'horloge devient un obstacle fondamental pour un modulateur $\Sigma\Delta$ à temps continu [40, 43, 108–110]. La dégradation de la performance dû à la gigue peut varier en fonction de l'architecture et des caractéristiques de modulateur surtout celles de CNA interne. Dans cette section, nous voulons évoquer le phénomène de la gigue d'horloge. Les résultats de cette analyse

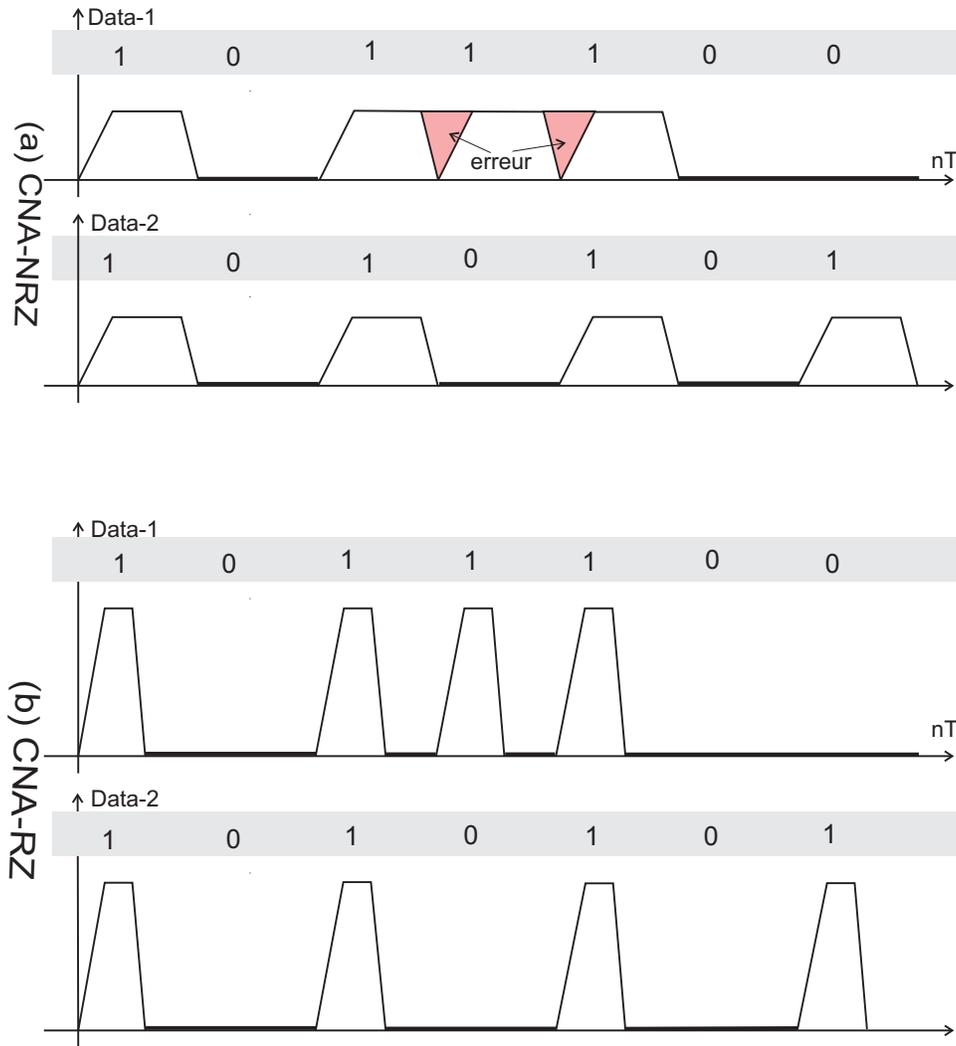


FIG. 3.24 – Erreur de la sortie d'une cellule du CNA pour deux séries de donn e  quivalent : a)NRZ, b)RZ.

seront utiles afin de pr senter une nouvelle solution d'am lioration au chapitre 6.

3.6.1 Effet de la gigue d'horloge dans un modulateur   temps discret

Dans un modulateur   temps discret, l'erreur de la phase d'horloge est essentiellement induite par le SAH. Pour une entr e $x(t) = A_{in} \sin(2\pi ft)$, le spectre et la puissance dans la bande de cette erreur peuvent s'exprimer comme suit [44] :

$$\begin{aligned}
 S_{j,DT}(f) &= \frac{A_{in}^2}{2} \cdot \left(\frac{2\pi f}{f_e}\right)^2 \cdot \left(\frac{\sigma_t}{T_e}\right)^2 \cdot \left(\frac{1}{f_e}\right)^2 \\
 N_{j,DT,max}^2 &= \frac{\pi^2 A_{in,max}^2}{2OSR^3} \cdot \left(\frac{\sigma_t}{T_e}\right)^2
 \end{aligned} \tag{3.76}$$

o , σ_τ^2 est la variance de gigue d'horloge. Contrairement au SAH, l'erreur de gigue a un effet n gligeable par rapport au filtre et au quantificateur interne puisqu'ils ont un temps d' tablissement assez court

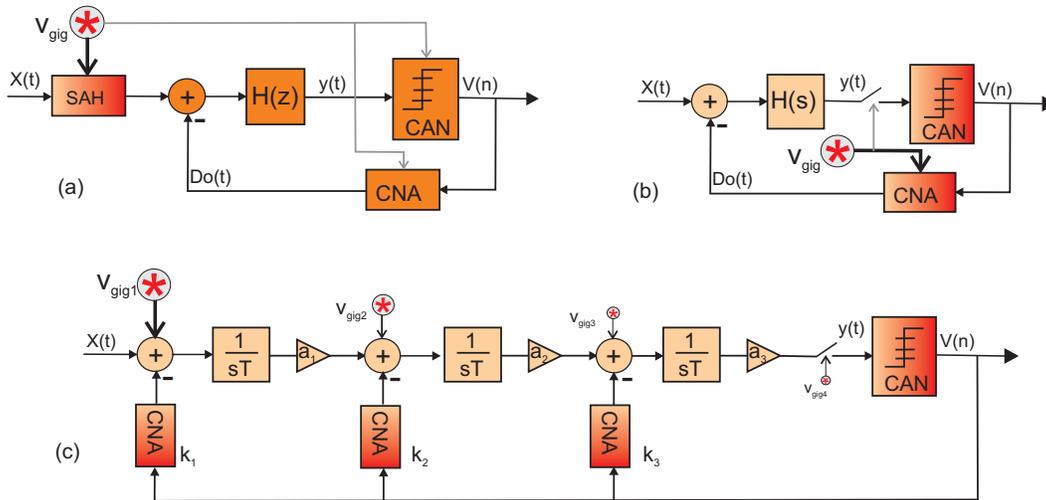


FIG. 3.25 – Lieu d'erreur de gigue d'horloge dans un modulateur: a) DT, b) CT général, c) exemple d'un CT de troisième ordre.

en utilisant des signaux produits par le circuit à capacité commutées "SC", comme sur la figure 3.26. Récemment, [44] a, pour la première fois, montré que si la bande passante de l'AOP utilisé dans le

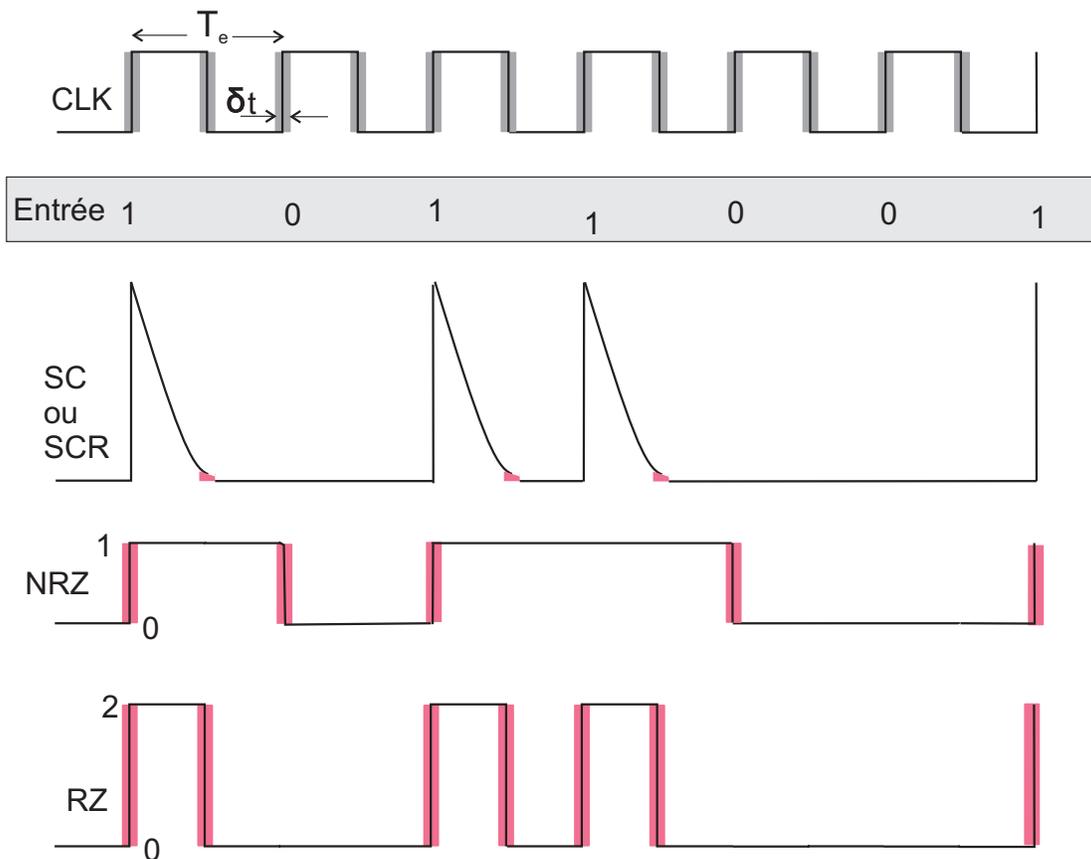


FIG. 3.26 – L'erreur de gigue d'horloge pour une cellule de CNA a) SC b) NRZ, c) RZ avec une durée de 50%.

filtre d'un modulateur DT n'est pas suffisamment large, la constante de temps des circuits SC, τ_{SC} peut varier vers une valeur effective plus grande faisant de sorte que l'effet de la gigue augmente.

3.6.2 Effet de la gigue d'horloge dans un modulateur CT

Dans le modulateur à temps continu "CT" (figure 3.25-b), deux sources d'erreur dues à une horloge non idéale existent, au sein du quantificateur et du CNA interne. L'erreur d'échantillonnage, qui se produit dans le quantificateur est rejetée par le gain de boucle $|NTF(s)|$. Cependant, l'importance de l'effet de gigue d'horloge dans un modulateur CT est liée aux commutations du CNA, comme cela a été montré dans beaucoup de publications. Ceci dû à la forme rectangulaire des signaux de sortie des cellules du CNA généralement utilisées dans le cas CT. La sortie de la cellule est modulée par la variation du temps de commutation comme sur la figure 3.26. Dans le cas d'une cellule NRZ, la gigue influence la sortie une fois par période si l'entrée de la cellule change, sinon il n'y a pas d'erreur. En revanche, la gigue d'horloge influence la sortie non nulle de chaque cellule type RZ deux fois par période au front montant et au front descendant. Notons que les erreurs induites aux deux fronts différents sont en général indépendantes. L'amplitude d'un signal RZ (avec une durée de 50%) est aussi deux fois plus grande que celle de signal NRZ équivalent.

Afin de mettre en équation l'effet de la gigue d'horloge, plusieurs méthodes analytiques ont été développées [33, 40, 43, 108–110]. Dans une approche complète, l'effet de la gigue doit être examiné pour toutes les branches de rebouclage comme dans l'exemple d'un modulateur CT d'ordre trois sur la figure 3.25-c. Pour chacune des sources, l'erreur de gigue ainsi que la fonction de transfert de chaque branche à la sortie du modulateur doivent être estimées séparément. Cependant, on peut constater que l'effet de la gigue est une erreur de type additionnel, et donc, celle du premier étage joue un rôle plus important que celles des étages suivants. Nous considérons ici seulement l'erreur du premier étage dans les deux cas NRZ et RZ. En raison de la présence d'au moins un intégrateur (ou résonateur) entre la sortie du CNA et l'entrée du quantificateur, seule la surface totale du signal, (charge du $\int_{nT_e}^{[(n+1)T_e+\delta\tau]} I(t)dt$), produite par le CNA peut définir l'erreur de gigue. Donc, la sortie de chaque cellule est considérée comme un rectangle d'amplitude Δ et de taille $\tau + \delta\tau$ au lieu de la forme réelle. De plus nous supposons que la gigue d'horloge est une erreur de type Gaussien (avec une moyenne nulle). En réalité, elle n'est pas blanche puisqu'elle a une distribution de type $\frac{\pm 1}{(f-f_e)}$ comme le montre la figure 3.27. Afin de prendre en compte la forme réelle de la gigue, il faudrait faire

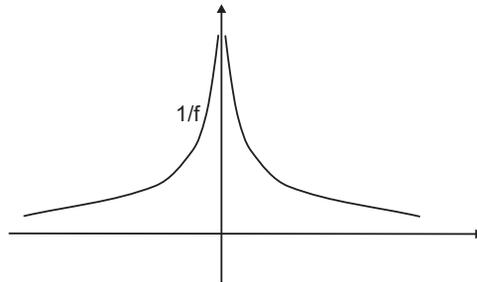


FIG. 3.27 – La densité spectrale de gigue d'horloge réelle, ce qui n'est pas blanche

une analyse plus détaillée, ce qui n'est pas envisagé dans le travail actuel.

Pour un CNA de type NRZ ($\tau = T_e$), avec une entrée $v(n)$, l'expression de cette erreur est [43, 108–111] :

$$\begin{aligned} e_{j,NRZ}(n) &= [v(n) - v(n-1)] \frac{\delta\tau}{T_e} \\ N_{j,CT,NRZ}^2 &\simeq \text{var}[e_{j,NRZ}] \times \frac{1}{OSR} = \sigma_{\Delta_v}^2 \times \frac{\sigma_\tau^2}{T_e^2} \times \frac{1}{OSR} \end{aligned} \quad (3.77)$$

où le spectre de la gigue est supposé blanc, $\sigma_{\Delta_v}^2$ est la variance de la variation d'entrée du CNA, puis $N_{j,CT,NRZ}^2$ montre la totalité du bruit de la gigue dans la bande utile. Par conséquent, la résolution maximale du CNA s'exprime comme suit :

$$ENOB_{j,CT,NRZ} = \text{Log}_2 \left[\frac{M}{2\sqrt{3}N_{j,CT,NRZ}} \right] = \text{Log}_2 \left[\frac{M\sqrt{OSR}}{2\sqrt{3} \sigma_{\Delta_v} \frac{\sigma_\tau}{T_e}} \right] \quad (3.78)$$

La valeur de σ_{Δ_v} doit être évaluée en fonction de l'architecture du modulateur et des caractéristiques du signal d'entrée. Elle dépend en pratique de deux facteurs différents : la vitesse de variation d'entrée, et l'erreur de quantification. Dans le cas d'un modulateur monobit, σ_{Δ_v} est estimé de manière analytique [108] et de manière numérique [109] (de l'ordre de $\frac{2}{3}$). Dans le cas d'un CNA multibit, la valeur exacte de σ_{Δ_v} vaut entre "1" et $\frac{2M}{3}$ [108]. Si on considère que l'entrée du modulateur $x(t)$ est limitée dans la bande avec un OSR assez grand (par exemple $OSR > M$), σ_{Δ_v} reste modérée dans la limite de 1 LSB. Cela confère au CNA multibit de type NRZ; cela limite efficacement l'effet de la gigue. Cette expression se vérifie par différentes simulations qui montrent, par exemple, qu'une gigue de l'ordre de 1% limite la résolution finale d'un modulateur monobit à 8 bit quand l'OSR vaut 64. Cette même gigue d'horloge limite la résolution du même modulateur mais avec un CNA 4-bit, à 12 bit.

Dans le cas d'un CNA de type RZ, il faut considérer trois différences par rapport au cas NRZ :

- le nombre de fronts pour lesquels la gigue intervient est de deux par période d'horloge,
- l'amplitude d'une cellule RZ de largeur d'impulsion τ qui doit être multipliée par $\frac{T_e}{\tau}$ pour fournir la même énergie,
- σ_{Δ_v} doit être remplacé par la variance d'entrée absolue σ_v , car l'effet de gigue d'une cellule RZ dépend de la valeur absolue d'excitation.

Ainsi, l'effet de la gigue dans la bande utile devint au moins deux fois plus élevé pour le cas RZ que celui du cas NRZ. Cela peut s'exprimer par la relation suivante [108, 110] :

$$N_{j,CT,RZ}^2 \simeq 2\sigma_v^2 \times \frac{\sigma_\tau^2}{T_e^2} \times \frac{1}{OSR} \quad (3.79)$$

Dans le cas monobit, σ_v vaut 2 au lieu de 0.7 pour σ_{Δ_v} dans l'expression 3.77.

Cependant, une très grande différence peut apparaître entre le CNA NRZ et le CNA RZ quant ils ont une structure multibit. Comme on l'a vu précédemment, la variation de l'entrée d'un CNA multibit est de l'ordre d'un LSB, alors que la valeur d'absolue de l'entrée varie de "0" à "M". Autrement dit, la résolution maximale d'un CNA de B bit de type RZ peut être limitée par la gigue d'horloge

d'environ $(B+1)$ bits inférieure à celle d'un CNA similaire de type NRZ.

Il existe un cas encore plus mauvais, qui est celui de l'utilisation de deux CNA RZ complémentaires. Cette structure est un bon candidat pour la conception d'un modulateur à temps continu, on peut grâce à elle ajuster l'équivalence entre la fonction de transfert DT et CT en faisant varier les coefficients d'un rebouclage multiple [101]. La puissance de bruit de la gigue peut doubler par rapport au niveau donné par l'équation 3.79 [109].

Il faut rappeler que l'effet de la gigue analysé dans cette section est général de sorte qu'il n'y a pas différence selon le type d'application passe-bas ou passe-bande. Cependant, en raison de la présence directe de la valeur absolue de la période d'échantillonnage ($T_e = \frac{1}{f_e}$), dans tous les cas, la performance d'un modulateur passe-bande est plus susceptible d'être limitée par la gigue que celle d'un modulateur passe-bas qui fonctionne en général en basse fréquence.

Par ailleurs, nous avons une remarque importante non rapportée dans la littérature sur la différence entre le fonctionnement d'un modulateur multibit passe-bande et celle d'un modulateur multibit passe-bas vis à vis de la gigue d'horloge. En utilisant un CNA RZ dans le cas passe-bas, l'effet de la gigue est en général indépendant de l'amplitude de l'entrée, alors que, dans le cas passe-bande, cela est le contraire. Plus de détails sur cette remarque seront donnés au chapitre 6.

Solution proposée afin de diminuer l'effet de la gigue

En gros, cinq solutions ont été déjà présentées par lesquelles l'effet de gigue d'horloge, dans un modulateur $\Sigma\Delta$ à temps continu, peut être limité.

- utilisation d'un CNA ayant une forme de signal du sortie exponentiel décroissant ("decaying"), par exemple, de type SCR [43] ou de type sinusoïdal [112,113],
- optimisation de la fonction de transfert de bruit de quantification,
- utilisation d'un CNA multibit de type semi RZ,
- utilisation d'un filtre de type FIR après le CNA,
- utilisation des filtres de type "Transmission-line"

Un CNA de type SCR utilise une forme exponentielle produite par une capacité commutée (C) au travers d'une résistance (R), comme sur la figure 3.28. Afin de minimiser l'effet de la gigue, la constante

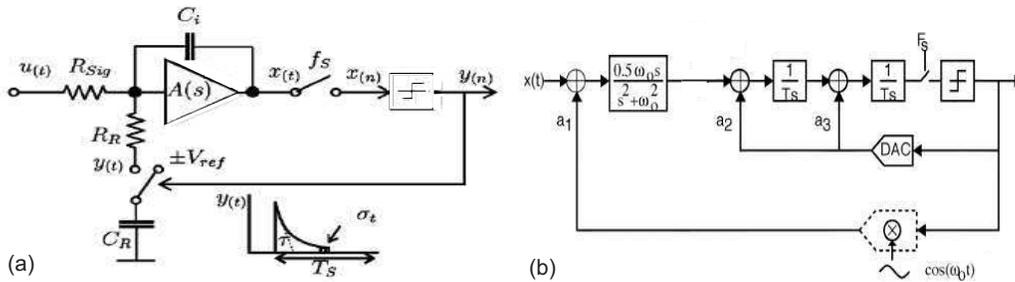


FIG. 3.28 – Modulateur à temps continu monobit a) avec un CNA de type SCR, b) avec un CNA sinusoïdal

de temps de $\tau_{SCR} = RC$ doit être d'environ 10 à 15 fois inférieure à la période d'échantillonnage,

et la dynamique (amplitude) de sortie du CNA doit être d'un ordre plus élevé que dans le cas normal pour le même niveau d'entrée. Cette structure rend un modulateur CT aussi robuste vis-à-vis de la gigue qu'un modulateur DT, cependant, certains avantages de réalisation en mode CT sont malheureusement éliminés. Par exemple, la bande passante, la vitesse de balayage et la plage dynamique des AOP doivent être conçus avec environ les mêmes limites que celles du modulateur DT. D'autre part, la réalisation d'un CNA ayant une sortie de forme par exemple sinusoïdale est en pratique complexe et coûteuse. La fréquence de sortie d'un CNA de forme sinusoïdale doit être synchronisée par la fréquence de l'horloge d'échantillonnage, ainsi que l'amplitude de ce type de CNA est au moins 3 fois grande que la celle d'un CNA NRZ normal.

La deuxième solution, c.-à-d. l'optimisation de la NTF du modulateur, a été présentée dans [114]. Cette solution est basée sur l'équation 3.77. En négligeant la faible variation d'entrée d'un modulateur devant le bruit de quantification σ_q , le terme $\sigma_{\Delta_v}^2$ peut être estimé par la relation suivante :

$$\begin{aligned} \sigma_{\Delta_v}^2 &= \text{var}[v(n) - v(n-1)] = \text{var}[(1 - z^{-1})V(Z)] \\ &\simeq \frac{\sigma_q^2}{2\pi} \int_0^{2\pi} |(1 - e^{-j\omega}) \cdot NTF(e^{j\omega})|^2 d\omega \end{aligned} \quad (3.80)$$

$\sigma_{\Delta_v}^2$ peut être minimisé en minimisant l'ensemble de la multiplication ci-dessus dans la bande, comme décrit sur la figure 3.29. Elle permet une amélioration limitée en jouant avec NTF, que peut être

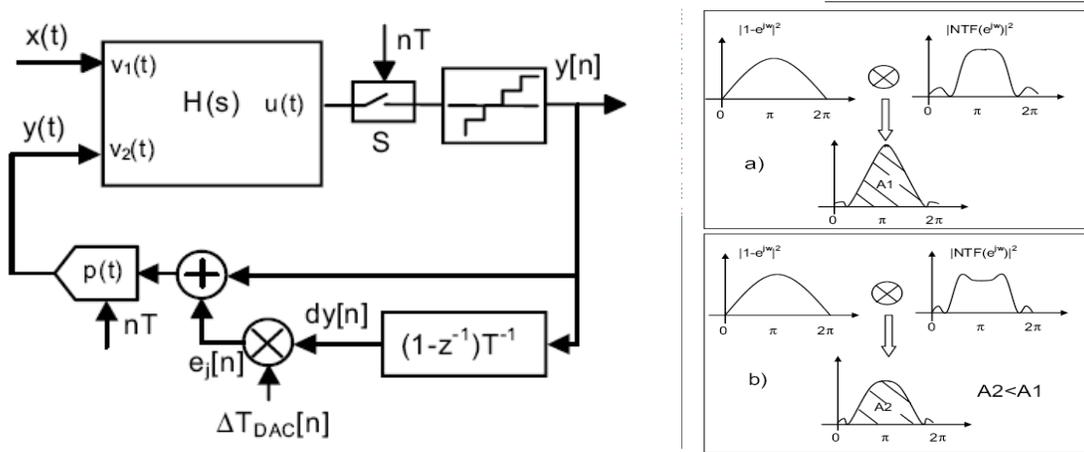


FIG. 3.29 – Solution proposé afin de réduire l'effet de gigue d'hologe en faisant varier NTF a) modèle de gigue d'un CNA NRZ, b) l'optimisation de NTF dans la bande.

disponible si le bruit de quantification n'est pas à la limite de résolution du système. En plus, si la gigue a une forme non Gaussienne cette solution devient plus intéressante.

Cependant, nous estimons que le meilleur moyen de lutter contre l'effet de la gigue d'horloge consiste en pratique à utiliser d'un CNA semi-RZ grâce auquel les avantages du mode RZ peuvent être obtenus sans subir la contrainte d'effets supplémentaires de gigue plutôt qu'un CNA NRZ [48]. Cette solution peut aussi être employée avec la solution précédente, pour laquelle nous proposons des circuit au chapitre 6.

[115] a présenté une autre solution par laquelle l'effet de gigue d'horloge peut être repoussé en dehors

de la bande utile par le filtrage du signal de rebouclage. Cette méthode peut être utilisée pour un CNA monobit. Dans le cas d'un CNA multibit, les circuits supplémentaires deviennent complexes. Le principe est représenté sur la figure 3.30 et les détails se trouvent dans les références [111,115]

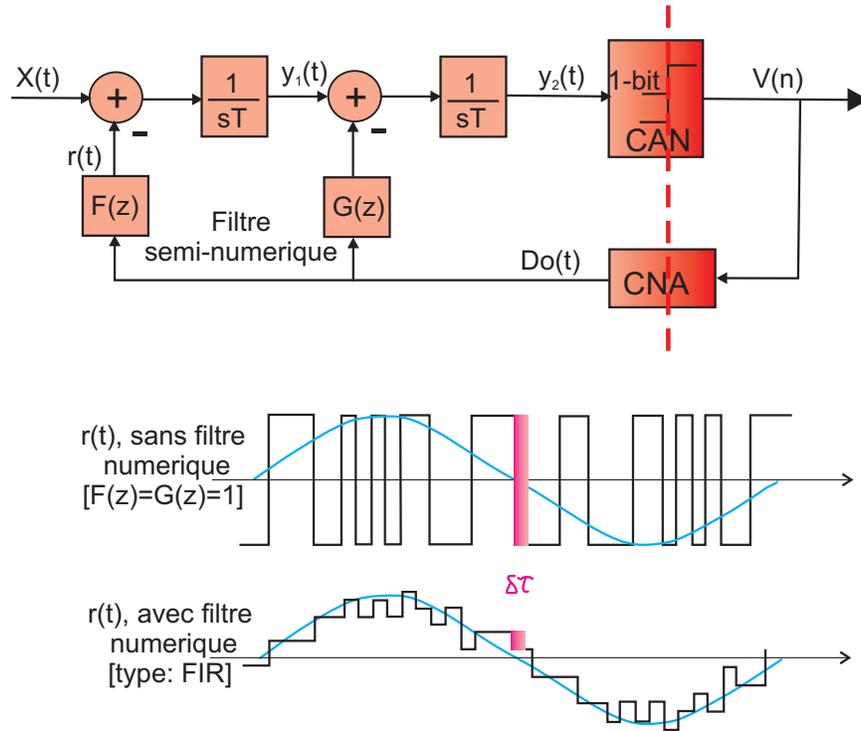


FIG. 3.30 – Solution proposée afin de réduire l'effet de gigue d'horloge par un filtrage Semi-FIR sur le signal de rebouclage a) position du filtre, b) un exemple de fonctionnement de filtre.

La dernière solution, qui est présentée dans les références [25,26], tente de tirer bénéfice de la fonction de transfert des guides d'ondes, par exemple, celle des "Transmission-line". Une telle fonction de transfert a des caractéristiques similaires aux circuits à capacités commutées SC. Cette solution n'entre pas dans le cadre de notre étude parce que, les éléments des guides d'ondes ne sont pas compatibles avec la technologie actuelle CMOS, de plus ils sont utiles pour des signaux au-delà du Giga-Hertz. On peut trouver plus de détails dans les références données.

3.7 Retard de la boucle

Chacun des blocs du modulateur $\Sigma\Delta$ peut en pratique produire un retard qui dépend de la structure et de la technologie. La tolérance du système vis à vis du retard dépend de sa valeur relative par rapport à la période d'échantillonnage. L'ensemble du retard de boucle doit être considéré dès l'étape de conception du modulateur [107]. En général, le retard du quantificateur et le retard du CNA constituent la bonne partie du retard de la boucle. La somme de ces deux retards peut être examinée comme un retard unique nommée "*ELD, Excess Loop Delay*" [32–36,99,107]. Un retard inférieur à une demi période d'échantillonnage ne pose pas de problèmes dans un modulateur discret. Cependant, la performance d'un modulateur à temps continu est très sensible au retard des éléments

dans la boucle. L'ELD contribue à l'augmentation du niveau de bruit dans la bande ainsi que de l'instabilité de la boucle. Un retard comparable à une période d'échantillonnage peut potentiellement introduire un pôle supplémentaire pour la boucle, et par conséquent, un modulateur théoriquement stable peut devenir instable.

Il existe deux méthodes de compensation du retard dans un modulateur à temps continu. Dans la première solution présentée dans [34], le modulateur utilise deux chemins de rebouclage ayant une demi-période de décalage comme sur la figure 3.31-a. Afin d'obtenir un modulateur CT équivalent à

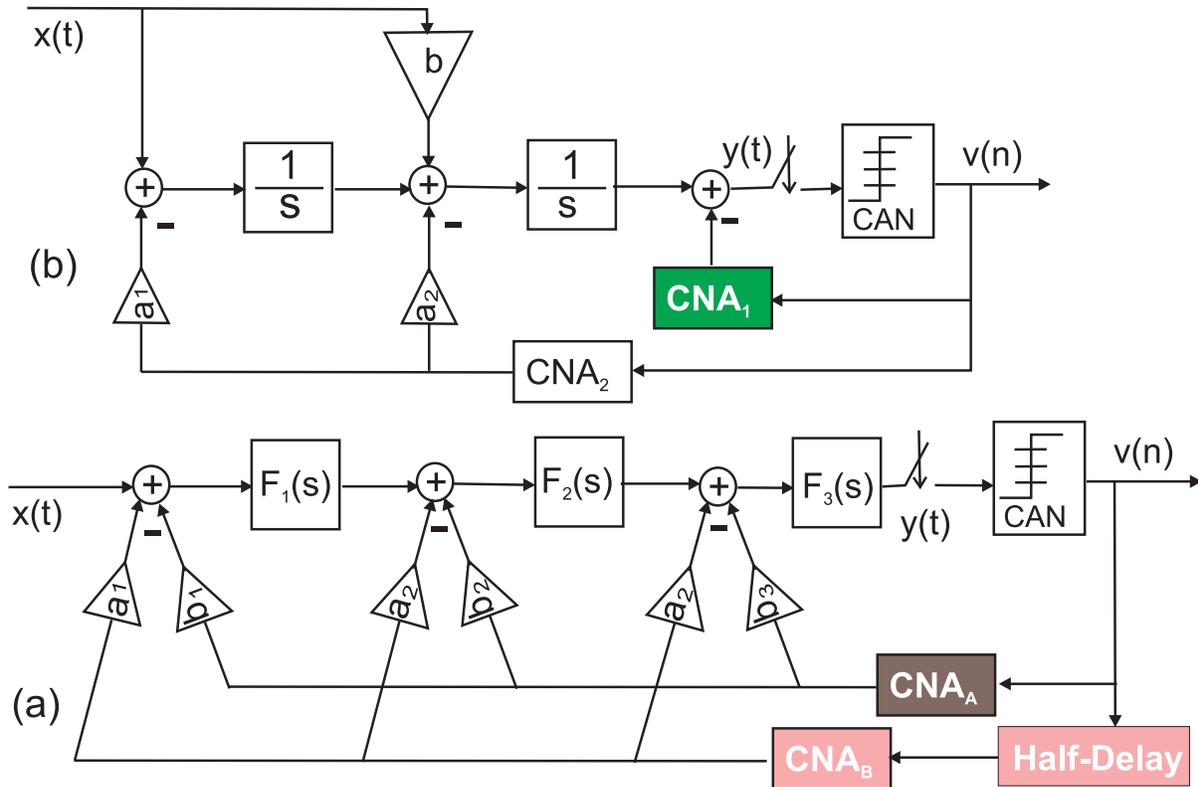


FIG. 3.31 – Solutions proposées afin de réduire l'effet de retard dans la boucle d'un modulateur CT: a) rebouclage multiple [34], b) rebouclage rapide [30]

celui DT, les coefficients de la boucle sont ajustés en tenant compte des valeurs des retards. Cette solution présente deux faiblesses. D'abord, la structure devient complexe et elle a besoin de deux séries de CNA appariés, ce qui est difficile à mettre en œuvre en pratique. Deuxièmement, comme on l'a dit auparavant, l'effet de la gigue d'horloge est plus important dans une structure à rebouclage multiple (qui sont souvent en mode RZ).

La meilleure solution [30], qui est utilisée dans plusieurs réalisations, consiste à utiliser un rebouclage entre la sortie du CNA et l'entrée du quantificateur comme le montre l'exemple de la figure 3.31-b [36,50,51,99,107]. Dans cette solution, ni l'effet de gigue ni le défaut d'appariement entre les CNA ne pose de problèmes. Pour plus de détails, on peut se répondre aux références [24,29].

3.8 Conclusion

Nous venons de décrire dans ce chapitre les différences sources d'erreur dans un modulateur $\Sigma\Delta$. Elles sont nombreuses et réparties dans la structure du modulateur. Nous avons analysé séparément l'influence de chacune de ces sources, puis nous avons donnée un calcul analytique, plus ou moins approximatif, pour la résolution maximale en tenant compte de chacune des sources d'erreur. De plus, nous avons discuté des différentes solutions possibles pour réduire l'influence de ces erreurs.

Un modulateur monobit a besoin d'un *OSR* élevé et souffre ainsi de l'effet de la gigue d'horloge et de la non linéarité des AOP, alors qu'un modulateur multibit n'apporte pas d'amélioration si les cellules du CNA ne seront pas appariées.

Le tableau 3.1 récapitule les caractéristiques recherchées sur chaque bloc. Les divers points que nous

TAB. 3.1 – *Effets de la non idéalité des blocs du modulateur $\Sigma\Delta$ sur ses performance comme : " f_e " - "ENOB ou SNDR" - "SFDR"*

Bloc	Non idéalité	Effets en DT	Effets en CT	Solutions
SAH	<ul style="list-style-type: none"> • bruit thermique • gigue • C,R-parasites • consommation • retard 	$f_e \downarrow$ - ENOB \downarrow $f_e \downarrow$ - ENOB \downarrow SFDR \downarrow $f_e \downarrow$ - ENOB \downarrow \uparrow $f_e \downarrow$	- - - - - -	- - - - -
filtre (RLC) (OTA)	<ul style="list-style-type: none"> • bruit thermique • précision-absolu • précision-relative • consommation • BW • stabilité 	- ENOB \downarrow - ENOB \downarrow \uparrow $f_e \downarrow$ - SFDR \downarrow $\downarrow L > 2$	- ENOB \downarrow - ENOB \downarrow - \downarrow $\downarrow \downarrow L > 2$	- - multibit
CAN	<ul style="list-style-type: none"> • $t_{on} + t_{off}$ • gigue • gain du comp. • c-parasite • hystéries • appariement (R_i) • offset • consommation 	- - - - - - ENOB \downarrow \uparrow	$f_e \downarrow$ - SFDR \downarrow - $f_e \downarrow$ - ENOB \downarrow - SFDR \downarrow $f_e \downarrow$ $f_e \downarrow \downarrow$ - SFDR \downarrow - ENOB \downarrow \uparrow	ping-pong - ping-pong - -
CNA	<ul style="list-style-type: none"> • appariement-(α_i) • impédance • $t_{on} + t_{off}$ • gigue • CFT • glitch • consommation 	ENOB $\downarrow \downarrow$ - SFDR $\downarrow \downarrow$ - - - -	ENOB $\downarrow \downarrow$ - SFDR $\downarrow \downarrow$ ENOB \downarrow $f_e \downarrow$ - ENOB \downarrow - SFDR \downarrow ENOB $\downarrow \downarrow$ - SFDR \downarrow ENOB \downarrow - SFDR \downarrow ENOB \downarrow - SFDR \downarrow \uparrow	DEM - mode RZ multibit-NRZ SRD, T.fantôme mode RZ

avons présentés dans ce chapitre ont eu pour but de décrire les blocs les plus sensibles d'un modulateur $\Sigma\Delta$ selon l'application. Dans le cas d'un convertisseur haute résolution avec un *OSR* modéré, on

est obligé d'employer une quantification multibit ainsi qu'un rebouclage multibit. Par conséquent, les erreurs dominantes proviennent des aspects pratiques du CNA multibit dont les plus importants sont : l'appariement des cellules, les glitches, les temps de transitions non équilibrés et l'effet de la gigue d'horloge comme le montre le tableau 3.1. Les deux chapitres suivants présenteront des solutions pour limiter l'effet d'appariement des cellules du CNA. Afin de limiter les erreur dynamiques, certaines techniques de conception de circuit seront proposées au chapitre d'après.

Chapitre 4

Méthodes d'appariement dynamique de composantes dans un CNA multibit (*DEM*)

4.1 Introduction

Dans le chapitre précédent, nous avons présenté les différentes sources d'erreurs présentes dans les modulateurs $\Delta\Sigma$ multibits. L'une des erreurs les plus importantes est introduite par les défauts d'appariement entre les cellules du CNA. Avec les technologies intégrées actuelles, sans méthode de correction, les avantages des modulateurs multibits vus à la section 2.3.9, sont compensés par l'effet de ces erreurs, qui ne peuvent ainsi être meilleurs que leurs concurrents monobits. Dans ce chapitre nous allons étudier différentes méthodes de correction des défauts d'appariement applicables aux CNA thermométriques.

Ces méthodes sont classées en trois catégories :

1. ajustement physique des composant désappariés, par exemple : "*laser trimming*",
2. étalonnage "*calibration*" et la correction numérique,
3. méthode d'appariement dynamique de composantes ou réarrangement dynamique d'interconnexions de composants, "*Dynamic element matching*", DEM

Chacun de ces types de méthodes possède des particularités différentes, comme leur précision, leur coût, leur vitesse, leur consommation, leur fiabilité contre le vieillissement des composants, etc. Après une présentation générale des CNA thermométriques, nous nous sommes limités à la troisième catégorie, c'est à dire la DEM. Cette stratégie de correction, pour laquelle plusieurs méthodes de réalisation ont été proposées, est actuellement la plus utilisée dans le cas des convertisseurs de type sigma delta suréchantillonnés. Dans un premier temps, les algorithmes de DEM seront présentés en mettant l'accent sur leurs avantages et leurs faiblesses. Ensuite, les aspects complémentaires de chaque méthode seront discutés. Ces discussions sont nécessaires pour nous guider vers les nouveaux algorithmes qui seront développés au chapitre suivant. Des résultats de simulation ainsi que des comparaisons entre ces algorithmes aideront à développer des analyses et à valider les conclusions.

Certaines techniques de correction qui en dérivent seront aussi rappelées dans une section séparée.

Par ailleurs, les méthodes d'ajustement physique, l'étalonnage et la correction numérique seront brièvement étudiées en fin de chapitre.

4.2 Structure thermométrique du CNA

La majorité des modulateurs $\Delta\Sigma$ multibits utilisent un CNA thermométrique dans le chemin de contre réaction car sa vitesse et sa précision sont de ce fait plus élevées. Cependant, la structure CNA thermométrique présente certains inconvénients : le plus important réside dans la complexité des commutations. Tant que le nombre de cellules du CNA reste peu élevé (CNA de moins de 6-bits), cet inconvénient ne pose pas beaucoup de problèmes, bien au contraire, cela peut faciliter l'utilisation de méthodes de correction de son éventuelle non linéarité. La figure 4.1 montre la structure d'un CNA thermométrique avec un bloc de prétraitement, et un brassage des sources qui sélectionne dynamiquement les éléments unitaires (DEM). Comme le schéma le montre, une entrée

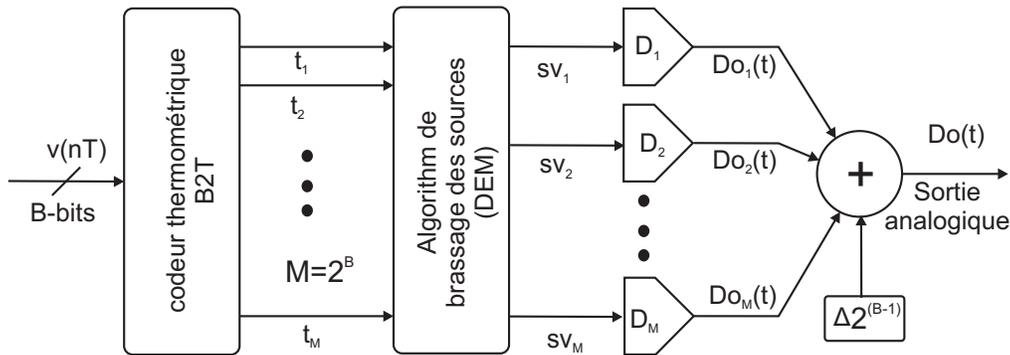


FIG. 4.1 – Schéma simple d'un CNA thermométrique avec DEM

B -bits (normalisée : $0 \leq v(nT) \leq M$, $M = 2^B$) est d'abord transformée en un code thermoélectrique ($t(n) = [t_1(n) \ t_2(n) \ \dots \ t_M]$, où $t_i(n) \in \{0, 1\}$, et $i \in \{1, 2, \dots, M\}$). Le bloc intermédiaire de DEM réarrange le vecteur $t(n)$ pour obtenir une nouvelle sélection $sv(n) = [sv_1(n) \ sv_2(n) \ \dots \ sv_M]$, $sv_i(n) \in \{0, 1\}$, et $i \in \{1, 2, \dots, M\}$. Le nombre des lignes actives, c.-à-d. pour lesquelles $sv_i(n) = 1$, est identique à celui du vecteur $t(n)$. Ces lignes actives sont choisies de sorte que la puissance moyenne du défaut d'appariement des cellules du CNA soit minimisée dans la bande utile. Le traitement numérique des deux premiers blocs de la figure 4.1 doit préserver la valeur du signal d'entrée du CNA à chaque instant, comme suit :

$$\begin{aligned}
 v(n) &= t(n) \ t^T(n) = \sum_{i=1}^M t_i(n) \\
 &= sv(n) \ sv^T(n) = \sum_{i=1}^M sv_i(n)
 \end{aligned} \tag{4.1}$$

où T est l'opérateur de transposition de matrice. Les erreurs de non idéalité du CNA peuvent être considérées comme étant de nature additionnelle stochastique, comme cela est supposé dans l'équation 3.38 et sur la figure 3.15. Dans le cas où il n'y a pas de brassage des sources, l'erreur maximale

de défaut d'appariement du CNA est donnée par les équations 3.60 et 3.61.

Prenons par exemple un modulateur sigma delta avec $OSR=64$, $M=16$, et une erreur d'appariement des cellules du CNA de variance $\sigma_\alpha = 1\%$. Cet exemple typique, qui offre idéalement une résolution de plus de 18 bits, sera examiné plusieurs fois dans ce chapitre. Sans aucune correction d'erreur, nous obtenons une résolution finale inférieure à 8bits, $ENOB \leq \text{Log}_2[\frac{\sqrt{M}}{\sqrt{3}\sigma_\alpha}] \simeq 8$ bit, si on suppose que le SNDR maximum se trouve environ 3dB inférieur à celui de la pleine échelle¹. Même si dans ce calcul nous avons négligé tous les autres bruits, une telle dégradation de la performance exige d'employer au moins une méthode de correction dans un modulateur multibit.

Une technique de DEM convenable peut améliorer cette résolution jusqu'à 16 bit, ou, avec une erreur d'appariement moins importante, par exemple $\sigma_\alpha = 0.001$ jusqu'à 20 bits. Les techniques de correction de défaut d'appariement les plus utilisées, basées sur la sélection dynamique d'éléments unitaires, sont classées comme suit :

- réarrangement aléatoires des éléments unitaires,
("Randomized selection", RDEM)
- réarrangement des composantes au niveau de l'horloge,
("Clock Averaging", CLK-DEM)
- réarrangement des composantes au niveau des codes utilisés,
("Individuel Level Averaging", ILA - DEM)
- réarrangement des composantes au niveau des données,
("Data Weighted Averaging", DWA-DEM)
- réarrangement des composantes par tri vectoriel,
("Vector feedback, ou Sorting based algorithm", SDEM)
- réarrangement des composantes en structure arborescente
("Tree structured", TDEM)

Chacun des algorithmes ci-dessus peut avoir plusieurs variantes selon les applications. Nous en parlerons par la suite.

Cependant, une perturbation dans l'algorithme de brassage de sources peut aggraver la dégradation des performances. Cette situation peut se produire si l'algorithme n'est pas suffisamment stable. En fait, cette perturbation est l'un des soucis les plus importants dans l'application des différentes méthodes de DEM, particulièrement pour les méthodes d'ordre élevé.

Un autre souci éventuel, qui peut devenir dominant dans l'utilisation des méthodes DEM simple du premier ordre, est le comportement tonal des algorithmes de réarrangement (périodique pour certaines valeurs d'entrée).

Dans les sections suivantes en présentant les différentes méthodes de correction proposées dans la littérature, nous essaierons d'analyser la probabilité de ces perturbations.

1. Dans le cas d'une entrée DC, la résolution finale peut être estimée par le $INL_{max} = \frac{\sigma_\alpha}{2\sqrt{M}}$ par la loi qui impose que $INL_{max} \leq \frac{LSB}{2}$. Cela donne aussi une résolution finale à l'ordre du $ENOB \simeq 8.3bit$.

4.3 Brassage aléatoires des sources, RDEM

Le défaut d'appariement des cellules du CNA interne introduit une erreur à la sortie du modulateur $\Delta\Sigma$, qui a été analysée dans le chapitre précédent (voir section 3.5.3). Cette erreur consiste en un bruit supposé blanc et des tonalités indésirables apparaissant dans la bande. Elles proviennent de la modulation entre les petites erreurs statiques des cellules du CNA et les erreurs du quantificateur de la boucle, ainsi qu'un certain nombre d'harmoniques des signaux d'entrées. Ces harmoniques sont produites par la modulation du défaut d'appariement des cellules par les entrées du modulateur. L'effet le plus grave est la dégradation dramatique du SFDR du système. La première idée pour combattre les harmoniques est de blanchir ces erreurs autant que possible [7]. Cette idée fonctionne bien en pratique dans le cas des modulateurs passe-bas et passe-bande, mais apporte peu d'amélioration au SNDR.

La figure 4.2 montre le schéma d'un assemblage du réarrangement aléatoire des sources unitaires, *Randomized Dynamic Element Matching*, (*RDEM*)². Un générateur introduit un jeu de possibilités

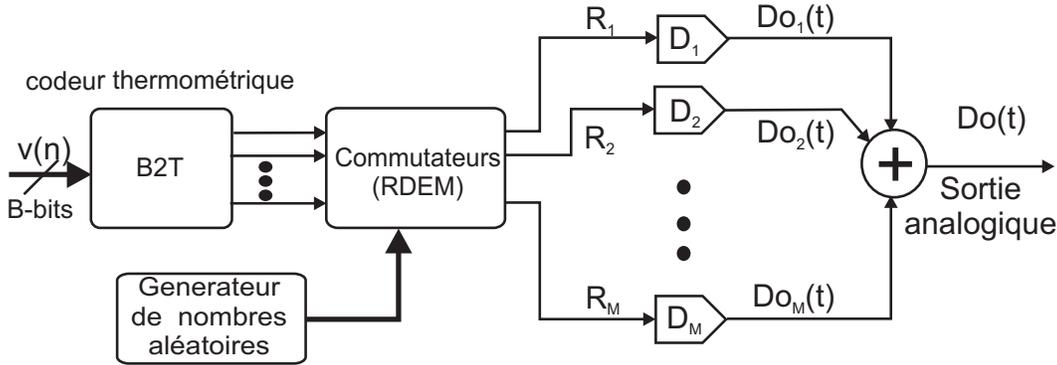


FIG. 4.2 – Schéma d'algorithm RDEM

$R(n)$, un tirage aléatoire de M -éléments, pour changer les connexions internes des codes d'entrée $t(n)$ au code sortant aléatoires $sv(n) \in \{R_1(n), R_2(n), \dots, R_v(n)\}$ en respectant l'équation 4.1.

Idéalement, ce générateur doit générer $M!$ différentes priorités aléatoires. Par exemple, si $v(n) = 5$ et $M = 8$, la table suivante montre quelques tirages possibles :

- $R(n1) = \{ 1,2,3,4,5,6,7,8 \}$, $sv(n1) = \{ 1,1,1,1,1,0,0,0 \}$
- $R(n2) = \{ 7,5,3,8,1,2,4,6 \}$, $sv(n2) = \{ 1,0,1,0,1,0,1,1 \}$
- $R(n3) = \{ 4,3,8,2,5,1,6,7 \}$, $sv(n3) = \{ 0,1,1,1,1,0,0,1 \}$
- $R(n4) = \{ 1,3,8,6,5,4,2,7 \}$, $sv(n4) = \{ 1,0,1,0,1,1,0,1 \}$
- ...

La séquence $R(n)$ se soit d'être non corrélée avec les entrées. En appliquant cet algorithme, la distorsion du signal est transformée en un bruit blanc à la sortie du CNA. Autrement dit, les tons générées par le CNA sont blanchis, mais, le niveau du bruit augmente pour toutes les fréquences. Comme en

2. La quasi totalité des informations présentées dans cette section ont déjà été publiées. Nous y ajoutons peu de nouveaux éléments. Mais comme le RDEM est une méthode de base, elle est présentée avec plus de détails.

l'absence de méthode d'appariement (équation 3.60 et 3.61), ce bruit ne peut pas être mis en forme par la fonction de transfert du bruit de quantification NTF , car il a été produit dans le chemin de rétroaction. En partant de l'équation de défaut d'appariement 3.46 pour une entrée normalisée, $x = v(n)$ avec supposée $\epsilon \cong 0$, l'erreur totale du CNA avec son étage de prétraitement de RDEM, peut être écrite au moyen de l'équation suivante :

$$e_{RDEM}(n) = \sum_i \alpha_i s v_i(n) \quad \text{pour : } i \in [x \text{ premiers indices du } R(n)], \quad (4.2)$$

Du fait du sur-échantillonnage, la puissance totale de bruit introduite dans la bande est divisée par l'OSR du système. De plus, grâce à l'application de la méthode RDEM, les éléments $s v_i(n)$ sont bien décorrélés (c.-à-d. $Covar_{i \neq j}[s v_i(n), s v_j(n)] = 0$), contrairement à ceux de l'équation 3.54. Par conséquent, en partant du pire cas de l'équation 3.61 au milieu de la pleine échelle $x = \frac{M}{2}$,³ la puissance maximum des erreurs du CNA normalisé, qui est composé de M cellules similaires ayant un défaut d'appariement σ_α , est obtenue au moyen de la relation suivante :

$$\sigma_{e_{RDEM,max}}^2 = \frac{\sigma_\alpha^2}{4OSR.M} \quad (4.3)$$

A partir de cette relation, la résolution maximale du CNA en dehors des erreurs de quantification devient la suivante [89] :

$$ENOB \leq \text{Log}_2 \left[\frac{\sqrt{M.OSR}}{\sqrt{3}\sigma_\alpha} \right] \text{ bits} \quad (4.4)$$

Par exemple, admettons une erreur de défaut d'appariement de 1%, $M=16$ et $OSR=64$, la résolution possible est d'environ 11.3 bits, l'amélioration ne plus valoir plus que 2.5 bits. La figure 4.3 montre le spectre d'erreur d'ensemble du CNA. Dans cet exemple, l'algorithme de RDEM améliore le SNDR d'environ 12dB dans les deux cas passe-bas et passe-bande.

Nous pouvons résumer cette méthode par la description suivante :

– **Avantages :**

- RDEM est simple,
- elle ne produit pas de tonalités fantomes,
- elle a le même effet sur les modulateurs passe-bas et passe-bande.

– **Inconvénients :**

- un générateur aléatoire idéal n'est pas facile à intégrer dans un circuit, on utilise donc un algorithme quasi aléatoire, ce qui en diminue l'efficacité,
- même si on met en œuvre l'algorithme quasi aléatoire, il sera encore difficile à réaliser pour un nombre des cellules élevé $M > 8$.
- il devient moins efficace pour un OSR petit.
- les erreurs dynamiques, en particulier le phénomène de glitch, détruisent l'efficacité de la RDEM pour des applications à fréquence élevée (comme c'est le cas pour toutes les méthodes de DEM).

3. ou $P(s v_i(n) = 1) = P(s v_i(n) = 0) = \frac{1}{2}$

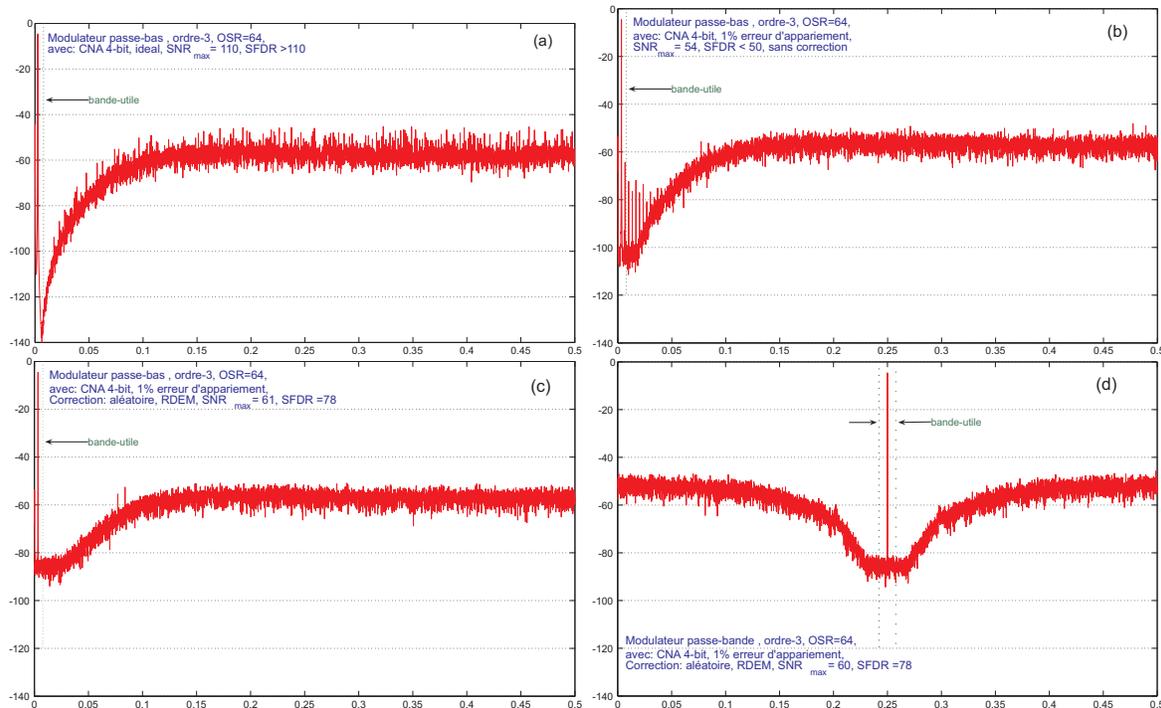


FIG. 4.3 – Spectre de la sortie du modulateur ayant 1% d'erreur de désappariement : a) idéal b) réel sans aucune correction c) correction par l'algorithme de RDEM d) cas similaire à (c) mais passe-bande

Les solutions proposées :

Comme cela a été dit auparavant, pour effectuer un algorithme de RDEM idéal, il faut générer un tirage aléatoire parmi $M!$ séries de codes aléatoires et appliquer ce code afin de choisir $v(n)$ -cellules indépendantes quasi instantanément à chaque période d'échantillonnage ($t = nT$). Malgré tout, même si l'idée de la RDEM et son implantation sont simples d'un point de vue algorithmique, sa réalisation au moyen de composants électroniques réels n'est pas si simple. Elle peut prendre beaucoup de place et consommer beaucoup de courant, surtout quand M est grand, par exemple quand $M \geq 8$. Une RDEM ayant un tel nombre de cellules (256), peut aussi limiter la vitesse de commutation. On cherche alors à simplifier les méthodes RDEM.

Une approche plus simple est celle du "Barrel shifter" [2, 8] par laquelle les cellules subissent une rotation de manière aléatoire (voir la figure 4.2). En fait, cette méthode limite le nombre de codes aléatoires possibles à M au lieu de $M!$ ⁴. Cela peut être une solution simple si les erreurs des différentes cellules, c.-à-d. α_i , sont non corrélées. Malheureusement, ce n'est pas le cas et bien au contraire, dans la pratique, une bonne partie de ces erreurs sont souvent corrélées car elles ont les mêmes origines,

4. Si le générateur utilisé dans la figure 4.2 fonctionne comme un compteur numérique normal (*digital-counter*) et que les interconnexions des cellules changent d'une façon circulaire en décalant avec le nombre introduit par le générateur, une "Barrel-Shifter" de type non-aléatoire sera composée [2, 8, 116, 117]. La méthode barrel-shifter avec un compteur non aléatoire est particulièrement favorable à pratiquer dans le convertisseur analogique-numérique parallèle (flash), mais, elle présente un fort comportement tonal qui dégrade la performance du système. A fin d'empêcher la présence des tons parasites produits par la méthode barrel-shifter non-aléatoire dans un modulateur sigma-delta suréchantillonné, l'OSR du modulateur doit être assez grand. Par exemple pour un CNA de 5-bit, il faut que l'OSR soit supérieur à 500. Cette méthode n'est pas traitée ici. Des détails sont disponibles aux références : [8, 116, 117]

par exemple un gradient de température. Un bon compromis entre ces deux extrêmes, c.-à-d. nombre des codes aléatoire entre M et $M!$, est la méthode de commutation aléatoire en papillon ("*Butterfly Shuffler*"), dont un exemple est montré sur la figure 4.4 pour $M = 8$. Elle consiste en une série

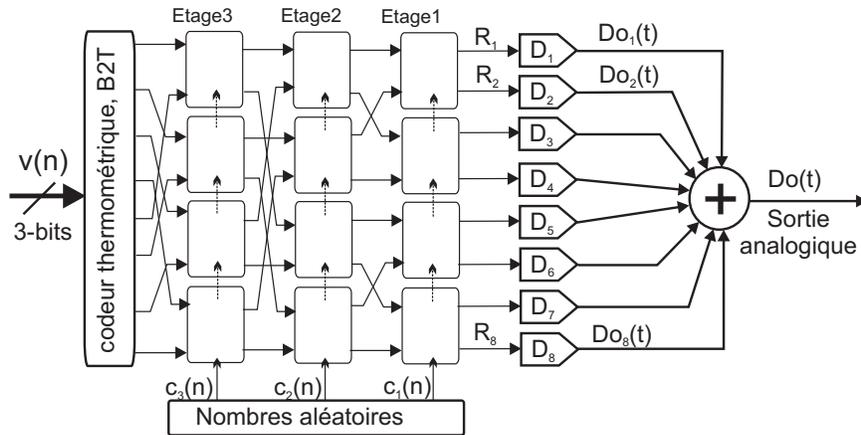


FIG. 4.4 – Méthode de commutation aléatoire quasi-RDEM en papillon ("*Butterfly Shuffler*")

de blocs similaires. Chaque bloc a deux entrées et deux sorties. Le couplage entre les sorties et les entrées de chaque boîte doit être changé de façon aléatoire, cependant un signal de commande d'un bit est suffisant. Le schéma d'un bloc papillon est décrit à la figure 4.5. Le nombre d'étages en série,

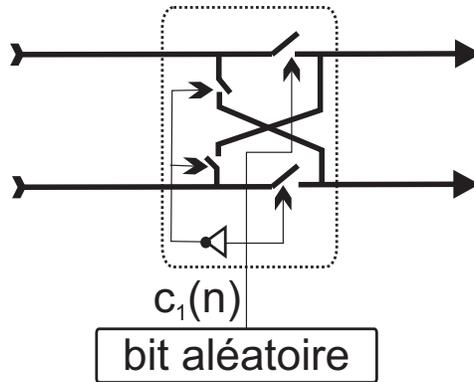


FIG. 4.5 – Schéma interne d'une boîte en papillon pour réaliser RDEM

c.-à-d. le nombre des colonnes à la figure 4.4, est au moins égal au nombre de bits de $v(n)$. Pour augmenter encore plus de capacité à blanchir les erreurs, on peut ajouter d'autres papillons en série. En revanche, quand le nombre d'étages augmente, le temps de propagation augmente aussi, ce qui limite la vitesse du système.

Par ailleurs, si M est trop grand, la méthode des papillons devient beaucoup plus complexe. Par exemple dans le cas d'un convertisseur numérique analogique classique 10 bits ou plus, le quasi RDEM n'est plus utilisable.

Récemment, un autre compromis nommé "*Segmented Randomiser DEM*" a été développé dans [1]. Dans cette méthode, les cellules du CNA sont attachées à des groupes séparés. Chaque groupe a sa

propre RDEM ou quasi RDEM. Pour coupler l'entrée aux différents groupes, on utilise un diviseur soit fixe soit aléatoire. Le reste de la division par deux de l'entrée peut être ajouté à l'une des deux sorties du diviseur de manière systématique ou aléatoire. un exemple de réalisation de la méthode RDEM segmentée est montré sur la figure 4.6. Une RDEM segmentée n'a pas la même performance qu'une

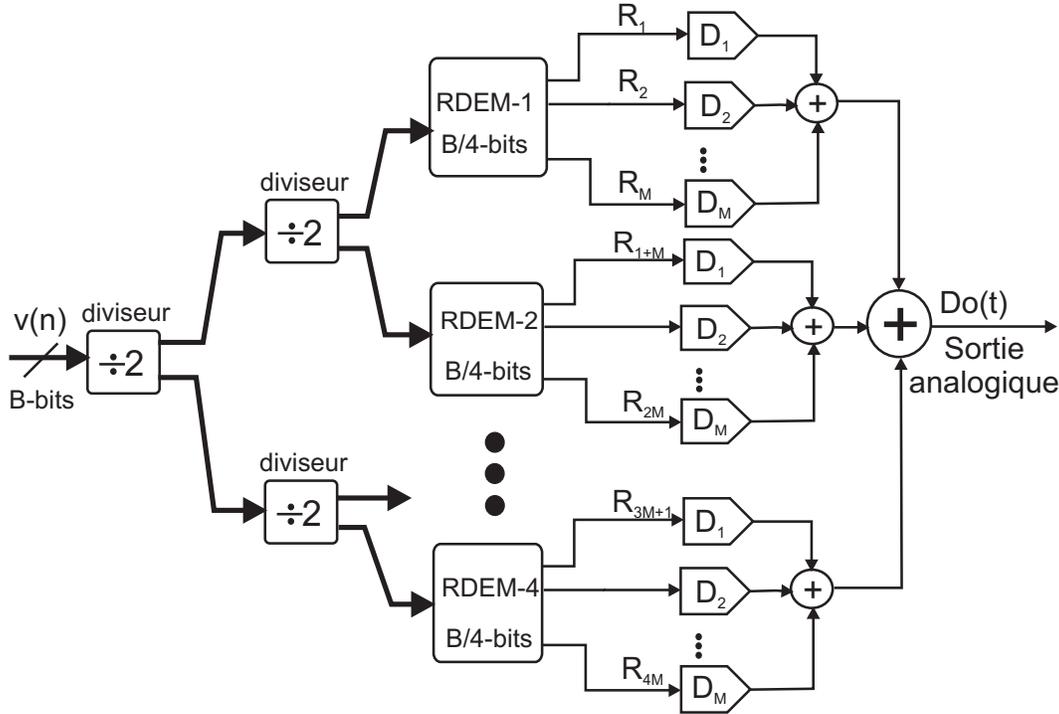


FIG. 4.6 – Schéma du brassage des sources par la méthode RDEM-segmentée, $M = 2^{\frac{B}{4}}$

RDEM idéale, mais elle présente l'avantage de réduire la surface du circuit sans que la performance en soit beaucoup diminuée pour des M grands.

Une autre difficulté commune à toutes les techniques de DEM y compris la RDEM, est que le nombre de commutations nécessaires augmente par rapport à une solution sans brassage. Autrement dit, le nombre des cellules qui doivent passer de l'état haut à l'état bas $\{Sv_i(n) = 1 \Rightarrow Sv_i(n+1) = 0\}$ ou inversement passer de l'état bas à l'état haut $\{Sv_i(n) = 0 \Rightarrow Sv_i(n+1) = 1\}$ est beaucoup plus élevé que dans le cas sans DEM. Chaque commutation ajoute des contraintes à la performance dynamique du système. Par exemple, le glitch et le temps de montée différents du temps de descente peuvent contribuer à la création d'erreurs dynamiques, ce qui limite le bénéfice réel de la DEM. Il y a trois solutions possibles dans le cas d'un brassage aléatoire :

- **RDEM en permettant seulement la commutation guidée "RDEM-guidée" :**

Dans cet algorithme, qui a été proposé dans [66], la commutation des sources demandée par la technique RDEM est limitée seulement à une direction pour chaque période. En outre, les nombres communs des sources entre les deux périodes, c.-à-d. $\min\{(M-v(n)), (M-v(n+1))\} = 0$ et $\min\{v(n), v(n+1)\} = 1$, restent à l'état précédent et les nouvelles commutations exigées par la

variation d'entrée, c.-à-d. $|v(n) - v(n + 1)|$, sont choisies en utilisant la méthode RDEM. Par conséquent, à chaque période, seules les nouvelles cellules qui étaient précédemment à l'état zéro sont activées, ou bien certaines cellules qui étaient à l'état "1" sont désactivées, mais pas en même temps dans la même période. Cela minimise efficacement le phénomène de glitch, mais au prix d'une RDEM moins performante.

- **DEM avec un nombre de commutation constant ou limité, "RDEM-modifiée" :**
une technique capable de réduire l'effet du nombre de commutations excessives des méthodes de DEM est obtenu par l'imposition d'un nombre de commutations constant à chaque période. Cet algorithme a été proposé pour la DEM en utilisant un autre algorithme que la RDEM [68], mais elle est aussi compatible avec la RDEM. Par cette méthode, l'erreur dynamique produite à chaque période est gardée quasiment fixe ce qui peut annuler l'effet dynamique des erreurs. En pratique, une telle contrainte, n'est pas facile à mettre en œuvre, donc, quelle que fois la dérivation de l'algorithme doit être autorisée.

- **utilisation des cellules de type RZ, "CNA-RZ" :**

Une méthode générale, qui rencontre un succès dans plusieurs applications [3, 4, 44, 47, 97, 118], emploie un ou deux CNA à retour à zéro, "*return-to-zero, RZ-DAC*". Dans cette technique, chaque cellule doit être remise à l'état zéro avant qu'elle soit utilisée à la période suivante. Par conséquent, l'effet des glitches et du temps de transition non idéal sont complètement annulés. Ainsi, le brassage de sources aléatoires peut être utilisé sans souci des effets secondaires produisant des erreurs dynamiques.

Cependant, la méthode de CNA-RZ a deux inconvénients;

- l'amplitude de la sortie de chaque cellule doit être augmentée (afin de maintenir constante la puissance moyenne de sa sortie), ce qui risque d'atteindre la limite de linéarité des circuits actifs des filtres,
- l'effet de la gigue d'horloge est aussi augmenté (elle intervient deux fois par période).

Le premier inconvénient peut être compensé par l'utilisation de deux CNA-RZ en alternance. Les deux inconvénients peuvent être compensés aussi par l'augmentation de la durée d'activité de chaque cellule, par exemple jusqu'à 75% ou même jusque 98% de la période. Des détails seront donnés au chapitre 6.

4.4 Appariement à moyenne d'horloge, (*Clock Averaging, CLK-DEM*)

B.H. Leung a proposé une méthode de DEM grâce à laquelle l'effet de la distorsion harmonique du CNA de rebouclage peut être écarté de la bande utile du signal d'entrée, plutôt que d'être blanchie [8]. Le schéma de cette méthode, qui est appelée "*Clocked Averaging, CLK-DEM*", est montré dans la figure 4.7 pour un CNA simple composé de deux cellules. Le principe est d'échanger le rôle de chaque moitié du CNA à vitesse élevée de sorte que la distorsion harmonique soit éloignée de la bande utile. Par exemple, dans le cas où les signaux de contrôle Φ_1 et Φ_2 sont complémentaires et

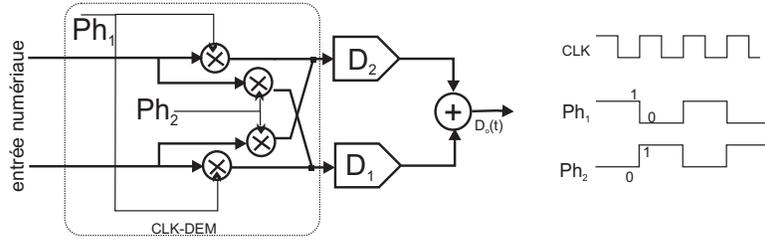


FIG. 4.7 – Schéma de l’algorithme Clock Averaging, CLK-DEM, pour deux cellules.

fixés à la moitié de la fréquence d’échantillonnage $\Phi_i = \frac{f_e}{2}$. L’extension de cette méthode devient alors plus complexe. L’exemple d’un CNA de 8-cellules est représenté sur la figure 4.8 où, les deux

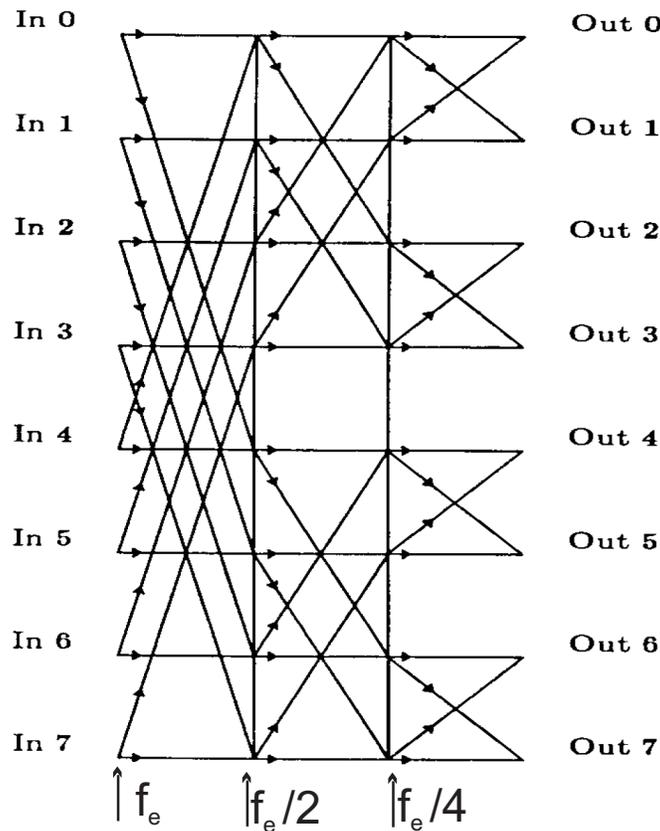


FIG. 4.8 – Schéma général de l’algorithme Clock Averaging, dans le cas de 8-cellules.

premières moitiés du CNA, ($\frac{M}{2}$ premières entrées et $\frac{M}{2}$ dernières entrées) doivent échanger leurs rôles à une cadence de $\frac{f_e}{2}$. De la même façon, les divisions suivantes ($\frac{M}{4}$ premiers et $\frac{M}{4}$ deuxième) doivent faire le même changement mais à une cadence de $\frac{f_e}{4}$. Cette façon de faire doit s’appliquer à toutes les divisions en progressant horizontalement et verticalement vers les différentes cellules. Par exemple, durant 8 périodes consécutives, le rôle de la cellule D_1 est changé de telle sorte que $sv_1 \equiv t_1, sv_1 \equiv t_5, sv_1 \equiv t_3, sv_1 \equiv t_7, sv_1 \equiv t_2, sv_1 \equiv t_6, sv_1 \equiv t_4$ et $sv_1 \equiv t_8$.

On constate que, par rapport à la méthode RDEM, la réalisation de l’algorithme CLK-DEM devient plus compliquée et peut générer des problèmes éventuels liés aux commutations. Cependant,

le problème majeur reste l'éventuelle réapparition de la distorsion harmonique du signal en bande utile dû au phénomène de modulation lors du changement de fréquence. Ce type de modulation fait apparaître des tonalités indésirables dans la bande passante surtout en cas de faible *OSR*. La figure 4.9 représente un exemple d'application de cet algorithme dans un modulateur multibit.

Bien que cette technique de réarrangement dynamique CLK-DEM soit une idée intéressante, elle n'a pas été suivie et a été rapidement remplacée par la méthode suivante, ILA-DEM.

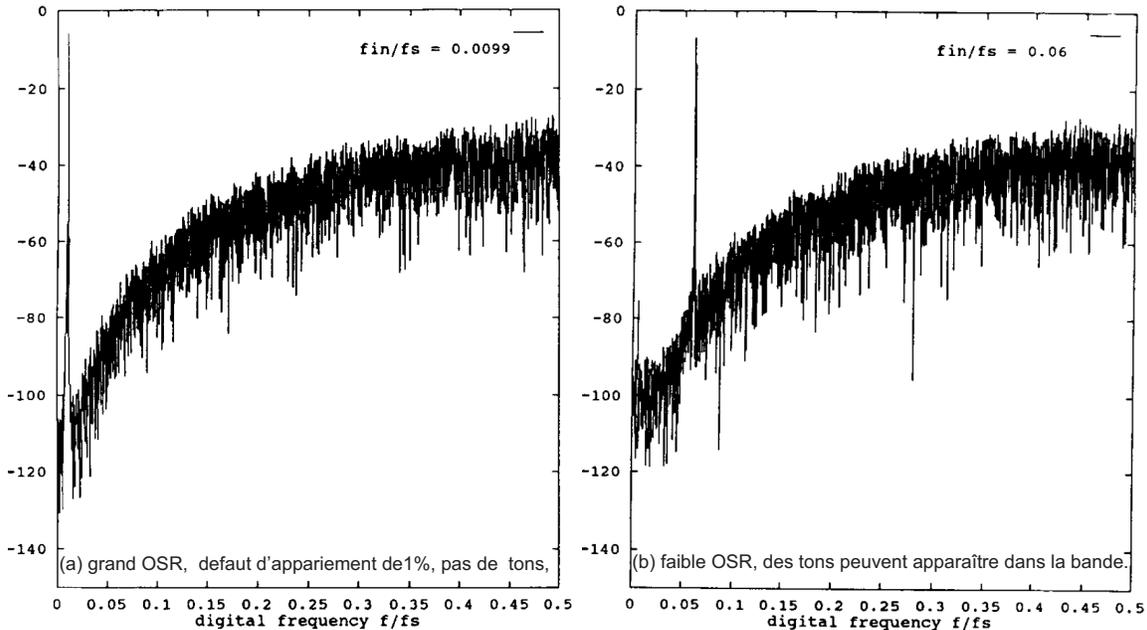


FIG. 4.9 — Spectre de la sortie du modulateur multibit passe-bas avec un défaut d'appariement du CNA de 1% et correction par la méthode CLK-DEM, pour deux différents OSR.

4.5 Brassage de source sur la moyenne des codes utilisés, "Individuel Level Averaging", ILA-DEM

Le point faible de la méthode CLK-DEM est de dépendre directement de la fréquence d'échantillonnage. Donc une meilleure solution doit être basée sur la minimisation de l'impact de la fréquence du système. Une solution possible, qui a été introduite par les mêmes auteurs, est nommée "méthode de réarrangement des éléments sur la moyenne des codes utilisés" ou simplement *ILA-DEM*, "Individual Level Averaging" [8,9]. Le principe est de faire en sorte que chaque cellule soit utilisée pour chacune des valeurs d'entrée avec une probabilité identique, c.-à-d. $P(sv_{i|x}) = \frac{1}{M}$. De cette manière, la somme des erreurs associées à chaque code d'entrée vaut zéro après M -périodes si l'entrée varie librement entre 0 et M , soit $\sum_{i=1}^M \alpha_i \equiv 0$ quand on néglige ϵ .

Afin d'indiquer le point de départ, chaque niveau d'entrée " $x = j$ " a besoin d'un registre de B -bits (nommée $R_j(n)$ pour le code j , $j \in [1, \dots, M]$).

Il existe deux approches permettant de réaliser l'algorithme ILA-DEM : circulaire et additionnelle. Dans le cas circulaire, $R_j(n)$ augmente d'une unité lorsque le CNA reçoit une nouvelle fois l'entrée

"x=j". Autrement dit, à l'instant $t = nT$ si un code "j" arrive, les cellules de numéro $R_j(n)$ au numéro $R_j(n) + j$ sont choisies, alors qu'à l'instant $t = (n + 1)T$ pour le même code, les cellules du numéro $R_j(n) + 1$ au numéro $R_j(n) + 1 + j$ seront choisies. Par conséquent, après M reprises du code "j" la circulation sera complète, les erreurs pourront donc être annulées.

Si on suppose que chaque code "j" arrive une fois tout les M -périodes, le temps nécessaire pour une telle annulation des erreurs du CNA est M^2 périodes d'échantillonnage ce qui est trop long. Ce temps d'annulation a un rôle important dans l'application des différentes méthodes de DEM. Il est nommé cycle d'annulation "CC, Cancellation Cycle". Il faut que le CC soit plus petit que l'OSR du système, si non, l'effet des erreurs sera présent dans la bande utile.

Dans la deuxième approche (additionnelle), on décale l'indice R_j de "j" pour chaque nouveau code "j". Afin de réduire le cycle d'annulation, on considère une deuxième approche. Cette fois-ci, à l'instant " $t = nT$ ", les cellules de l'indice (numéro) $R_j(n)$ à l'indice $R_j(n) + j$ sont choisies, alors qu'à l'instant $t = (n + 1)T$, les cellules du numéro $R_j(n) + j + 1$ au numéro $R_j(n) + j + j$ sont choisies pour deux réceptions consécutives du code "j". Il est évident que les positions des cellules sont toutes fois considérées circulaires, c'est-à-dire, qu'après l'élément de numéro M , on passe à l'élément de numéro 1. En ce qui concerne le cycle d'annulation, il est assez réduit pour des codes proches de $j = \frac{M}{2}$. Par ailleurs pour des codes éloignés, par exemple $j = 1$ ou $j = M$, on ne note pas d'amélioration. Le tableau 4.10 montre les éléments choisis pour une série de codes donnés, $x(n) \rightarrow x(n + 8) \equiv \{2,5,4,7,5,2,3,4\}$. Le spectre d'un modulateur passe-bas est représenté pour ces

Cellules du CNA 3-bit

		D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇	D ₈	
↓	Entrée	2	5	4	7	5	2	3	4	Rj(n)
		2	5	4	7	5	2	3	4	R2(n)=2
		5	5	4	7	5	2	3	4	R5(n)=5
		4	5	4	7	5	2	3	4	R4(n)=4
		7	5	4	7	5	2	3	4	R7(n)=7
		5	5	4	7	5	2	3	4	R5(n)=2
		2	5	4	7	5	2	3	4	R2(n)=4
		2	5	4	7	5	2	3	4	R2(n)=6
		4	5	4	7	5	2	3	4	R4(n)=8
		5	5	4	7	5	2	3	4	R5(n)=7
		⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

FIG. 4.10 – Tableau : l'approche additionnelle dans la méthode ILA-DEM, les éléments choisis pour une série des codes donnés, $x(n) \rightarrow x(n + 9) \equiv \{2,5,4,7,5,2,2,4,5\}$, quant $M=8$

deux approches ainsi que pour le cas CLK-DEM et pour le cas idéal, la figure 4.11 permet de faire les comparaisons. On peut noter que l'approche additionnelle apporte environ 10dB d'amélioration

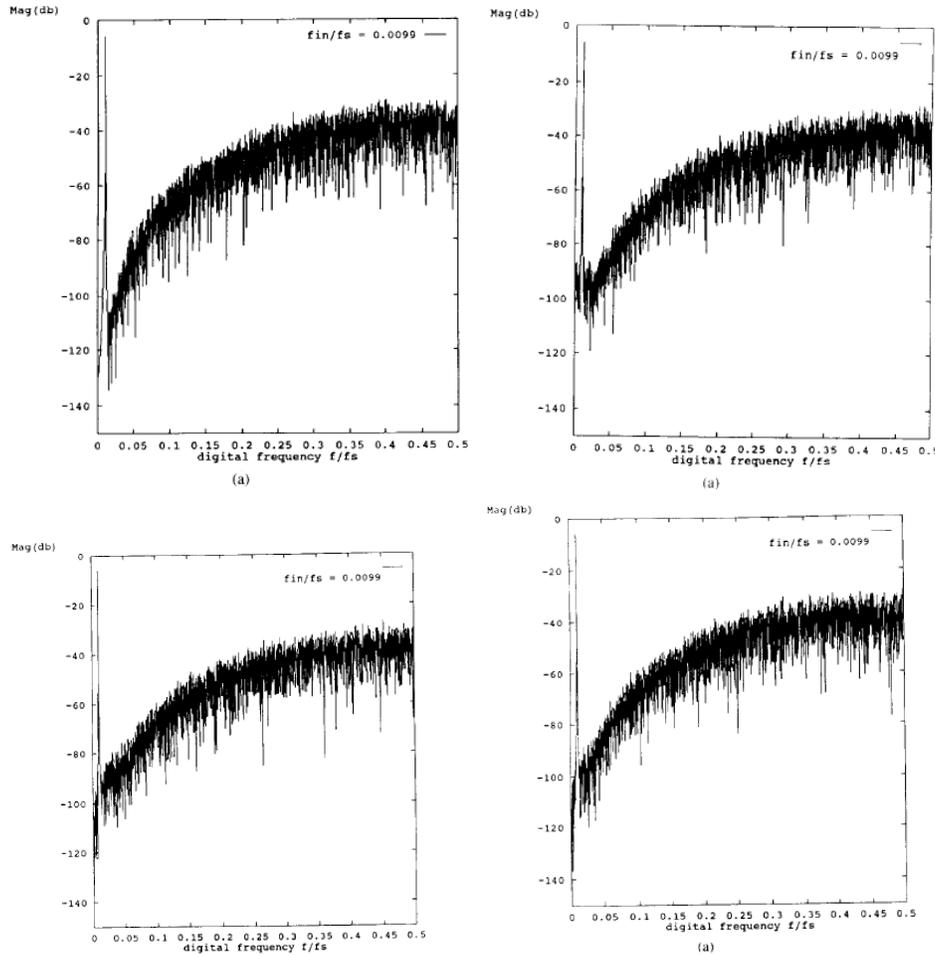


FIG. 4.11 – Spectre du modulateur passe-bas, a) idéal, b) méthode CLK-DEM, c) méthode ILA-DEM approche circulaire, d) méthode CLK-DEM, approche additionnelle, quand $M=8$

par rapport à l'approche circulaire. La méthode CLK-DEM ne peut pas atteindre la performance de la méthode ILA-DEM⁵.

Pour conclure cette section, nous pouvons remarquer que la méthode ILA-DEM n'est pas souvent utilisée dans les applications pratiques; elle est remplacée par d'autres propositions qui seront présentées à la section suivante. On peut ainsi résumer la technique ILA-DEM :

– **Avantages :**

- son mode d'exécution ne dépend plus de la fréquence d'échantillonnage contrairement à la méthode précédente CLK-DEM.
- elle peut mettre en forme l'erreur d'appariement par une fonction de transfert passe-haut (ce qui sert aux applications des modulateurs passe-bas).

– **Inconvénients :**

- elle nécessite plus de circuits que les méthodes précédentes,

5. Il me semble que la différence entre ces deux approches de la ILA-DEM a mal été traité dans certains documents comme par exemple dans [2]

- le cycle d'annulation est très long,
- elle ne convient pas aux applications passe-bande.

Cependant dans certains cas, on a montré que cette technique peut être intéressante, car elle offre un meilleur SFDR pour les systèmes à faible OSR [119]. La résolution maximale pouvant être atteinte par cette méthode ne présente pas une forme analytique simple. La simulation montre que la méthode ILA offre une résolution d'environ 12.5 bits pour le système avec : $M=16$, $\sigma_\alpha = 0.01$ et $OSR=64$.

4.6 Appariement dynamique sur la moyenne de données, DWA

Dans la section précédente, nous avons constaté qu'un cycle d'annulation long diminue la performance du brassage des sources. La méthode serait améliorée si l'annulation des erreurs était réalisée sur l'ensemble des codes utilisés et non pas séparément pour chacun des codes. Autrement dit, au lieu d'attendre une éventuelle circulation de chacun des codes d'entrée, la circulation est accompagnée à chaque période de n'importe quel code d'un seul indicateur $Ptr(n)$, "*Pointer*". Cette méthode se traduit par le brassage des sources sur la moyenne des données et est connue sous le nom de "*Data-Weighted-Averaging, DWA*", proposée en 1992 [69] puis en 1995 [10].

Le tableau 4.12 montre les éléments que l'application de la méthode DWA sélectionne pour la série de codes suivant, $x(n) \rightarrow x(n+8) \equiv \{2,5,4,7,5,2,3,4\}$ dans un CNA de 3-bit. Contrairement aux

Cellules du CNA, 3-bit

	$v(n)$	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	$Ptr(n)$
↓	2									3
↓	5									8
↓	4									4
↓	7									3
↓	5									8
↓	2									2
↓	3									5
↓	4									1
↓	⋮									⋮

FIG. 4.12 – Les éléments choisis pour une série des codes donnés, $x(n) \rightarrow x(n+8) \equiv \{2,5,4,7,5,2,3,4\}$ par la méthode DWA, quand $M=8$

méthodes précédentes, dans cette technique, pour différentes valeurs de code d'entrée les éléments sont sélectionnés de sorte que la moyenne du temps d'utilisation des différentes cellules soit identique. La gestion de ces moyennes est contrôlée directement par la séquence du code d'entrée. Cela est la raison pour laquelle, cette méthode a été nommée "Data-Weighted-Averaging". Elle nécessite un seul registre de B -bits $P_{tr}(n)$ afin de mémoriser l'indice de la première cellule qui devra être choisie dès que le système reçoit une entrée différente de zéro. Cette faible taille de circuit nécessaire est une particularité très intéressante et économique de la méthode DWA puisqu'elle est plus performante que les méthodes ILA-DEM, CLK-DEM et RDEM. En fait, on peut facilement montrer que la DWA met en forme les erreurs introduites par le CNA du fait de sa fonction de transfert passe-haut.

On suppose que le signal de contrôle de l'algorithme $P_{tr}(n)$ est un registre modulo M , ($1 \leq P_{tr}(n) \leq M$). Il doit être évalué par le code d'entrée $x(n)$ à chaque période de façon circulaire. C'est à dire, quand $P_{tr}(n-1) + x(n) < M$ le nouvel indicateur $P_{tr}(n)$ vaut $P_{tr}(n-1) + x(n)$, sinon il vaut $P_{tr}(n-1) + x(n) - M$, comme l'exprime la relation suivante :

$$P_{tr}(n) = [P_{tr}(n-1) + x(n)] \quad \text{modulo } M \quad (4.5)$$

Ce signal de contrôle est en quelque sorte l'intégrale du signal d'entrée $x(n)$. Bien sûr, le registre $P_{tr}(n)$ débordera ("*overflow*") de temps en temps, mais cela n'a pas d'importance car nous nous intéressons seulement à la différence entre deux échantillons consécutifs. Comme l'entrée $x(n)$ doit logiquement posséder une valeur entre "0" et "M", il n'y aura pas dépassement dans une période. La sortie analogique du CNA est alors donnée par la relation suivante :

$$D_o(t) = [P_{tr}(n) - P_{tr}(n-1)](1 + \bar{\alpha}) + e(n) \quad (4.6)$$

où $e(n)$ est l'erreur de défaut d'appariement qui s'est ajoutée à la période numéro " n ", et $\bar{\alpha}$ est le gain moyen du CNA. En partant des équations 3.52 et 3.53, la valeur $e(n)$ peut être estimée en terme d'erreur de linéarité intégrale :

$$e(n) = INL[P_{tr}(n) - 1] - INL[P_{tr}(n-1)] \quad (4.7)$$

où la fonction " $INL(p)$ " est définie pour les cellules de "1" au " p ", par la relation suivante :

$$INL[p] = \sum_{i=1}^p \alpha_i \quad (4.8)$$

En utilisant l'équation 4.8, on peut également décrire le terme " e " en utilisant la transformée en " z " :

$$E_{CNA}(z) = (1 - z^{-1})INL[P_{tr}(z)] \quad (4.9)$$

Maintenant, il devient plus clair que l'erreur de défaut d'appariement du CNA a été mise en forme par la méthode DWA-DEM à la manière d'un filtre passe-haut. Théoriquement, c'est, d'une part, un fonctionnement identique à celui d'un modulateur passe-bas du premier ordre vis à vis du bruit de quantification. D'autre part, il est probable que la DWA-DEM présente un comportement tonal non

désirable qui est le point faible du modulateur du premier ordre.

Malheureusement, cet inconvénient peut être un grand obstacle pour la technique DWA, quand le CNA reçoit une entrée fixe ou périodique ce qui est souvent le cas dans des applications telles que l'audio, la vidéo, les mesures.⁶

En ce qui concerne la résolution totale du système, idéalement l'erreur du CNA peut être estimée en utilisant l'équation 2.27 pour σ_α au lieu de q_{rms} et en supposant $L = 1$:

$$P_{e_{DWA, dans la bande}} = \frac{P_{e_{DWA}} \pi^2}{3 OSR^3} \quad (4.10)$$

où la puissance d'erreur dans le cas d'une entrée aléatoire entre 0 et M est exprimée par la relation suivante [88,89] :

$$P_{e_{DWA}} = \frac{\sigma_\alpha^2 (1 - \frac{1}{M})^2 \cdot M}{12} \quad (4.11)$$

Alors, la résolution maximale peut être estimée comme suit :

$$ENOB_{DWA} \cong \text{Log}_2 \left[\frac{\sqrt{3 M \cdot OSR^3}}{\pi \cdot \sigma_\alpha \cdot (1 - \frac{1}{M})} \right] \text{ bits} \quad (4.12)$$

Si nous supposons que l'erreur de défaut d'appariement est dominante, la résolution maximale avec une DWA-idéal, $\sigma_\alpha = 1\%$, $M = 16$ et $OSR = 64$, on obtient : $ENOB_{max} \simeq 16.7$ bits. Évidemment, les relations ci-dessus sont valides seulement pour le cas DWA idéal, c.-à-d. le cas où ses entrées ne sont ni fixes ni périodiques et sont distribuées uniformément entre 0 et M . Dans le cas pratique cela n'est pas toujours vrai, donc, des tons apparaissent et dégradent la performance du système. La figure 4.13 montre trois situations différentes pour l'application de la méthode DWA avec trois entrées différentes.

La répétition d'une sélection des éléments du CNA de manière cyclique se transforme en tons dont les amplitudes dépendent des caractéristiques stochastiques du défaut d'appariement des cellules. Les tons apparaissent sur l'axe des fréquences $[0, \frac{f_e}{2}]$. Leurs lieux sont directement liés à la périodicité éventuelle de la valeur de $Ptr(n)$. Pour une entrée fixe " j ", ces tons se trouvent aux fréquences données par la relation suivante :

$$f_{tons} = k \frac{GCD}{M} f_e, \quad k = 1, 2, 3, \dots \quad (4.13)$$

où, le terme GCD est le plus grand diviseur commun de l'entrée j , qui est supposée périodique et de nombre des cellules M . Ces tons peuvent moduler le signal d'entrée et revenir dans la bande utile. En général, pour un CNA de M cellules, la plus petite fréquence tonale produite par cette relation est $\frac{f_e}{M}$, c.-à-d. $k = 1$, $GCD = 1$. La fréquence maximale de l'entrée peut être $f_{BW} = \frac{f_e}{2 \cdot OSR}$. Ainsi,

6. Considérons une entrée fixe " $x = j$ ", la séquence périodique est la suivant :

$$\left\{ \sum_{i=1}^j \alpha_i, \sum_{(i=j+1)}^{2j} \alpha_i, \dots, \sum_{(i=M-j+1)}^M \alpha_i, \sum_{i=1}^j \alpha_i, \dots \right\}$$

où, $\frac{M}{r \cdot f_e}$ est la période, r est le diviseur commun de M et j .

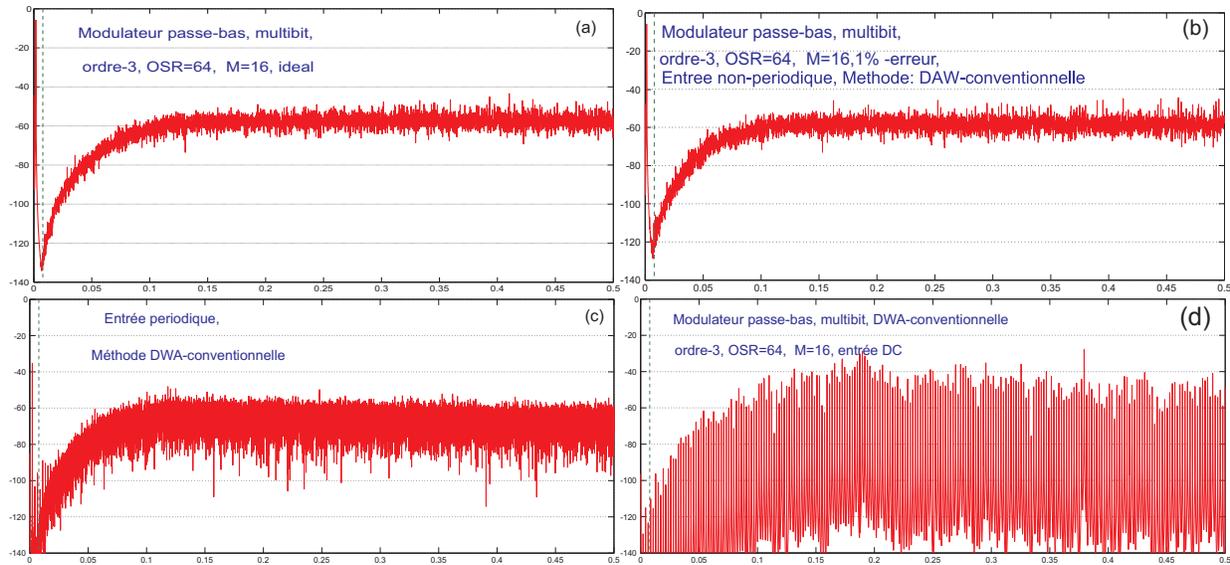


FIG. 4.13 – Spectre du modulateur passe-bas multibit ($M=16$) et $\sigma_\alpha = 1\%$: a) idéal, b) méthode DWA et l'entrée non périodique, c) méthode DWA avec une entrée périodique, d) méthode DWA avec une entrée DC.

les tons modulés par l'entrée ou l'une de ses harmoniques ne seront pas en bande utile si la condition suivante est validée.

$$\frac{f_e}{M} - k \cdot f_B > f_B \Leftrightarrow OSR > \left(\frac{k+1}{2}\right)M \quad (4.14)$$

Par exemple, pour un CNA de 16-cellules, l'OSR du système doit être plus grand que - {16,24,32,40,...} pour au moins les 5 premiers harmoniques, ce qui va limiter la généralité de la méthode DWA.

Un autre problème de la méthode DWA est que sa réalisation présentée ci-dessus est utilisable pour des applications passe-bas, elle ne supporte donc pas directement les applications passe-bande ou passe-haut.

On constate que la méthode DWA a besoin de quelques changements et adaptations complémentaires pour mieux fonctionner dans le cas pratique, c.-à-d. pour une entrée arbitraire.

Avant d'expliquer les différentes solutions proposées pour empêcher le comportement tonal du DWA, nous présentons un schéma bloc de la DWA-DEM de base. Une DWA de conception simple est représentée à travers la figure 4.14, dans laquelle l'entrée est supposée binaire. Comme on l'a dit, l'algorithme de DWA est toujours préférable pour les modulateurs multibit passe-bas car il peut être implanté avec un minimum de logique par rapport aux autres méthodes.

4.6.1 Solutions proposées pour l'empêcher les tons

Depuis l'introduction en 1995 de la technique DWA, au moins 9 solutions différentes ont été proposées pour résoudre le problème liée au comportement tonal :

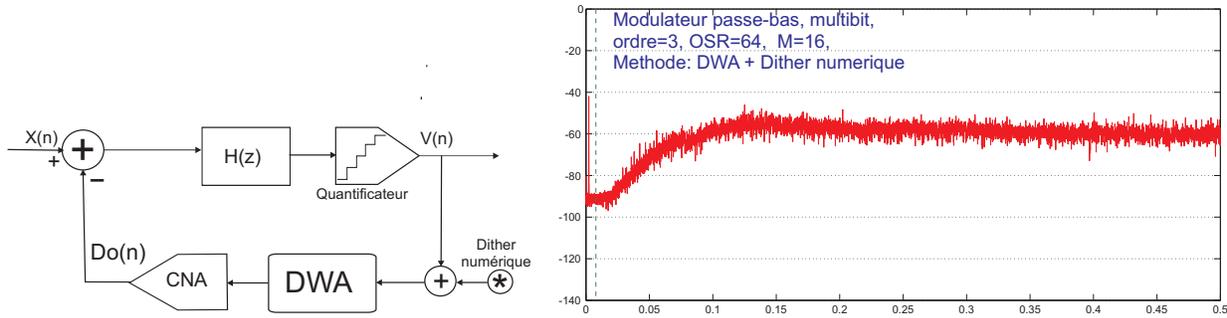


FIG. 4.15 – Prévention de tons produit par la DWA dans un modulateur multibit à l'aide d'un dither, b) Le spectre du modulateur passe-bas multibit avec $M=16$, $|Dither|_{max} = LSB$, $std(Dither) \simeq 0.01LSB$ et l'entrée périodique.

2- Technique de rotation des éléments, RDWA

Une autre méthode nommée "Rotated Data Weighted Averaging", RDWA, qui a déjà été utilisée pour blanchir les tons produits par une éventuelle séquence cyclique, est basée sur le principe de la perturbation consécutive à l'évaluation de $Ptr(n)$ [120]. Le principe du RDWA est celui du DWA, sauf qu'il partage quelques cycles différents dans le temps. Chaque cycle, qui est choisi une fois pour toute puis enregistré dans une mémoire "ROM", indique la façon dont $Ptr(n)$ est évalué. La figure 4.16 montre le diagramme de quelques cycles pour un CNA de 8-cellules. Le cycle numéro "C1" est celui de DWA classique puis les autres "C2, C3,...,C5" sont arbitrairement choisis. En effet, un CNA de M -cellules ($M = 2^B$ bit), a $(M - 1)!$ cycles différents. Par exemple, un CNA 3-bit ($M = 8$) a 5040 cycles. Cependant, un nombre de cycles si important n'est pas nécessaire pour le RDWA. L'expérience

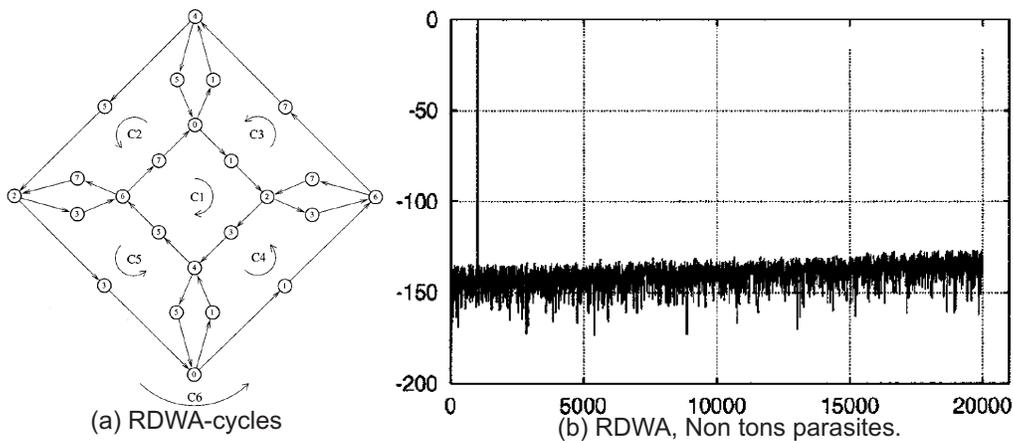


FIG. 4.16 – Quelques exemples de cycles utilisés pour l'algorithme de RDWA, b) le spectre d'un CNA multibit [120].

montre qu'elle n'en a besoin que de quelques uns, par exemple, de " M " différents cycles. L'indicateur $Ptr(n)$ suit chaque cycle durant un nombre limité des périodes. Ensuite, l'évaluation de $Ptr(n)$ passe à un autre cycle et ainsi de suite. Le temps entre deux changements consécutifs de cycles, qui est nommé MTBS (Mean Time Between Switching), a un rôle important dans la détermination

de la performance. Parce que, d'une part, un MTBS court fait augmenter le niveau de bruit du système [120]. Autrement dit, le principe de DWA n'est pas respecté, par conséquent, la RDWA avec un MTBS trop court s'approche de la méthode RDEM. En revanche, un MTBS trop grand n'est pas capable de blanchir complètement les tons, donc la RDWA avec un MTBS assez grand ressemble à la méthode DWA traditionnelle. Un MTBS raisonnable peut être choisi de l'ordre du quart de l'OSR [120]. Cette solution nécessite une logique additionnelle considérable, et les publications à ce sujet montrent que cette technique n'est pas la meilleure [88].

3- Technique aléatoire, RnDWA

Une méthode plus complète que la technique précédente (RDWA) est présentée dans la référence [88]. Cette méthode nommée RnDWA, "*Randomized DWA*", utilise à la fois le principe de DWA (c.-à-d. l'équation 4.6), et le principe de la méthode RDEM. Le principe de DWA fait qu'aucune cellule n'est utilisée tant qu'une circulation complète de toutes les cellules du CNA n'a pas été effectuée. Cependant, pour compléter chaque circulation plusieurs choix sont possibles. Dans chaque période, l'algorithme RnDWA sélectionne le nombre de cellules nécessaires parmi des éléments qui n'ont pas été choisis dans ce cycle, d'une manière aléatoire. Par exemple pour une séquence d'entrée, l'un des choix possibles est montré sur la figure 4.17 en comparaison avec la méthode DWA. Bien que

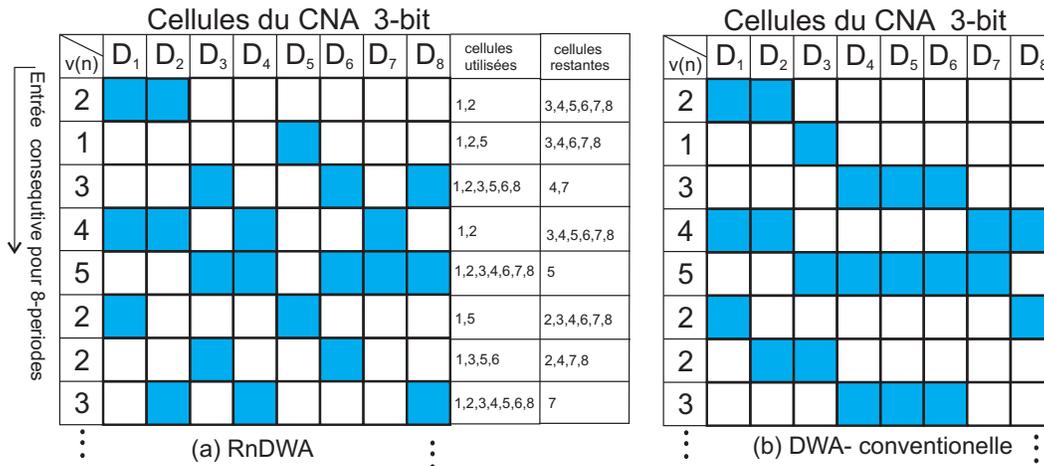


FIG. 4.17 – Cellules sélectionnées pour la séquence donnée {2,1,3,4,5,2,2,3,...}, quand M=8, a) algorithme de RnDWA, b) méthode de DWA classique.

l'algorithme RnDWA idéal puisse complètement empêcher le comportement tonal, il n'est pas facile à implanter au niveau des circuits.

Afin de simplifier la réalisation des circuits de RnDWA, on peut commencer par une cellule, puis suivre la méthode de la circulation normale de DWA jusqu'à ce que l'indicateur revienne à la même cellule, c.-à-d. après une ou plusieurs circulations complètes et régulières. On remplace alors, l'indicateur $Ptr(n)$ par une valeur aléatoire pour faire d'autres circulations dans le cycle du " $Ptr(n)$ " au " $Ptr(n) - 1$ ". Selon [88] où l'analyse de ce principe est détaillée, cette méthode fonctionne mieux que les méthodes précédentes. Cette technique n'est adaptée que pour les modulateurs passe-bas en gardant la mise en

forme du premier ordre de $(1 - z^{-1})$ pour l'erreur $e(n)$, (voir la figure 4.18).

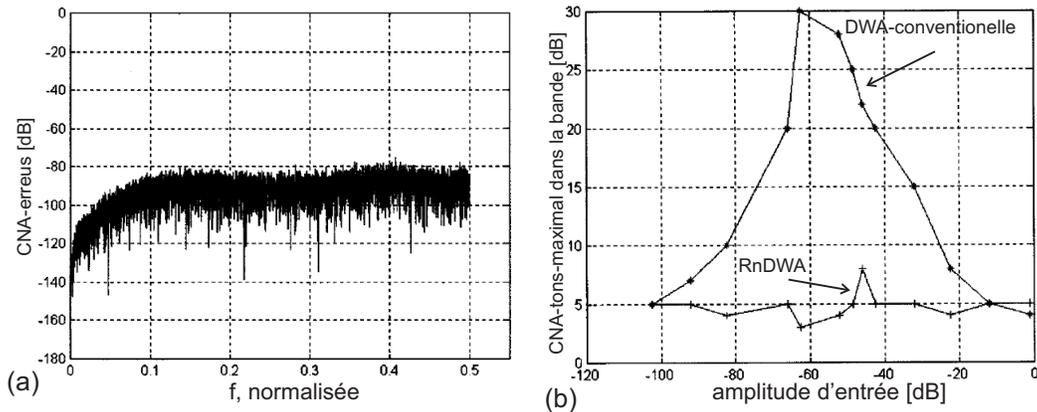


FIG. 4.18 – Spectre de la sortie d'un modulateur multibit ayant l'algorithme RnDWA, b) l'amplitude maximale des tons dans la bande obtenue par la méthode RnDWA est comparée avec celle de DWA classique.

4- Technique de double indice, DIA

Un nouvel algorithme pour éliminer le comportement tonal de DWA-DEM a été présenté dans la référence [121]. Il est adapté plus particulièrement aux CNA différentiels, dont un exemple est représenté sur la figure 4.19, avec $M=8$. Cette structure fondée sur la technique des capacités commutées "SC", utilise $\frac{M}{2}$ capacités ($c_i, i \in [1, 2, \dots, \frac{M}{2}]$), au lieu des M capacités dans chaque branche. Cependant, chaque cycle c_i a la possibilité de se charger à une valeur négative, nulle ou une valeur positive, c.-à-d. $[-V_R, 0, +V_R]$. On applique le principe de DWA mais avec deux indicateurs, un pour

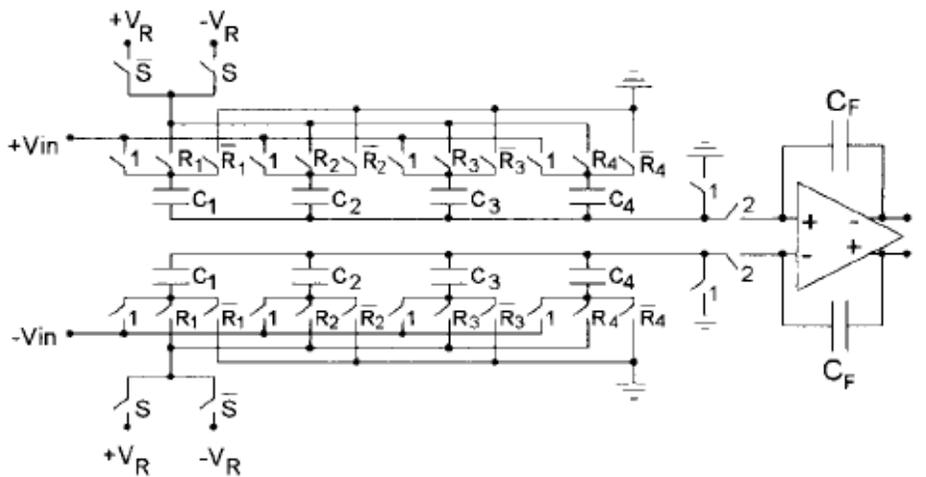


FIG. 4.19 – Structure du CNA entièrement différentiel à l'entrée d'un intégrateur SC, où la méthode de DIA peut être appliquée.

les entrées négatives (moitié inférieur $x < 0$) et un autre pour les entrée positives (moitié supérieur

$x \geq 0$). Cette méthode est nommée *DIA*, *Double Index Averaging* et peut être considérée comme un bon compromis entre les méthodes DWA idéal et ILA. La méthode DIA résout le comportement tonal mais n'arrivera jamais à atteindre la performance d'une DWA idéal. La figure 4.20 de la référence [121] donne des éléments de comparaison. Il faut noter que cette technique ne peut pas être

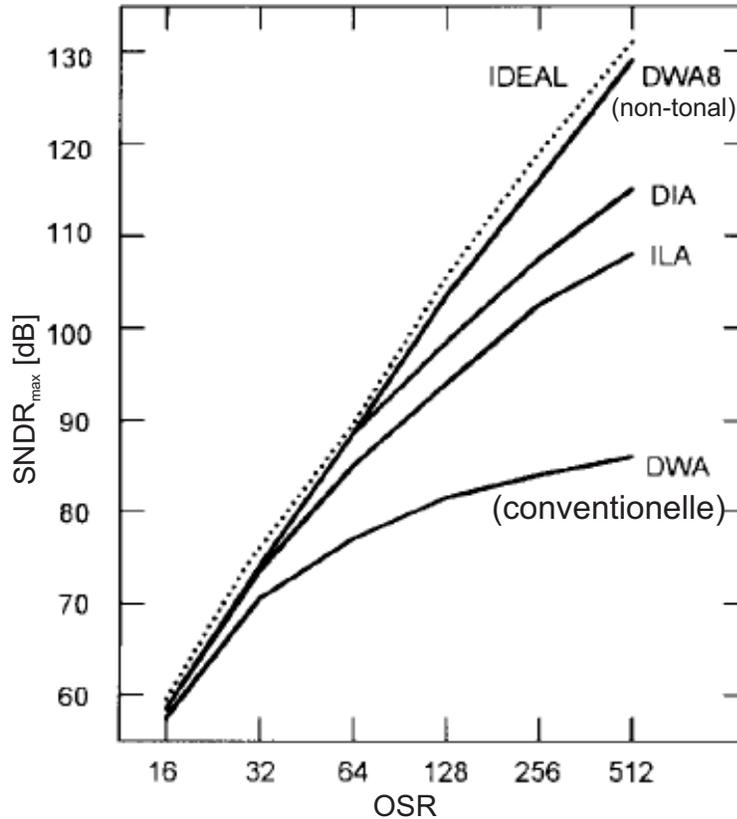


FIG. 4.20 – Résultats des simulations rapportés par [121]; SNDR maximum en fonction de l'OSR.

appliquée aux CNA à cellules de courant ("Current steering DAC"), car les cellules de courant fonctionnent uniquement dans une seule direction $[0, I_u]$. Dans ce cas, on doit diviser le CNA en deux CNA avec deux indicateurs séparés.

5- Technique à accroissement, IDWA

Comme nous l'avons dit précédemment, l'algorithme de DWA conventionnel peut produire des tons indésirables aux fréquences données par l'équation 4.15, $f_{tons} = k \frac{GCD}{M} f_e$, $k = \{1, 2, 3, \dots\}$. Dans le cas d'un *OSR* faible, ces tons apparaissent en bande passante à cause de la modulation par la fréquence d'entrée. Une méthode pour éloigner ces tons modulés est proposée aux références [122, 123], où le nombre des cellules du CNA est plus grand que M . Par exemple, un CNA 4-bits utilise 17 cellules au lieu de 16 dans le cas normal. Ainsi, en appliquant l'algorithme DWA pour un tel CNA, ayant " $M+m$ " éléments, les fréquences des éventuels tons sont éloignées, selon la relation suivante :

$$f_{tons, IDWA} = k \frac{GCD}{M+m} f_e, \quad k = 1, 2, 3, \dots, \quad (4.15)$$

où, GCD est le plus grand diviseur commun entre l'entrée " x " et le nombre des cellules " $M + m$ ". A cause de l'accroissement du nombre des cellules, cette méthode est nommée *IDWA, Incremental Data Weighted Averaging*. La figure 4.21 montre le spectre de sortie du modulateur pour $m = \{0,1,3\}$. On

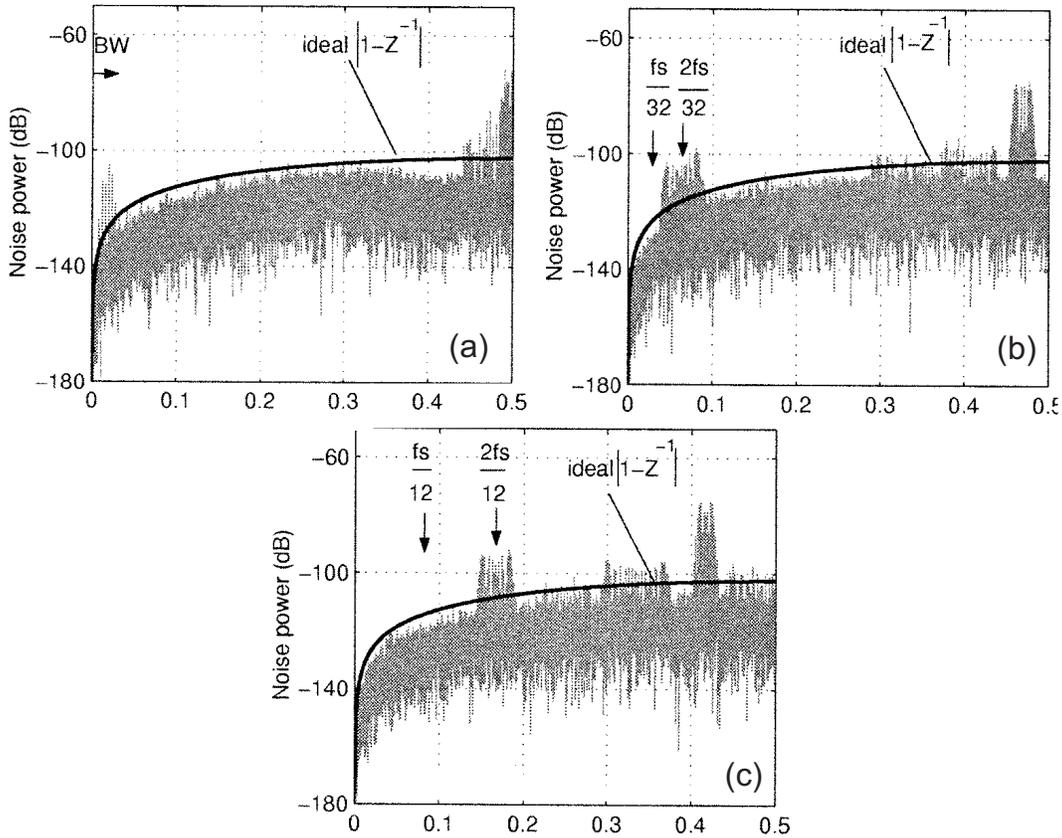


FIG. 4.21 – Spectre de la sortie du modulateur multibit [123]: a) DWA conventionnelle $M=16$ et $m = 0$, b) IDWA avec $M=16$ et $m = 1$, c) IDWA avec $M=16$ et $m = 3$.

voit bien que les tons sont éloignés de la bande utile. En pratique, une ou deux cellules supplémentaires suffisent.

Même si la méthode IDWA fonctionne assez bien pour la plupart des valeurs d'entrée, et surtout les entrées proche de la mi-échelle, elle a encore du mal à blanchir les tons produits par les entrées très faibles (proche de $x = \frac{1}{M+m}$). Une réalisation pratique d'IDWA est présentée dans la référence [123]. Cependant, la condition de l'*OSR* doit être respectée, soit $OSR > \frac{1}{2}(M + m)$, ce qui en définitive n'apporte pas grand-chose, $\frac{1}{M}$ amélioration relative par chaque cellule supplémentaire.

6- Technique bi-directionnelle, BIDWA

Afin d'empêcher l'apparition des tons produites par la méthode DWA conventionnelle, une stratégie complémentaire consiste à perturber la périodicité de l'algorithme en utilisant deux indicateurs de DWA de deux directions opposées. Cette méthode est nommée "*BIDWA, Bi-directional Data weighted averaging*" [124]. Le principe de BIDWA est décrit sur le tableau 4.22 : deux indicateurs alternés mémorisent le rang de l'élément qui doit être sélectionné au prochain cycle. Par conséquent, le cycle

d'annulation de la DEM est deux fois plus long que dans le cas du DWA conventionnel, ce qui altère sa performance et son efficacité surtout pour les applications de faible *OSR*. Cependant, elle offre un meilleur SFDR que d'autres méthodes similaires, par exemple DIA, c.-à-d. moins de tons dans la bande du signal [119].

La technique BIDWA a d'abord été proposée pour un système passe-bas, sans grand succès. Cette méthode convient cependant pour une application passe-bande. Une méthode de DEM passe-bande d'ordre deux est envisageable sur ce principe ; elle sera proposée au prochain chapitre.

Cellules du CNA, 3-bit

	v(n)	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇	D ₈	Ptrj(n)
↓	2									Ptr1(n)=2
	5									Ptr2(n)=4
	4									Ptr1(n)=6
	7									Ptr2(n)=5
	5									Ptr1(n)=3
	2									Ptr2(n)=3
	2									Ptr1(n)=5
	4									Ptr2(n)=7
	5									Ptr1(n)=2
	⋮									⋮

Entrée consecutive pour 9-periodes

FIG. 4.22 – a) Une sélection possible par la méthode BIDWA pour $x=\{2,5,4,7,5,2,2,4,5,\dots\}$, $M=8$

7- Technique DWA segmentée, P-DWA

Si on accepte de sacrifier un peu de la performance de la DWA afin de d'empêcher le comportement tonal, l'une des meilleures techniques est l'utilisation de deux demi DWA [76]. Dans cette méthode appelée DWA segmentée "*P-DWA, Partitioned DWA*", l'entrée du CNA est divisée par 2 et chaque partie utilise son propre DWA, la circulation s'effectuant entre $\frac{M}{2}$ cellules. Quand l'entrée est impaire, le reste de la division peut être ajouté à l'une ou l'autre des deux parties. La figure 4.23 montre la simplicité de la structure P-DWA. Cette technique présente une petite difficulté pour obtenir un fonctionnement symétrique, car les deux parties de DWA ne sont pas excitées de la même façon. Une partie reçoit toujours le bit avec le plus faible poids de l'entrée $LSB_{x(n)}$, c.-à-d. $(\frac{x(n)-LSB_{x(n)}}{2} + LSB)$, l'autre ne reçoit que $(\frac{x(n)-LSB_{x(n)}}{2})$. Autrement dit, la technique ne respecte pas les conditions nécessaires de DEM comme cela est analysé dans [125]. La contribution de cette erreur vaut :

$$e_{P-DWA}(n) = \frac{LSB_{x(n)}}{2} \cdot \left(\frac{2}{M} \sum_{i=1}^{\frac{M}{2}} \alpha_i - \frac{2}{M} \sum_{i=\frac{M}{2}+1}^M \alpha_i \right) \quad (4.16)$$

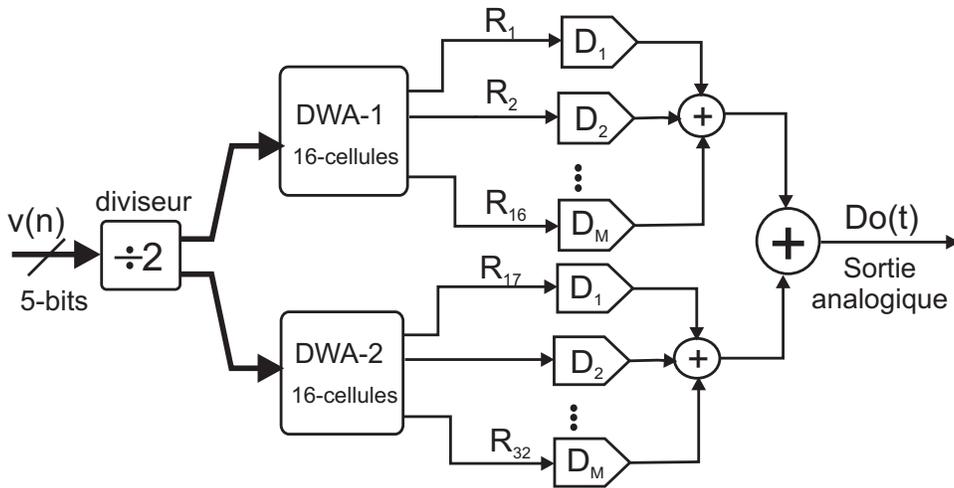


FIG. 4.23 – a) Structure simple proposée pour la méthode P-DWA pour un CNA du 4-bits divisée en deux moitiés de 3-bit, chaque moitié utilise sa propre DWA conventionnelle

où le bit LSB est supposé être appliqué à une des deux moitiés. On peut espérer que la contribution de ce terme sera faible, car, d’une part, la valeur du LSB est $\frac{M}{2}$ fois plus petit que la moyenne de x , et d’autre part, l’opération de sommation des erreurs de défaut d’appariement sur le nombre de $\frac{M}{2}$ des erreur stochastiques α_i diminue normalement leur effet d’environ $\sqrt{\frac{1}{M}}$. Les résultats des simulations confirment ce raisonnement. Cependant, ils montrent que la méthode P-DWA ne donne pas une fonction de transfert passe-haut parfaite, puisque le spectre d’erreur n’est pas vraiment égal à zéro à la fréquence " $f = 0$ ". Ils confirment par contre l’efficacité de la P-DWA à blanchir les tons générés par la DWA classique avec un taux d’erreur négligeable. La figure 4.24 montre un exemple dans lequel le spectre d’erreur du CNA ($Do(n) - x(n)$) n’atteint pas zéro. Il ne serait nul ($-\infty$ en dB) que dans le cas idéal. Nous allons proposer une autre approche similaire à la P-DWA, au chapitre

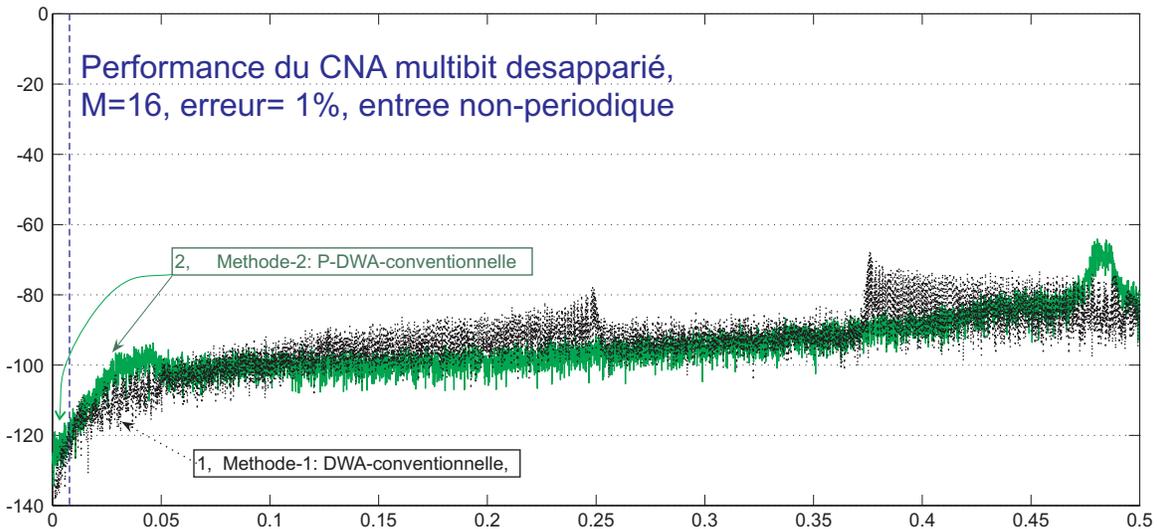


FIG. 4.24 – un exemple d’application passe-bas de la méthode P-DWA où le spectre d’erreur du CNA ($Do(n) - x(n)$) n’atteint pas zéro à la fréquence proche au DC ($f = 0$).

suivant.

8- Technique DWA partielle

Le signal d'entrée d'un système audio-vidéo présente normalement une caractéristique Gaussienne. Cela signifie que les codes très proches des deux extrémités, $x = 0$ ou $x = M$, ont une probabilité d'apparition plutôt faible. Si M est assez grand ($M > 16$), on peut alors utiliser une DWA partielle. Par exemple, au lieu de faire circuler l'indicateur $Ptr(n)$ modulo M , l'algorithme le fait circuler modulo $\frac{4M}{5}$. Par cette approche, deux effets intéressants sont obtenus :

- les tons ne sont pas générés, car de temps en temps, la circulation est perturbée par l'activité des cellules en dehors de l'algorithme DWA partiel.
- la circulation s'effectue plus rapidement que dans le cas normal, donc l'erreur due au défaut d'appariement peut être annulée plus vite et plus efficacement.

Un exemple de réalisation est rapporté dans la référence [84] (voir la figure 4.25). Même si pour la

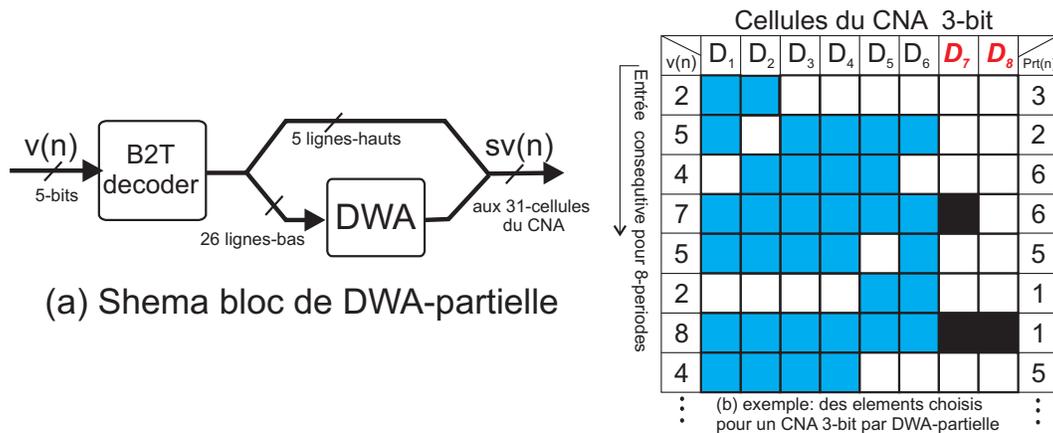


FIG. 4.25 – Méthode de DWA-partielle : a) le schéma bloc, b) un exemple de sélection des cellules pour un serai d'entrée

plupart des entrées, la DWA-partielle donne un meilleur résultat que l'algorithme DWA original, elle ne peut pas fonctionner pour une entrée arbitraire, par exemple $x \simeq \frac{4M}{5.k}$ quand $k \in \{2,3,4..\}$, car l'équation 4.14 devient plus limitative. Sans vérification plus précise, on ne peut cependant négliger les risques de cette technique dans les applications pratiques. On peut constater que cette solution est en quelque sorte à l'opposé de la technique IDWA.

9- Technique Pseudo-DWA

Afin de simplifier les circuits proposés pour empêcher le comportement tonal de DWA, une autre idée, toujours basée sur la perturbation de la périodicité de la circulation entre les cellules du CNA, consiste à sauter de temps en temps une unité dans la valeur de $Ptr(n+1)$ vers l'avant ou vers l'arrière selon le LSB de l'entrée $x(n)$ [126]. En fait, on inverse le LSB de l'entrée avant qu'il ne soit appliqué à l'additionneur numérique dans l'indicateur $Ptr(n)$. La figure 4.14 montre un exemple de

la réalisation de ces circuits logiques pour $M = 32$. Par conséquent, en présence du signal de contrôle d'inversion $Com_{inv}(n) = 1$, si l'entrée est impaire ($LSB = 1$), $Ptr(n+1)$ vaut $[Ptr(n) + x(n) - 1]$ et si l'entrée est paire ($LSB = 0$), $Ptr(n+1)$ vaut $[Ptr(n) + x(n) + 1]$. L'optimisation de la période de la commande d'inversion dans la technique de Pseudo-DWA ne peut être obtenue au moyen d'un procédé analytique, mais par simulation. Si la période du signal de commande $Com_{inv}(n)$ est très longue, les tons ne sont pas blanchis ce qui rapproche cette technique de celle de la DWA conventionnelle (cas $f_{Com_{inv}} = \infty$). Par contre, si l'inversion est très fréquente, la fonction de DWA peut être perdue, ce qui rapproche de la méthode RDEM. Un bon compromis consiste à choisir $f_{Com_{inv}} = \frac{4 * f_c}{OSR}$ selon les prévisions de l'article [126].

Cet algorithme ne peut pas complètement résoudre le problème des tons dans le DWA, mais offre un

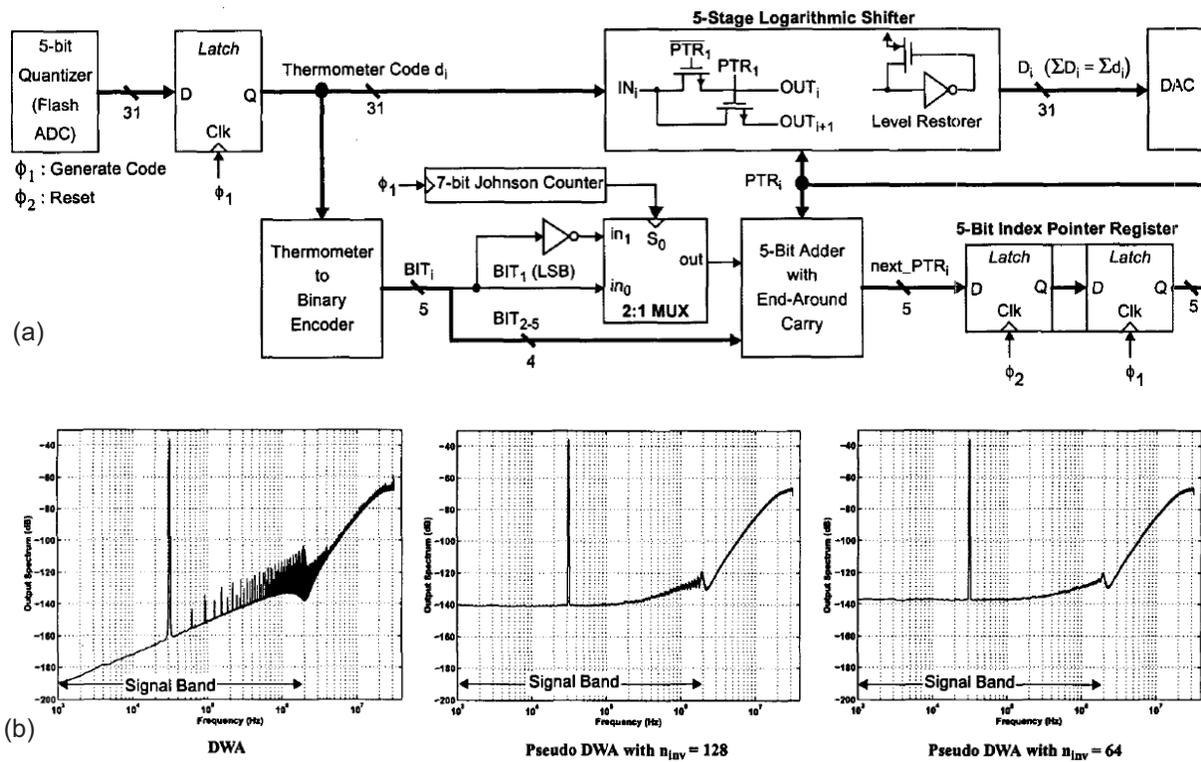


FIG. 4.26 – a) Une structure simple proposée pour la méthode Pseudo-DWA pour un CNA du 4-bits, b) spectre de la sortie avec Pseudo-DWA et $f_{Com_{inv}} = \frac{4 * f_c}{OSR}$ puis $f_{Com_{inv}} = \frac{10 * f_c}{OSR}$

bon SFDR sans beaucoup dégrader le SNDR. Il offre l'avantage d'un circuit simple, c.-à-d. les circuits de DWA plus un numérateur de "Johnson", une porte "NOT" et un Multiplexeur 2×1 comme cela est présenté sur la figure 4.26.

Comparaison des différentes solutions destinées à empêcher le comportement tonal de DWA

Le tableau 4.1 résume les résultats de simulation pour un modulateur passe-bas avec différents types de DWA, pour $M = 16$, $OSR = 64$, $L = 3$, $m = 1$. On remarque que les techniques de

TAB. 4.1 – *Comparaison des différentes solutions pour blanchir les tons de DWA dans un modulateur passe-bas*

DEM-méthode	SNDR-maximal [dB]
CNA-Idéal	92.8
Sans DEM	60
DWA, sans entrée périodique	91
Pseudo-DWA, $n_{inv} = 128$	89.9
Pseudo-DWA, $n_{inv} = 64$	90
Bi-DWA	85
RDWA	85
RnDWA	84
IDWA	84
DIA	84
DWA-partielle	82
DWA+dither	80
DWA+offset	80

DWA-segmentée et pseudo-DWA (optimisée) donnent le meilleur SNDR. Pour le critère de simplicité des circuits et la faisabilité, peut-être préfère-t-on la technique DWA-segmentée qui ne possède pas beaucoup de différences avec le cas du meilleur SNDR, et le cas du meilleur SFDR.

4.7 Brassage de sources d'ordre élevé par DWA

Dans, les références [89, 90], une méthode générale basée sur le concept de DWA-DEM pour les applications d'ordre élevé passe-bas et passe-bande, est proposée et analysée. Au delà du premier ordre (cette méthode peut être réduite à un algorithme DWA classique) elle devient si complexe par ailleurs qu'elle ne peut plus s'appliquer. Par exemple, pour un système DEM de deuxième ordre passe-bas, elle a besoin au moins des deux conditions suivantes dans les cas simplifiés :

- le CNA doit être commuté au moins avec une horloge 4 fois plus rapide, c.-à-d. $f_{CNA} \geq 4f_e$.
- chaque cellule du CNA doit avoir la possibilité d'intervenir avec une valeur positive ou négative sans différence absolue, c.-à-d. $\pm\Delta_i$.

Dans le cas présent, i.e. un modulateur passe-bande à haute fréquence, les conditions ci-dessus sont très difficiles à obtenir, voire impossible pour des applications dans lesquelles le CNA est composé de sources de courant commutées. Ce type de CNA peut poser en plus d'autres problèmes dynamiques qui limitent la performance du système. Cet algorithme ne nous intéresse donc pas et nous ne le traiterons pas ici.

4.8 Brassage des sources en rebouclage vectoriel, (Vector feedback, SDEM)

R. Schreier a présenté l'idée de filtrage des erreurs du CNA dans le domaine numérique [12]. Dans la littérature, elle est appelée le brassage vectoriel des sources "*Vector-Based Mismatch Shaping*", ou

brassage des sources par un tri, "*Sorting Dynamic Element Matching, SDEM*" [1,2,79,80,83,127–132]. Dans ce document, elle sera simplement nommée méthode SDEM⁸.

Cette idée a ouvert une nouvelle voie au développement des convertisseurs multibit. On peut théoriquement diminuer l'effet de la non linéarité du CNA grâce à un système de filtrage arbitraire, c.-à-d. différents ordres et différents types de bandes passantes. En pratique, cela fonctionne bien pour le cas passe-bas du premier et du deuxième ordre, ainsi que pour le cas passe-bande du deuxième ou du quatrième ordre. Cependant, pour une DEM d'ordre plus élevé, le problème de stabilité devient complexe et dominant.

La structure générale de la SDEM est représentée sur la figure 4.27. La sortie numérique $v(n)$ du quantificateur est un signal B -bit binaire appliqué à l'entrée de la SDEM, puis converti en $M = 2^B$ signaux $sv(n)$. Le bloc SDEM se compose d'un quantificateur vectoriel spécial, et de M boucles de filtrage ayant une fonction de transfert de $H(z)$. Chaque boucle de filtrage ressemble à une boucle

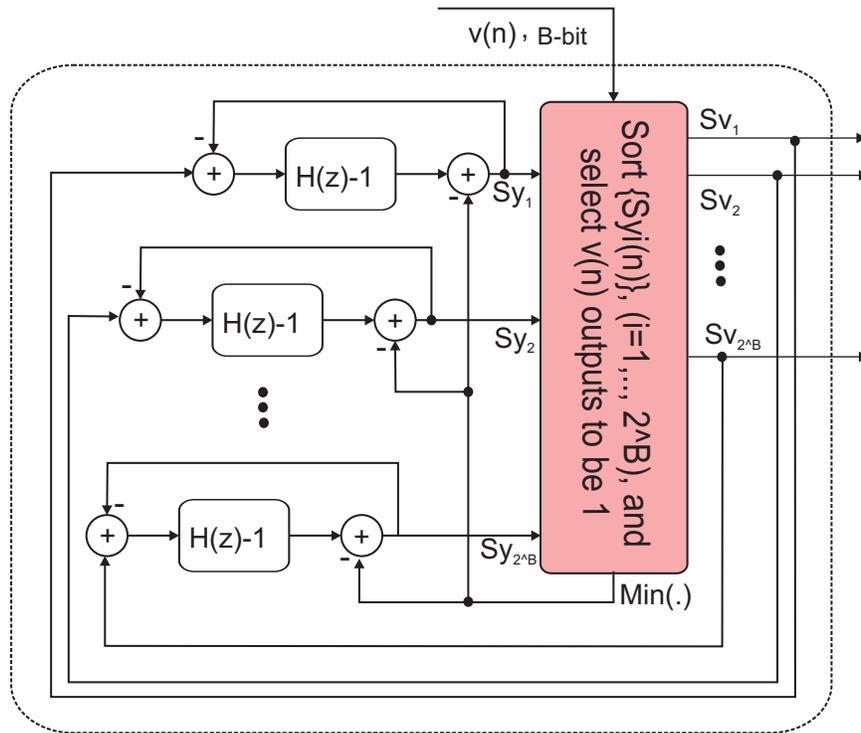


FIG. 4.27 – Structure général de SDEM

de modulateur $\Delta\Sigma$ (voir figure 4.28). Le quantificateur vectoriel compare les M sorties des filtres

8. L'hypothèse fondamentale est qu'une éventuelle erreur de gain $\bar{\alpha}$ et d'offset ϵ sont admises. Pour un CNA de M éléments normalisés, l'algorithme de DEM doit présenter une fonction de transfert comme celle de l'équation ci-dessous, où le terme $e(n)$ doit être minimisé en bande utile.

$$Do(n) = (1 + \bar{\alpha})sv(n) + e(n) + \epsilon, \quad \text{où: } \bar{\alpha} = \frac{1}{M} \sum_{i=1}^M \alpha_i, \quad \text{et } \epsilon = \sum_{i=1}^M \epsilon_i \quad (4.17)$$

Ainsi, dans la suite nous utiliserons le modèle linéaire du CNA comme dans le cas idéal, c.-à-d. $Do(n) = \sum_{i=1}^M sv(n) + e(n)$ pour laquelle seulement le terme de $e(n)$ doit être diminué dans la bande utile. En outre, nous supposons: $\bar{\alpha} \equiv 0$, ainsi que $\epsilon \equiv 0$. Cela ne change en rien l'analyse de DEM, car aucune méthode de brassage des sources ne peut corriger ces termes. De plus, ces termes $(\bar{\alpha}, \epsilon)$ ne causent pas d'erreur de la non linéarité pour le CNA.

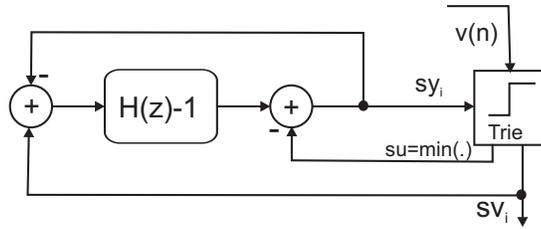


FIG. 4.28 – Une boucle simplifiée de technique SDEM.

$sy_i(n), i = \{1, 2, \dots, M\}$, les trie par ordre ascendant, choisit les $v(n)$ premières sorties pour activer $sv_i(n) = 1, i \in \{v(n)$ premières lignes pour lesquelles $sy_i(n)$ sont les plus petits que les autres }, puis inactive les autres $sv_i(n) = 0, i \in \{M - v(n)$ dernières lignes pour lesquelles $sy_i(n)$ sont plus grandes que le reste }. La structure du filtre utilisé dans chaque boucle est conçue selon le type d'application avec l'ordre nécessaire au SDEM souhaité. Par exemple, pour une SDEM passe-bas du premier ordre, le filtre $H(z)$ vaut $(1 - z^{-1})$, donc la logique nécessaire est réduite à un simple retard, car $(H(z) - 1) = z^{-1}$. Le tableau 4.2 présente quelques autres fonctions de transfert utiles.

Les points d'entrée de la boucle d'une SDEM et d'un modulateur sigma delta constituent une dif-

TAB. 4.2 – Différentes fonctions transferts pour SDEM d'ordre L

$[H(z) - 1]$	Mis en forme passe-bas (d'ordre $L = m$)	Mis en forme passe-bande (d'ordre $L = 2m$)
$m = 1$	z^{-1}	z^{-2}
$m = 2$	$z^{-2} - 2z^{-1}$	$z^{-4} + 2z^{-2}$
$m = 3$	$-z^{-3} + 3z^{-2} - 3z^{-1}$	$z^{-6} + 3z^{-4} + 3z^{-1}$

férence importante. Dans le cas de la SDEM, l'entrée peut être considérée comme égale à zéro pour toutes ses M boucles. En pratique, afin de minimiser les tailles de registres, la valeur minimale de la sortie des filtres $su(n) = \min\{sy_1(n), sy_2(n), \dots, sy_M(n)\}$ est soustraite de toutes les valeurs prises par les sorties des filtres. Cette valeur $su(n)$ ne change pas l'ordre du tirage des $sy(n)$; elle correspond à la valeur moyenne du signal d'entrée $v(n)$, $E[su] = \frac{1}{M}E[v]$.

L'opération qui consomme la majorité de la logique et de la surface de la méthode SDEM concerne la partie quantification vectorielle. Le quantificateur vectoriel doit trier M entrées au moyen de $(M - 1)!$ comparaisons numériques puis effectuer un décodage de $(M - 1)! \times M$. L'importance de la partie tri de l'algorithme peut justifier le nom donné à cette technique "*Sorting Dynamic Element Matching, SDEM*". Au fur et à mesure que M augmente, cette opération devient plus complexe : cette tendance est exponentielle.

Afin de mettre en équation le fonctionnement de la méthode SDEM, on procède à une analyse linéaire en remplaçant le quantificateur vectoriel par un additionneur de façon à additionner l'ensemble des erreurs de défaut d'appariement des cellules du CNA au signal d'entrée $sy(n)$. Considérons le schéma 4.28, on obtient dans le domaine échantillonné :

$$SV_i(z) = SE(z).H(z) + SU(z)$$

$$\begin{aligned}
SY_i(z) &= SE_i(z)[H(z) - 1] - SU(z) = SE_i(z).H(z) - [SV_i(z) - SY(z)] + SU(z) \\
Do(z) &= \sum SV_i(z).(1 + \alpha_i) = V(z) + H(z). \sum_{i=1}^M SE_i(z).\alpha_i + \sum_{i=1}^M SU(z).\alpha_i
\end{aligned} \tag{4.18}$$

où $SU(z)$ est identique pour toutes les cellules et indépendant de l'indice "i". Le terme $\sum_{i=1}^M \alpha_i$ peut être considéré comme une erreur de gain et être négligé, donc :

$$Do(z) = V(z) + H(z). \sum_{i=1}^M SE_i(z).\alpha_i \tag{4.19}$$

On voit bien que l'erreur de défaut d'appariement des cellules du CNA $\alpha_i, i \in [1, 2, \dots, M]$ est mise en forme par les filtres $H(z)$, à la condition que $SE(z)$ ait une valeur limitée. Autrement dit, si l'ensemble du système de SDEM est stable, les erreurs sont mises en forme par $H(z)$. Malheureusement, la stabilité de SDEM n'est pas, en général mathématiquement définie. Une SDEM du premier ordre est toujours stable. Cependant, les simulations et les expériences montrent qu'une SDEM d'ordre élevé est stable si la boucle du modulateur $\Delta\Sigma$ similaire est inconditionnellement stable avec un quantificateur monobit. Les signaux intermédiaires des $SE_i(z)$ ont en pratique un spectre quasiment blanc et la quantité de logique nécessaire à la réalisation des filtres du premier et deuxième ordre est assez modérée. Par conséquent la puissance des erreurs présente dans la bande utile du CNA (auteur d'une fréquence f_o) avec une SDEM peut être estimée selon :

$$P_{SDEM} = \frac{1}{2} \int_{f_o - f_B}^{f_o + f_B} |H(j\omega)|^2 \sigma_{se}^2 \sigma_\alpha^2 d\omega \tag{4.20}$$

Par exemple pour une SDEM du premier et deuxième ordre passe-bas, la puissance dans la bande se simplifie comme suit :

$$\begin{aligned}
P_{SDEM, 1^{ie} ordre} &= \frac{8\pi^2 \sigma_m^2}{3OSR^3} \\
P_{SDEM, 2^{me} ordre} &= \frac{16\pi^4 \sigma_m^2}{3OSR^5}
\end{aligned} \tag{4.21}$$

Alors, la résolution maximale peut être estimée par :

$$\begin{aligned}
ENOB_{SDEM, 1^{ie} ordre} &= \frac{3MOSR^3 \sigma_m^2}{8\pi^2} \\
ENOB_{SDEM, 2^{me} ordre} &= \frac{5MOSR^5 \sigma_m^2}{8\pi^4}
\end{aligned} \tag{4.22}$$

La figure 4.29 montre les résultats de simulation d'un modulateur passe-bas multibit avec une erreur de 1% corrigé par SDEM. La réalisation de SDEM est plus complexe que celle des cas précédents (DWA, ILA et RDEM) pour deux raisons. D'abord, le nombre de circuits de filtrage dans la boucle augmente avec le nombre de cellules du CNA (M). De plus, la complexité des circuits effectuant un tri de sorties des filtres augmente de façon exponentielle au fur et à mesure que M augmente (voir la figure 4.30). Certaines stratégies de tri simplifiées sont proposées; on en montre un exemple sur la figure 4.31. La méthode SDEM est la meilleure méthode de brassage de sources, à notre connaissance. Mais

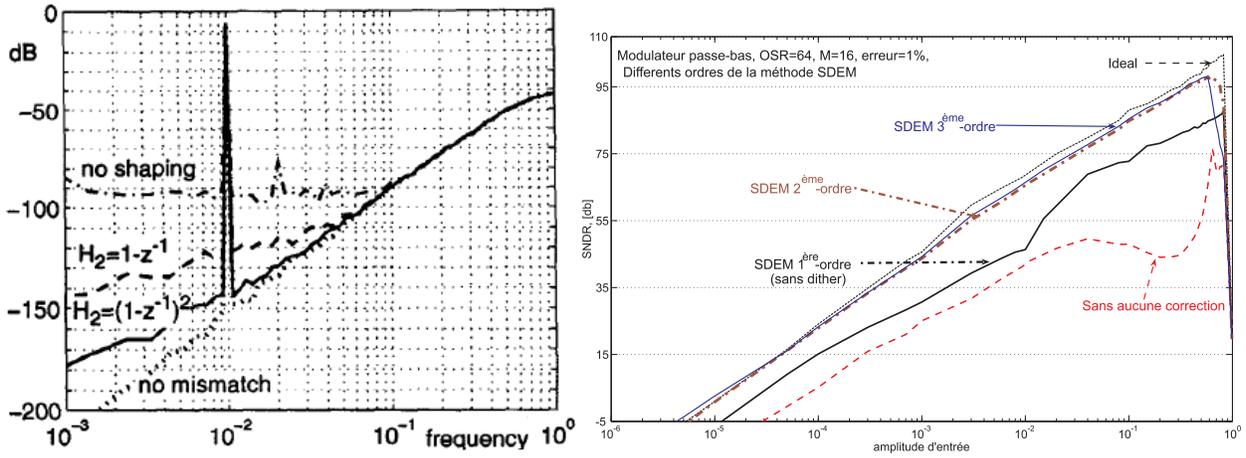


FIG. 4.29 – a) Spectre de la sortie de modulateur multibit avec 1% d'erreur corrigée par différents ordres de SDEM, b) la performance du CNA ayant un bloc SDEM qui est comparé par d'autres méthodes

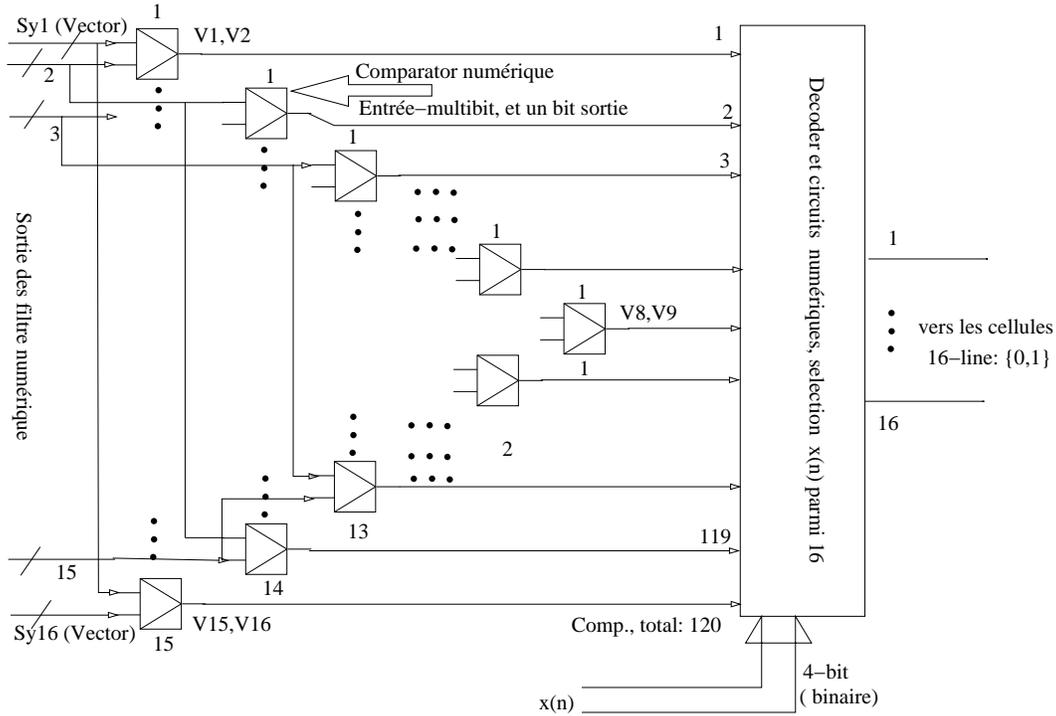


FIG. 4.30 – Schéma bloc de la réalisation directe d'un algorithme de tri pour SDEM, avec $M=16$

cette méthode est malheureusement difficile à réaliser, en particulier pour un CNA multibit ayant un grand nombre de cellules par exemple $M > 8$. C'est une des raisons pour laquelle la méthode suivante, qui a à peu près le même fonctionnement, est présentée.

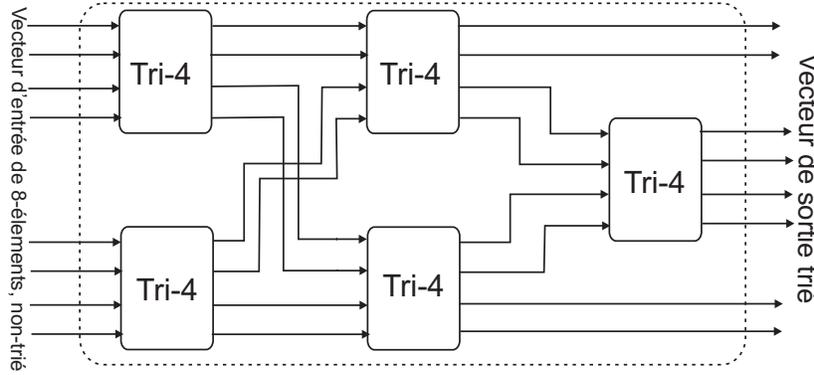


FIG. 4.31 – Une stratégie simplifiée du tri pour SDEM, avec $M=8$

4.9 Brassage de source en structure arborescente, (Tree structured, TDEM)

Malgré de bons résultats, la technique SDEM souffre de sa complexité. Par ailleurs, la structure arborescente est connue comme étant une technique efficace pour réduire la complexité de l'électronique numérique associée. I. Galton a proposé une stratégie de DEM à structure arborescente dans laquelle la mise en forme des erreurs du CNA par un filtre arbitraire analogue à SDEM est conservée [11]. Cette méthode est appelée TDEM "*Tree structured Dynamic Element Matching*". Afin de mieux expliquer son principe, on considère un schéma général de TDEM avec un CNA composé de $M = 2^B$ cellules, lequel est représenté sur la figure 4.32. Les B couches ("Layers") différentes sont numérotées "1" pour la couche la plus proche du CNA, à " M " pour la couche qui reçoit l'entrée de TDEM (nommée $v(n)$) en premier. L'entrée de TDEM des M éléments peut varier entre "0" et " M ", ce qui suppose que $v(n)$ est un signal de $(B + 1)$ bits. Chaque colonne d'indice " k ", $k \in \{1, 2, \dots, B\}$ est composée de $2^{(B-K)}$ blocs en ligne différents nommés blocs de commutation $SB_{k,r}$ ("Switching Block"), " k " indique le nombre de couches et " r " indique la position du bloc dans la couche. Un $SB_{k,r}$ a une entrée de $(k + 1)$ bits ($x_{k,r}(n)$) et deux sorties de k bits ($x_{k-1,2r-1}(n)$ et $x_{k-1,2r}(n)$). Afin que l'équation 4.1 soit respectée ($v(n) = \sum_{i=1}^M sv_i(n)$), la somme des deux sorties doit maintenir constante la valeur d'entrée de chaque bloc $SB_{k,r}$, comme l'exprime l'équation suivante :

$$\begin{aligned} x_{k-1,2r-1}(n) &= \frac{1}{2}[x_{k,r}(n) + s_{k,r}(n)] \\ x_{k-1,2r}(n) &= \frac{1}{2}[x_{k,r}(n) - s_{k,r}(n)] \end{aligned} \quad (4.23)$$

où $s_{k,r}(n)$ est une séquence définissant la différence entre deux sorties de bloc :

$$s_{k,r}(n) = [x_{k-1,2r-1}(n) - x_{k-1,2r}(n)] \quad (4.24)$$

Ce mécanisme s'effectue à l'intérieur même de chaque bloc $SB_{k,r}$. La séquence $s_{k,r}$ qui détermine la différence entre deux sorties est un nombre entier, par conséquent, elle doit satisfaire certaines conditions :

$$s_{k,r}(n) = \begin{cases} \text{pair} & \text{si } x_{k,r} \text{ est pair} \\ \text{impair} & \text{si } x_{k,r} \text{ est impair} \end{cases}$$

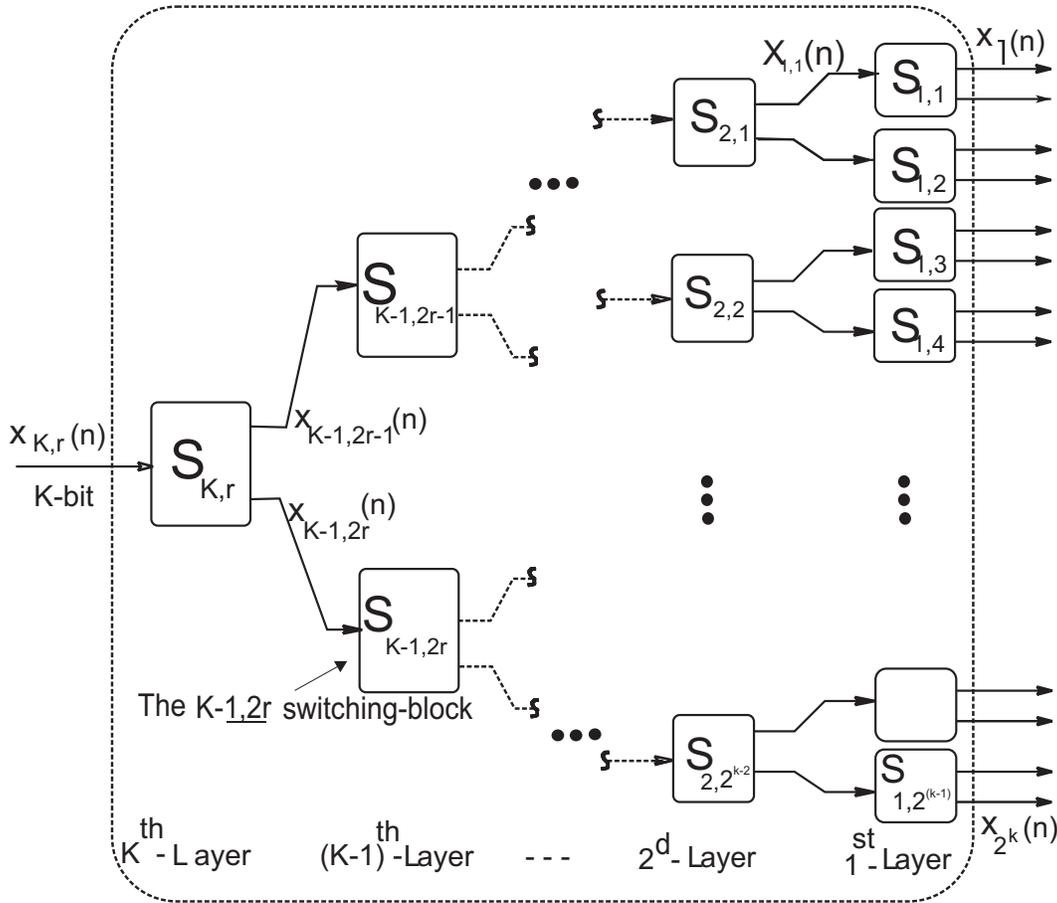


FIG. 4.32 – Structure générale de TDEM pour $M=16$

et,

$$|s_{k,r}(n)| \leq \min\{x_{k,r}(n), 2^k - x_{k,r}(n)\} \quad (4.25)$$

La structure capable de générer les sorties des blocs $SB_{k,r}$ et la séquence souhaitable $s_{k,r}(n)$ sont représentées à la figure 4.33.

où, un quantificateur spécial est chargé de présenter la valeur adéquate de $s_{k,r}(n)$ à chaque période en respectant les conditions 4.25. Le filtre numérique $H(z)$ est de type coupe-bande; il met en forme $S_{k,r}$ et dépend des application souhaitées.

La caractéristique du quantificateur spécial a été initialement définie par la relation suivante :

$$s_{k,r}(n) = \begin{cases} 0 & \text{si } x_{k,r}(n) \text{ est pair} \\ +1 & \text{si } x_{k,r}(n) \text{ est impair et } v_{k,r}(n) > 0 \\ -1 & \text{si } x_{k,r}(n) \text{ est impair et } v_{k,r}(n) < 0 \end{cases} \quad (4.26)$$

où $v_{k,r}(n)$ est la sortie du dernier étage du filtre de boucle. Même si la définition ci-dessus respecte les conditions 4.25, elle doit être modifiée pour des raisons de stabilité de l'algorithme ainsi que pour empêcher des tons dans la bande utile, comme on l'explique dans la sous section suivante. La figure 4.34 montre deux exemples d'implantation d'un bloc $SB_{k,r}$ pour une TDEM passe-bas du premier et deuxième ordre respectivement.

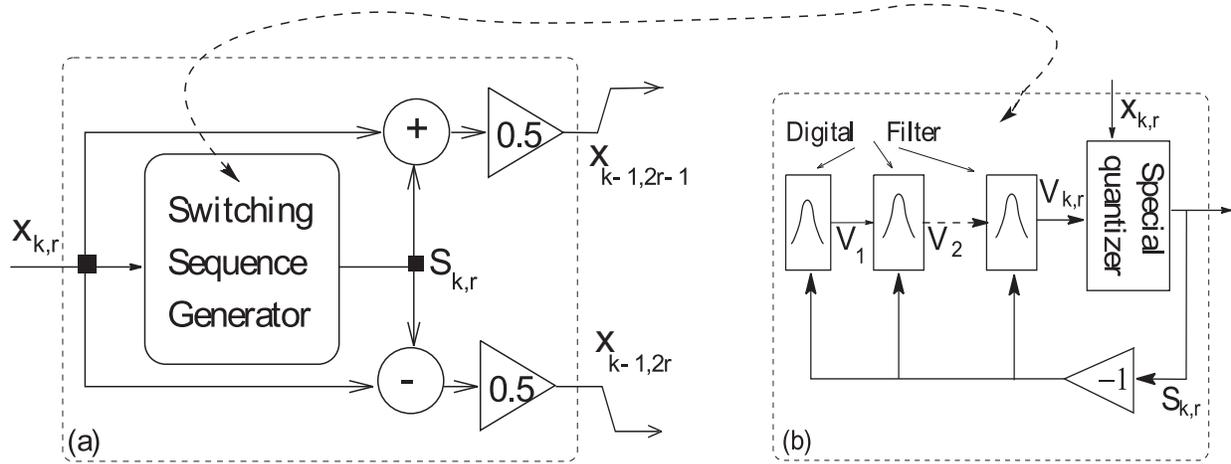


FIG. 4.33 – a) Structure générale de $SB_{k,r}$ pour l'algorithme TDEM, b) boucle de filtrage qui génère la séquence souhaitable pour $s_{k,r}(n)$

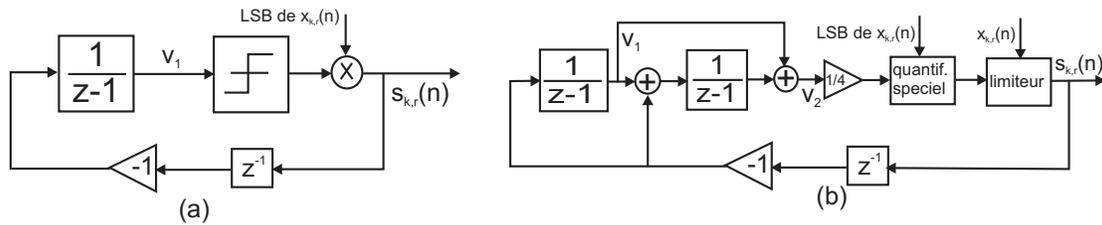


FIG. 4.34 – a) Schéma bloc simplifié d'une TDEM passe-bas : a) du premier ordre b) du deuxième ordre

Le détail de la mise en équation du fonctionnement de TDEM est présenté précisément dans les références [1, 11, 133–140]. La sortie finale d'un CNA à $M = 2^B$ cellules qui utilise un brassage de sources TDEM est donnée par l'équation 4.27 :

$$\begin{aligned}
 Do_{M\text{-cellules}}(n) &= (1 + \bar{\alpha})v(n) + e(n) + \epsilon \\
 \bar{\alpha} &= \frac{1}{2^B} \sum_{i=1}^{2^B} \alpha_i \quad , \quad \epsilon = \sum_{i=1}^{2^B} \epsilon_i \\
 e(n) &= \sum_{k=1}^B \sum_{r=1}^{2^{B-k}} \Delta_{k,r} s_{k,r}(n) \\
 \Delta_{k,r} &= \frac{1}{2^k} \sum_{i=(r-1)2^k+1}^{(r-1)2^k+2^k-1} (\alpha_i - \alpha_{i+2^k-1}) \quad (4.27)
 \end{aligned}$$

où $\bar{\alpha}$, qui est l'erreur du gain, et ϵ qui est l'offset du CNA, ne sont qu'une erreur linéaire. Ainsi comme dans les cas précédents, on peut les négliger afin de simplifier les équations, cela conduit à $\epsilon \equiv 0$.

En revanche, le terme $e(n)$ est l'erreur de non linéarité qui limite la résolution du CNA. Grâce à l'utilisation du brassage TDEM, $e(n)$ est bien mise en forme par des filtres $H(z)$ implantées dans les différents blocs de commutation $SB_{k,r}$, comme l'exprime l'équation 4.27. En outre, le spectre final d'erreur de non linéarité des cellules désappariées du CNA est mise en forme par une fonction de transfert d'ordre L si chacune des séquence $s_{k,r}(n)$ est mise en forme par un filtre d'ordre L . À

condition que toutes les séquences $s_{k,r}(n)$ soient délimitées par une valeur non infinie pour une durée J arbitraire.

$$\sum_{j=1}^J s_{k,r}(n) \leq P, \quad \text{pour toutes les } (M-1) \text{ blocs, quand : } |P| \neq \infty \quad (4.28)$$

Autrement dit, l'algorithme doit être stable pour toutes sortes de signaux à l'entrée du CNA. Cette condition est extrêmement critique pour une TDEM d'ordre $L > 1$. Nous allons traiter de cet inconvénient à la section suivante.

Que ces séquences soient indépendantes l'une de l'autre est une autre condition : $M-1$ filtres doivent être utilisés. Cela est déjà pris en compte dans la structure de TDEM par la présence de $M-1$ blocs $SB_{k,r}$.

En ce qui concerne la résolution offerte par la méthode TDEM, les équations 4.21 et 4.22 restent valables, mais à la condition que la TDEM soit stable.

4.9.1 Problème d'instabilité de l'algorithme TDEM

Bien qu'un filtrage d'ordre arbitraire puisse être facilement mis en application de façon numérique, la règle de conservation du nombre de sources sélectionnées (équation 4.25) rend difficile la conception des algorithmes TDEM. Ce problème de stabilité non résolu est tout à fait différent de celui des filtres classiques. La stabilité dans les filtres est garantie habituellement par une optimisation appropriée des coefficients, et peut être examinée mathématiquement en utilisant quelques méthodes comme la méthode de "Nyquist" ou le lieu des pôles. Cependant, dans le cas du bloc de commutation $SB_{k,r}$, l'instabilité réside principalement dans la méthode de quantification qui est fortement non linéaire selon les propriétés stochastiques de l'entrée $x_{k,r}(n)$. De plus, la décision propre sur chacun des $SB_{k,r}$ est différemment imposée par la règle de conservation du nombre, selon le nombre de couches (l'équation 4.25).

En général, la stabilité des algorithmes d'ordre élevé de DEM ne peut pas être vérifiée analytiquement; elle est habituellement vérifiée par des simulations. Comme le montre la figure 4.34, pour chaque SB, le filtre numérique se compose d'un ou plusieurs intégrateurs pour des algorithmes passe-bas et un ou plusieurs résonateurs pour des algorithmes passe-bandes. La fonction de transfert d'un quantificateur spécial, qui doit fonctionner comme un quantificateur et un limiteur, a été au départ choisie de la forme la plus simple possible selon l'équation 4.26. Pour une TDEM du premier ordre, cette définition est suffisante et mène à une implantation très simple, sauf que l'on doit y ajouter un générateur de bruit blanc additif (dither) pour empêcher les tons.

Cependant, si nous avons besoin d'un second ordre ou d'un ordre supérieur, la définition ci-dessus impose une limite importante pour l'algorithme. Par exemple, dans la TDEM du deuxième ordre passe-bas représentée sur la figure 4.34, quand $x_{k,r}(n)$ est une séquence de nombres pairs, les $s_{k,r}(n)$ sont égaux à zéro jusqu'à la fin de la séquence. Ces séquences critiques d'entrée peuvent être composées de n'importe quelle combinaison de $\{0,2\}$ et de $\{0,2,4\}$ dans les premières et deuxièmes couches respectivement. Cependant, si la valeur initiale du premier intégrateur (V_1) au début n'est pas égale

à zéro, le deuxième intégrateur peut être saturé en quelques périodes selon la taille de ses registres (V_2). On peut également observer cette saturation par l'effet d'une répétition mineure de séquences d'entrées impaires entre les séquences majoritaires paires. La conséquence de ceci est que la mise en forme du bruit sera suspendue jusqu'à ce que les filtres sortent de la saturation.

La figure 4.35 montre la dynamique de sortie du filtre ($V_2(n)$). Elle est obtenue pour une TDEM passe-bande de second ordre conventionnelle simulée avec $MTF = \frac{1}{1+H(z)} = (1+z^{-2})^2$ en utilisant l'équation 4.26⁹. On voit que les registres vont saturer, alors qu'un modulateur d'ordre deux est tou-

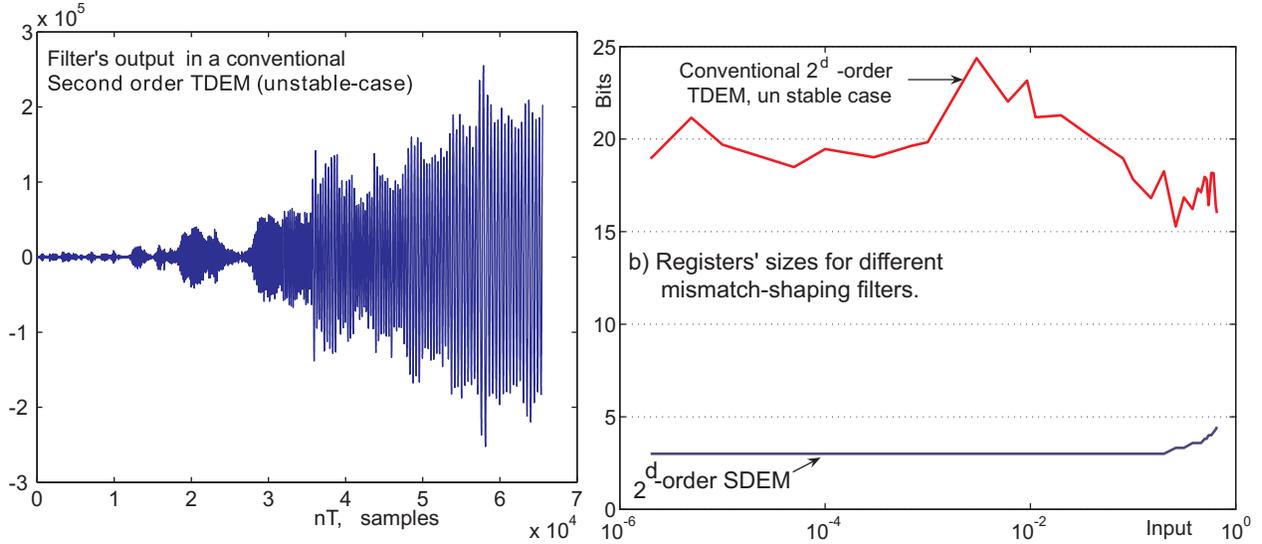


FIG. 4.35 – a) Sortie du filtre dans une TDEM conventionnelle de second ordre, b) taille nécessaire des registres pour une TDEM de second ordre conventionnelle et une SDEM similaire.

jours stable. Sur la figure 4.35-b les dynamiques de sorties d'une TDEM conventionnelle du deuxième ordre sont comparées avec celles d'une SDEM similaire en fonction de la variation de l'amplitude d'entrée. Cette figure met en évidence la différence entre un algorithme stable et un algorithme instable, point faible de la TDEM.

En effet, dans une situation de saturation, une décision efficace ne pourrait pas être correctement prise même si les registres étaient de dimension infinie. Ceci est dû au fait que le gain linéaire limité de la boucle ne permet pas que les nouvelles entrées impaires que produit un $s_{k,r}(n)$ différent de zéro puissent diminuer rapidement le volume accumulé au niveau du second intégrateur. Autrement dit, une fois qu'une telle saturation arrive, une boucle linéaire n'est pas capable d'occasionner un retour à la situation normale même si la séquence d'entrée revient au cas normal. Ce raisonnement indique que la situation serait encore plus grave si la TDEM utilisait des filtres du troisième ordre ou plus.

9. Le terme MTF dénote la fonction de transfert de mise en forme d'erreur du CNA, *c.-à-d.* "Mismatch Transfer Function"

Solutions proposées pour le Problème d'instabilité de TDEM

1- Afin d'empêcher l'instabilité de l'algorithme, l'équation 4.26 a été alors modifiée dans la première exécution de la TDEM [133], de la façon suivante :

$$S_{k,r}(n) = \begin{cases} +1 & x_{k,r} & \text{est impair et} & V_{k,r} > 0 \\ -1 & x_{k,r} & \text{est impair et} & V_{k,r} < 0 \\ +2 & x_{k,r} & \text{est divisible par 4 et} & V_{k,r} > 0 \\ -2 & x_{k,r} & \text{est divisible par 4 et} & V_{k,r} < 0 \\ 0 & & \text{dans tous les autres cas} & \end{cases} \quad (4.29)$$

La modification mentionnée ci-dessus a pour conséquence une meilleure stabilité de l'algorithme, mais cela n'est pas suffisant. D'autre part, les modifications proposées dans 4.29 sont seulement applicables aux couches de TDEM avant le second étage, c.-à-d. $x_{k,r} > 4$ ou $k > 2$. On peut proposer une autre modification en utilisant la condition d'être divisible par 2 au lieu de 4, mais le problème au niveau de la première couche reste non résolu. À notre connaissance, l'instabilité d'une TDEM d'ordre deux ou d'ordre plus grand n'a jamais été éliminée ou améliorée dans les deux premières couches de l'algorithme.

2- Les articles [133, 134] proposent d'autres modifications de l'algorithme TDEM de sorte qu'en cas de saturation au niveau du deuxième étage du filtre, le système n'utilise plus cet étage, mais rend à la place une décision basée uniquement sur la sortie du premier intégrateur. Par conséquent, un tel système fonctionne en tant que TDEM de premier ordre jusqu'à ce que le deuxième intégrateur sorte de la saturation. Malheureusement, ce problème arrive en pratique très souvent, et la mise en forme de l'erreur du défaut d'appariement d'un CNA est plus proche d'une fonction de transfert du premier ordre que de celle du deuxième ordre. Par conséquent, l'implantation d'une TDEM du deuxième ordre, qui nécessite plus de circuits, n'est pas justifiable.

La figure 4.36 montre le SNDR d'un modulateur $\Delta\Sigma$ multibit ayant une TDEM d'ordre 2 pour corriger le défaut d'appariement des cellules du CNA interne comparée à celui d'une TDEM conventionnelle et à une SDEM. Cette comparaison confirme bien la faiblesse de la TDEM ainsi que l'efficacité insuffisante des deux solutions citées ci-dessus même appliquées conjointement.

3- La troisième solution proposée par P.Benabes [141] utilise deux techniques en même temps. La première consiste en une modification de l'équation 4.26 introduite par la suite (voir sur la figure 4.37) :

$$S_{k,r}(n) = \begin{cases} +1 & x_{k,r} & \text{est impair et} & V_{k,r} > 0 \\ -1 & x_{k,r} & \text{est impair et} & V_{k,r} < 0 \\ +2 & 0 < x_{k,r} < M & \text{est pair et} & V_{k,r} > L_{seuil} \\ -2 & 0 < x_{k,r} < M & \text{est pair et} & V_{k,r} < -L_{seuil} \\ 0 & & \text{dans tous les autres cas} & \end{cases} \quad (4.30)$$

où L_{seuil} est un seuil optimisé par la simulation.

La deuxième technique consiste à utiliser un ou plusieurs rebouclages non linéaires λ , qui, en cas de débordement, forcent le gain de boucle à changer. De ce fait, les intégrateurs sont ramenés à la situation normale quand ils sont temporairement débordés par certaines séquences. La figure 4.37 représente un schéma montrant une application de ces techniques et la figure 4.38 montre un exemple

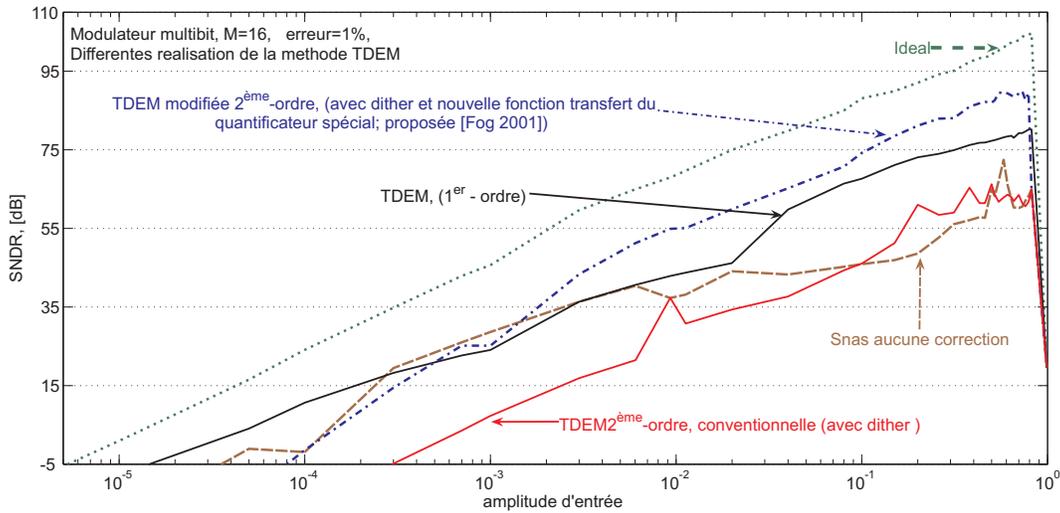


FIG. 4.36 – SNDR du modulateur multibit utilisant une TDEM conventionnelle ou modifiée selon [133,134].

de résultat de simulation au niveau système d'un CNA 4 bit mais en dehors d'une boucle sigma delta. Le CNA est alimenté par une entrée aléatoire $0 \leq x_{4,1} \leq 16$ optimisée pour une TDEM de troisième ordre avec $L_{seuil} = 1024$, $\lambda = 64$, $C_1 = -8$, $C_2 = -48$, $C_3 = -624$, où le volume nécessaire des

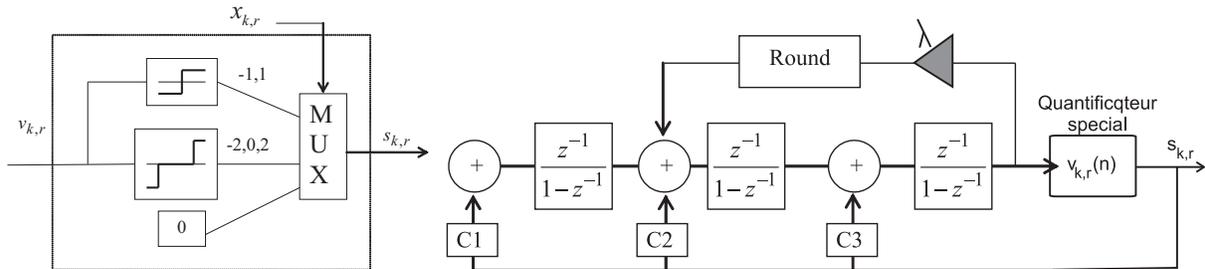


FIG. 4.37 – a) Schéma de la fonction de transfert du quantificateur spécial qui produit $s_{k,r}$, b) schéma d'une TDEM passe-bas du troisième ordre utilisant un rebouclage non linéaire pour stabiliser sa boucle.

registres est de l'ordre de 13 bits. Cette solution est très sensible aux caractéristiques de l'entrée et ne fonctionne pas toujours aussi parfaitement quelque soit l'entrée. De plus, le rebouclage non linéaire restreint la performance du système comme dans la deuxième solution expliquée ci-dessus.

En résumé, une TDEM de premier ordre avec bruit blanc additif peut être intéressante pour des applications passe-bas, mais ne fonctionne pas bien en pratique pour un système d'ordre élevé, que ce soit avec un algorithme conventionnel ou avec un algorithme modifié. Nous proposerons d'autres algorithmes basés sur le principe de la TDEM au prochain chapitre.

4.9.2 Structure TDEM segmentée

Il est possible d'utiliser une structure "TDEM segmentée", qui est généralisée aux différents type de CNA [138]. En fait, la structure TDEM segmentée est adaptée aux CNA qui utilisent des cellules

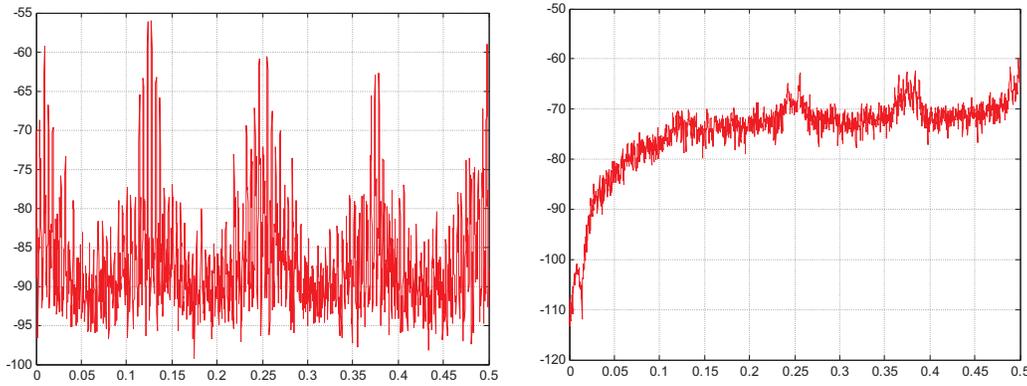


FIG. 4.38 – Spectre de la sortie de CNA lorsqu'on applique une entrée aléatoire: a) sans correction, b) corrigée par la TDEM d'ordre trois proposée en [141] et illustrée à la figure 4.37.

binaires plutôt que des cellules thermométriques. Cela est essentiel quand le nombre de bits du CNA est assez grand, par exemple $B > 6$ bit. La segmentation peut aboutir soit sur des cellules binaires soit sur deux sous partitions différentes. La figure 4.39 montre ces deux stratégies. Dans

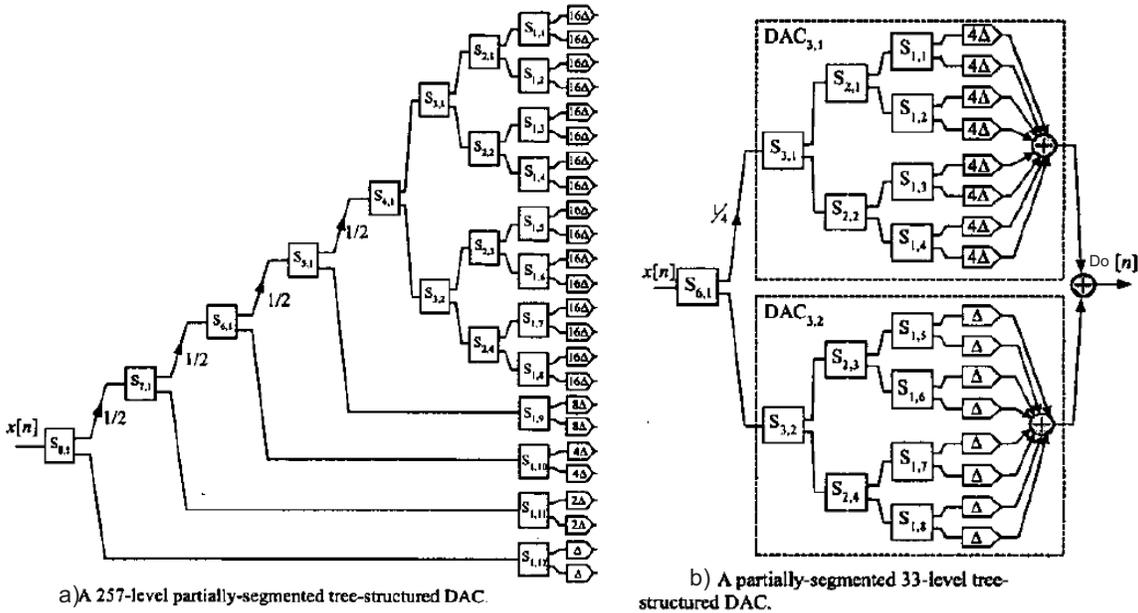


FIG. 4.39 – a) Schéma de TDEM segmenté, a) adapté aux éléments binaires, b) adapté aux CNA avec des éléments de poids différents.

le premier exemple 4.39-a, dans les blocs de commutation $SB_{k,r}$ des 4 premières couches, c.-à-d. $B > k > 4$, le code d'entrée est divisé d'une manière simple et sans aucune possibilité de mise en forme des erreurs. Cependant, les blocs de commutation $SB_{k,r}$ aux 4 dernières couches, c.-à-d. $k \leq 4$, fonctionnent exactement de la même manière que dans une TDEM conventionnelle. Le deuxième exemple représente une autre type de partition qui est adopté pour un CNA composé de cellules de poids différents plutôt que des cellules binaires ou unitaires. Son analyse est détaillée dans la

référence [138]. Le terme $e(n)$, c.-à-d. l'erreur du CNA, peut être calculé au moyen de l'équation 4.27 en considérant seulement les blocs $SB_{k,r}$ qui mettent en forme les erreur, c.-à-d. $k \leq 4$.

Le point faible de TDEM, c.-à-d. le problème d'instabilité des algorithmes d'ordre élevé, est aussi présent dans le cas de la TDEM segmentée.

4.10 Techniques dérivées

Dans cette section, nous présentons brièvement quelques techniques qui ne ressemblent pas aux techniques de DEM mais qui peuvent être utilisées pour éliminer les erreurs statiques des CNA. On les utilise rarement dans les applications classiques, mais certaines peuvent être intéressantes dans des applications particulières.

4.10.1 Double quantificateur

Une façon de résoudre le problème de la non-linéarité du CNA interne, est d'utiliser un CNA monobit dans le rebouclage pour les deux premiers intégrateurs et un CNA multibit aux autres points du rebouclage. Par conséquent, le défaut d'appariement du CNA multibit ne concernera plus le signal d'entrée, et sera mis en forme par les deux premiers intégrateurs, sachant que la sortie du CNA monobit ne produit pas d'erreur de non linéarité vis à vis de l'entrée. La structure d'un tel modulateur ressemble à la structure MASH. Trois façons de le réaliser sont brièvement décrites ci-dessous.

Remarque : cette technique améliore considérablement le SNR du système. En revanche, les avantages d'une structure multibit ne peuvent pas être complètement obtenus dans le cas d'une structure à double quantification, notamment la stabilité, l'abaissement des contraintes sur des circuits actifs, et la diminution de l'effet de la gigue d'horloge. La diminution de l'effet de la gigue d'horloge par une structure multibit est un avantage majeur qui peut justifier le passage d'une structure monobit à une structure multibit pour des circuits plus complexes, alors que ceci est inutile dans la technique de double quantification.

1) Cascade L-0, la technique "Leslie-Singh"

Une structure simple du modulateur $\Delta\Sigma$ multibit avec deux quantificateurs est représentée à la figure 4.40 [142]. C'est un modulateur d'architecture en cascade "MASH" une première boucle d'ordre $L = 2$, et une deuxième boucle d'ordre zéro. La deuxième boucle utilise simplement un quantificateur multibit. Les deux sorties des quantificateurs sont corrigées numériquement comme celle d'un système MASH normal. Bien sûr, on peut aussi utiliser une boucle d'ordre supérieur au premier étage.

De plus, les deux quantificateurs peuvent être combinés, puisque les deux quantificateurs ont exactement la même entrée. Le quantificateur monobit peut être physiquement supprimé parce que sa sortie est toujours égale au MSB du deuxième quantificateur. Ils peuvent être implantés ensemble, le MSB sert de sortie du quantificateur monobit. Leslie a fait l'analyse d'un modèle linéaire du modulateur utilisant une telle structure. La conclusion est que cette structure présente idéalement la même résolution qu'un modulateur multibit conventionnel. Cependant, ses problèmes pratiques tels que les

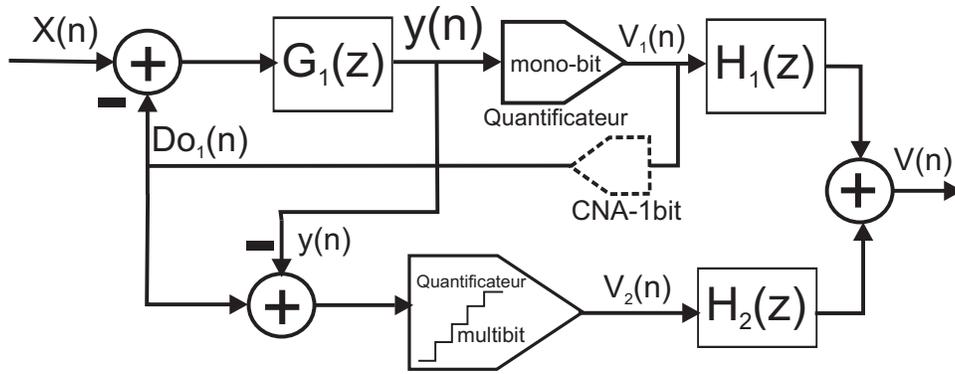


FIG. 4.40 – Structure d'un modulateur multi-boucle "2-0", avec double quantificateur

tons parasites, les contraintes sur les circuits actifs et l'effet de la gigue d'horloge restent proches de ceux du modulateur monobit.

2) Structure MASH à double quantificateur

En gros, dans un modulateur de type MASH, on peut utiliser le quantificateur multibit (suivi par un CNA multibit) pour toutes les boucles, sans être limité par l'erreur de non linéarité, sauf dans la première boucle. Par conséquent, l'erreur du CNA multibit est mise en forme par la fonction de transfert $H_2(z)$, comme le montre l'exemple de la figure 4.41. En plus, des inconvénients habi-

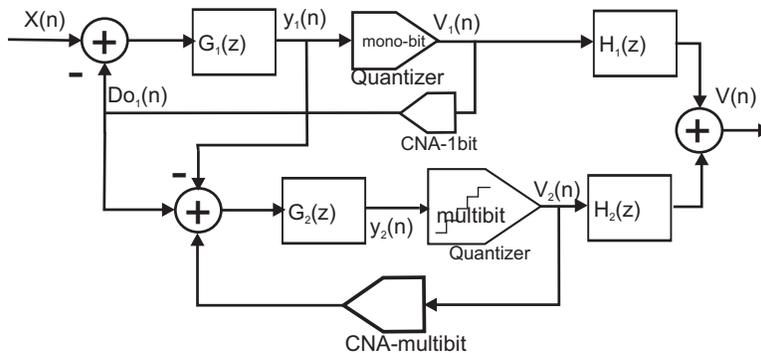


FIG. 4.41 – Double-quantification, structure MASH qui n'a pas besoin de correction des erreurs du CNA multibit car il est utilisé seulement au deuxième étage

tuels d'une structure à double quantificateurs expliqués ci-dessus, le défaut d'appariement entre les différents étages d'une structure MASH est un facteur limitant supplémentaire.

3) Double quantificateur à étage unique

Il s'agit d'une structure de modulateur de type CCFB décrit dans la référence [1], dans lequel les CNA sur les chemins de rétroaction vers les premiers intégrateurs (ou résonateurs) sont monobits, puis les autres sont multibit. Un exemple de la réalisation d'un modulateur sigma delta d'ordre trois passe-bas utilisant deux CNA différents est représenté sur la figure 4.42. Un tel modulateur, avec un

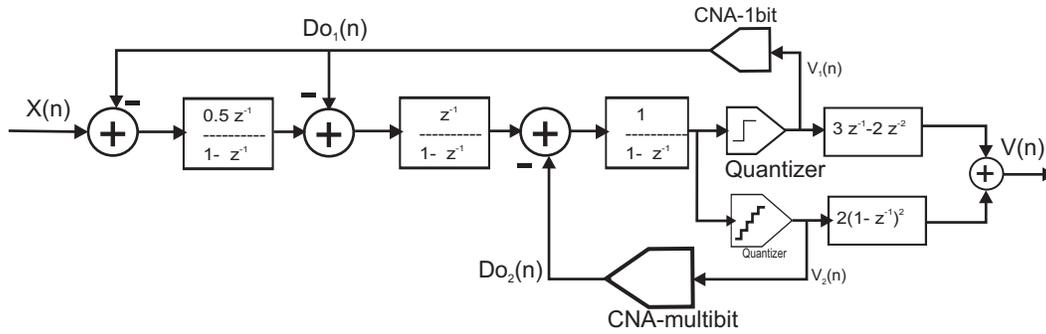


FIG. 4.42 – Double-quantification, structure CIFB utilisent deux type de reboilage; monobit pour les deux premiers intégrateurs et multibit pour les intégrateurs suivants.

quantificateur multibit, semble un bon compromis entre la stabilité de la boucle et le SNR du système. En résumé, cette technique possède quelques avantages d’un modulateur multibit, sans ajouter d’effet dégradant supplémentaire.

4.10.2 CNA série

Un autre moyen de résoudre le problème du défaut d’appariement consiste à utiliser un CNA monobit activé en série par l’entrée multibit. Bien sûr, la fréquence appliquée au CNA doit être au moins M fois supérieure au cas normal. En plus, il est nécessaire qu’un tel CNA fonctionne en mode RZ afin de limiter l’effet de la différence entre le temps de montée et le temps de descente. La figure 4.43 montre le schéma d’un CNA multibit utilisant une seule cellule excitée par une entre série [143].

On sait bien que dans un modulateur sigma delta conventionnel, pour que les circuits puissent

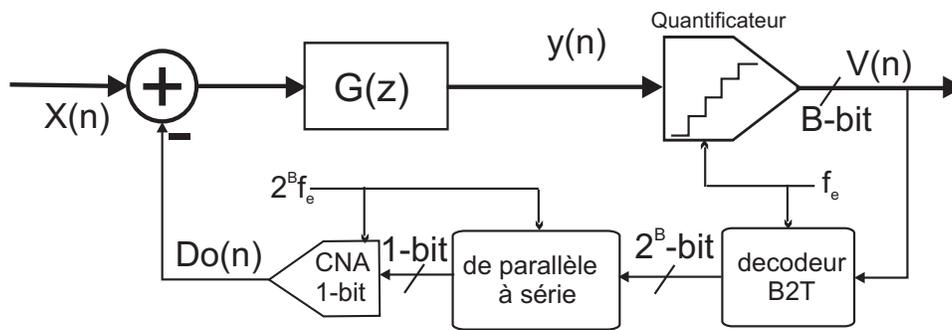


FIG. 4.43 – Schéma d’un CNA série

fonctionner à une fréquence supérieure, il est préférable d’utiliser un OSR plus élevé plutôt qu’un quantificateur multibit. Jusqu’ici, l’idée d’un CNA série semblait inutile, mais si on arrive à concevoir un modulateur avec M rétroactions en parallèle qui fonctionnent en phase $\frac{2\pi}{M}$, l’idée devient intéressante.

4.11 Ajustement physique des éléments du CNA

Une méthode directe d'amélioration de la linéarité du CNA est l'ajustement de ses différentes cellules par des procédés physiques. Les éléments communs du CNA qui peuvent être ajustées sont les résistances et les condensateurs. En général, un rayon laser est utilisé afin d'ajuster la valeur exacte des éléments autour de leur valeur nominale. Cette technique doit donc s'effectuer avant que le circuit ne soit mis en boîtier, et avant que le circuit ne soit placé dans un environnement réel. La technique coûte cher et elle ne peut pas compenser d'éventuels désappariements des cellules en cours de fonctionnement, par exemple à cause des gradients de température.

Une autre technique consiste de placer des petits éléments supplémentaires, par exemple, des petits condensateurs en parallèle avec le condensateur principal, afin de ramener la cellule réelle autour de sa valeur nominale. Cela peut se faire au moyen d'une information enregistrée, en mémoire de type ROM ou périodiquement à chaque mise en marche du système dans des registres numériques. Cette méthode a besoin d'un système de mesure et de traitement associé qui augmente fortement la complexité des circuits. En pratique, la technique d'ajustement physique n'est pas beaucoup utilisée, sauf dans des cas très précis et spéciaux [2].

4.12 Techniques d'étalonnage numérique

L'erreur statique du CNA multibit pourra être corrigée dans le domaine numérique si la grandeur d'erreur est estimée puis enregistrée dans une mémoire pour chaque code de sortie du modulateur $\Delta\Sigma$ (ou l'entrée du CNA). Ceci est l'idée cachée derrière des techniques connues comme l'étalonnage numérique ("Digital Correction of DAC nonlinearity") [1, 2, 77, 144].

Une première stratégie est basée sur l'acquisition des erreurs à la mise en marche du système. Le schéma d'un tel modulateur est représenté à la figure 4.44-a, où le code sortant du modulateur doit adresser la valeur réelle de la sortie $Do(n)$ supposée équivalente au signal d'entrée $x(n)$. En fait, chaque code sortant représente une valeur estimée de la sortie du CNA, qui a été calculée et enregistrée en phase d'estimation, dans une mémoire RAM.

La figure 4.44-b montre un schéma simplifié d'un système d'acquisition de valeurs exactes du CNA utilisant un modulateur principal à rebouclage monobit. L'entrée du modulateur en phase d'estimation est la sortie analogique du CNA principal excité par un compteur numérique qui produit tous les codes possible $[0, 1, \dots, M]$. Ce processus d'étalonnage est assez long car chaque code d'entrée du modulateur doit s'appliquer au moins 2^{ENOB} périodes. L'estimation des erreurs de chaque code doit être aussi précise que la précision finale, c.-à-d. ENOB, donc la mémoire RAM doit contenir des données de l'ordre de ENOB bit pour chaque adresse entre 0 et M . Par conséquent, les parties suivantes, par exemple la décimation, travaillent sur un nombre de bits plus élevé que dans le cas normal.

Une deuxième stratégie est communément appelée étalonnage en tâche de fond "*Digital correction with background calibration*", pour laquelle les erreurs des cellules du CNA sont mises à jour de façon continue. On la préfère donc, puisqu'elle considère les éléments d'erreur qui deviennent importants au

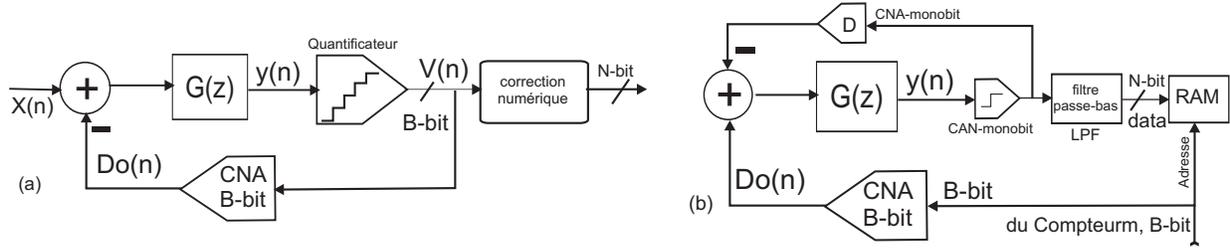


FIG. 4.44 – Première stratégie l'étalonnage, a) le schéma global, b) le système en phase d'estimation des erreurs du CNA multibit

cours de l'utilisation, par exemple, en cas de non équilibre thermique entre des cellules. Le principe de cette technique est représenté sur la figure 4.45, où un bloc de brassage aléatoire est inséré entre le quantificateur et le CNA thermométrique. La fonction de transfert " $ETF(z)$ " met en forme les erreurs du CNA selon l'application souhaitée :

$$\begin{aligned} v(n) &= x(n) * STF(n) + q(n) * NTF(n) + e(n) * ETF(n) \\ e(n) &= \sum_{i=1}^M \alpha_i \cdot sv_i(n) \end{aligned} \quad (4.31)$$

où, "*" dénote l'opération de la convolution en temps. Le bloc corrélateur détermine la valeur de l'espérance des erreurs des différentes cellules. La relation est la suivante, pour chaque cellule :

$$E\{v(n) \cdot sv'_i(n)\} = E\{u'(n) \cdot sv'_i(n)\} + E\{q'(n) \cdot sv'_i(n)\} + E\{e'(n) \cdot sv'_i(n)\} \quad (4.32)$$

En supposant que les séquences $u'(n)$, $q'(n)$, et $e'(n)$ sont non corrélées, les deux premiers termes de l'équation 4.32 valent zéro, ce qui se simplifie de la façon suivante :

$$E\{\alpha_i \cdot sv_i(n) * ETF(n) \cdot [sv_l(n) * ETF(n)]\} = \alpha_i \cdot E\{[sv'_i(n)]^2\}. \quad (4.33)$$

Par la suite, le défaut d'appariement de la cellule indice i de la valeur moyenne des M -cellules $\bar{\alpha}$ peut être estimé par la relation suivante :

$$\alpha_i \simeq \frac{E\{v(n) \cdot sv'_i(n)\}}{E\{[sv'_i(n)]^2\}} \quad (4.34)$$

Cette estimation peut être répétée de façon périodique ou plus sporadiquement. La valeur d'erreur de chacune des cellules doit être stockée dans une mémoire RAM. Ces erreurs stockées sont utilisées pour la correction de la sortie finale $v(n)$ en une valeur exacte $w(n)$. Malheureusement, l'hypothèse de non corrélation des séquences $u'(n)$, $q'(n)$, et $e'(n)$ n'est pas toujours valable. Alors, le résultat de la correction est moins bon que celui prévu par la théorie.

Afin de diminuer les corrélations les plus gênantes, quelques solutions sont proposées dans la référence [77], mais la complexité reste un obstacle pour bien bénéficier de cet algorithme.

En résumé, la méthode d'étalonnage peut être intéressante si l'OSR est très faible, par exemple $4 < OSR < 8$, mais elle est beaucoup plus complexe que les autres techniques comparables. Autrement dit, le point faible commun à toutes les méthodes de brassage des sources expliquées précédemment

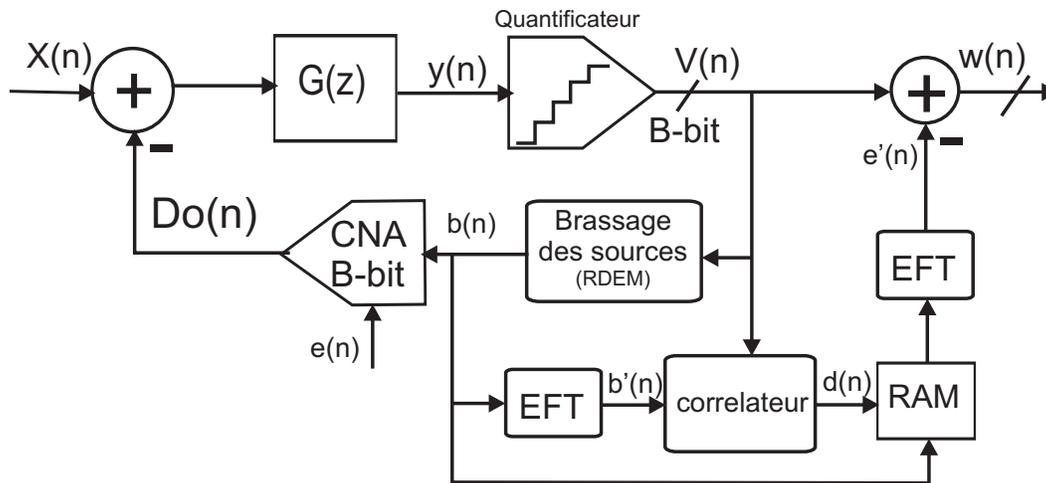


FIG. 4.45 – Deuxième technique de l'étalonnage (*background-calibration*)

dans ce chapitre, c.-à-d. inefficacité pour les faibles *OSR*, n'est pas une limite essentielle pour la méthode de l'étalonnage. Ainsi, un tel système complexe d'étalonnage, qui consomme plus de surface et de puissance, convient à la conversion très large bande. Mais cela n'est pas toujours vrai, en particulier pour les applications passe-bande dans lesquelles on ne peut abaisser l'*OSR* à cause d'autres contraintes pratiques.

Avantages de la méthode d'étalonnage :

- 1- elle fonctionne indépendamment de l'*OSR* du modulateur,
- 2- elle n'ajoute pas de retard dans la boucle,

Inconvénients de la méthode étalonnage :

- 1- elle est complexe à concevoir,
- 2- elle nécessite plus de surface et consomme plus,
- 3- elle nécessite un temps de mise en marche très long surtout pour une résolution élevée.

4.13 Conclusion

Nous avons décrit dans ce chapitre les différentes méthodes existantes de correction des erreurs du CNA multibit utilisé dans les modulateurs $\Delta\Sigma$. Parmi les différentes stratégies, nous nous sommes d'avantage attachés à étudier les algorithmes de brassage de sources les plus utilisées dans le développement de la conversion sigma delta. Durant les 14 dernières années, au moins 7 algorithmes de brassage (RDEM, CLK-DEM, ILA-DEM, DWA, SDEM, TDEM, Double-Quantification) ont été proposés, mis en application, puis améliorés peu à peu. Cependant, un algorithme général, qui présenterait à la fois tous ces avantages, ne peut pas être choisi de façon générale. L'algorithme doit être choisi selon l'application, la résolution finale, la fréquence d'échantillonnage, la technologie de réalisation, les différentes sources d'erreur et la consommation. Cela n'est pas toujours facile car il y a ni règle unique ni même de réponse unique. Le tableau 4.3, qui rassemble les caractéristiques principales des différentes méthodes de correction, peut aider à effectuer un choix correct.

TAB. 4.3 – Propriétés principal des différentes méthodes de DEM

DEM-méthode	avantages	inconvéniants	remarques :
RDEM	simple, bon SFDR	faible SNDR	< 10 bit
CLK	-	tons dépendant de l'horloge	moins intéressant
ILA	presque-1 ^{er} ordre	tons, faible performance	moins intéressant
DWA :	simple, performant	tons	meilleur DEM, <15bit
•dithered	simple	SNDR réduit	moins intéressant
•segmentée	simple	non zéro @ f=0	une meilleur solution
•Pseudo		besoin d'optimisation	
•Bi-directional		faible SNDR	moins intéressant
•Rotated	complexe	faible SFDR	
•Randomized		non général	difficile à pratiquer
•Incremented			moins intéressant
•DIA		sensible à la valeur d'entrée	difficile à pratiquer
•partielle			moins intéressant
TDEM	simple, efficace	instabilité	
SDEM	meilleure performance	complexe	c'est une référence

En conclusion, afin d'obtenir une bonne correction passe-bas du premier ordre, les meilleures méthodes sont : P-DWA (modifiée avec segmentation), ou TDEM avec un dither numérique, à la condition de posséder un OSR supérieur à 8.

En revanche, pour obtenir une mise en forme de défaut d'appariement des cellules du CNA du deuxième ordre, le seul moyen est pour l'instant la SDEM.

Dans le cas d'un modulateur passe-bande, la situation est à peu près la même sauf que la DWA n'est pas adaptée. De plus, pour une résolution finale d'environ 10 à 14 bit avec un $OSR > 32$ et $\sigma_\alpha = 1\%$, une correction d'ordre un suffit, mais pour une résolution élevée une correction de deuxième ordre est nécessaire si l'on veut bénéficier des avantages séduisants de la structure multibit.

En ce qui concerne les erreurs dynamiques, toutes les méthodes de brassage des sources dégradent le comportement, donc il faut absolument en tenir compte. Ce problème devient un obstacle majeur dans les applications haute fréquence, de sorte que n'importe quel algorithme de DEM peut dégrader la performance d'un système, même si aucune correction ne lui est appliquée. Les techniques proposées pour la limitation des erreurs dynamiques limitent fortement la performance des algorithmes essentiellement conçus pour la DEM. Seul un CNA de type RZ peut minimiser les effets secondaires imposés par une DEM arbitraire.

Finalement pour une application ayant un OSR très faible ($OSR \leq 8$, on peut en trouver dans les applications passe-bas très large bande), la technique de l'étalonnage est préférable à celle de la DEM. Elle est plus complexe en pratique, mais sa performance reste indépendante de l' OSR .

Chapitre 5

Amélioration de la linéarité du CNA, techniques proposées

5.1 Introduction

Nous venons d'étudier plusieurs méthodes de correction de la non linéarité du CNA multibit inséré dans le modulateur $\Sigma\Delta$, et en particulier les différents algorithmes de brassage des sources unitaires nommés couramment DEM¹. Chacun des algorithmes présente des avantages et des inconvénients. Mais, le principe de base reste toujours le même. La complexité des circuits nécessaires doit être raisonnable par rapport à l'amélioration amenée par rapport à un système similaire de structure monobit. La complexité des circuits devient handicapante à mesure que l'on augmente le nombre des cellules du CNA plus particulièrement pour les algorithmes d'ordre supérieur à 1. On s'efforce en général de rechercher un compromis afin d'obtenir un modulateur ayant des caractéristiques s'approchant le plus possible de celles du modulateur multibit idéal.

En utilisant les précédentes vérifications, nous présenterons au cours de ce chapitre deux algorithmes nouveaux pour la correction du premier ordre, l'un pour les systèmes passe-bas et l'autre pour les applications passe-bande. Ces deux algorithmes sont à la fois des techniques plus économiques et plus performante. Ils sont justifiées par la comparaison analytique et par les résultats des simulations au niveau système et au niveau transistor.

Ensuite, nous présenterons deux autres algorithmes de DEM d'ordre deux et même trois qui sont applicables en option passe-bas et passe-bande. Ces algorithmes sont assez prometteurs pour des modulateurs multibit de performance élevée, par exemple, pour des applications ayant une résolution de 15 à 20 bits avec un $OSR \geq 16$.

5.2 DWA-segmentée par une division mise en forme

La correction de la non linéarité, imposée par la présence de cellules désappariés dans un système suréchantillonné, peut être obtenue théoriquement par la méthode DWA [10]. Une telle méthode offre une mise en forme du premier ordre passe-bas des erreurs du CNA avec une faible complexité de

1. Appariement dynamique des composants, "Dynamic Element Matching"

réalisation. Cela fonctionne très bien en pratique sauf pour certaines entrées périodiques ou fixes, pour lesquelles des tons indésirables apparaissent dans la bande utile [120]. Afin d'adapter la méthode DWA à des entrées arbitraires, au moins neuf solutions complémentaires ont été récemment proposées; elles ont été exposé au chapitre précédent [10, 76, 84, 88, 120–123, 123, 124, 126]. Parmi ces solutions, la méthode la plus proche du cas DWA idéal est RnDWA, qui repose à la fois sur le même principe que la DWA et sur le brassage aléatoire RDEM, mais la réalisation sous forme de circuit n'est pas facile [88]. D'autres méthodes intéressantes utilisent deux indicateurs entrelacés, par exemple DIA, BiDWA et P-DWA contre un seul dans le cas DWA conventionnelle [121, 124]. Chacune de ces solutions suit le principe de DWA, mais, avec une diminution légère de sa performance afin de perturber son algorithme, qui d'ailleurs blanchit les tons non désirables. Nous avons estimé que la méthode P-DWA peut être une solution simple et proche à la méthode DWA idéale, car elle utilise deux indicateurs séparément pour deux demi DWAs idéaux [76]. Cependant, la P-DWA présente une autre faiblesse par rapport à la DWA idéal, puisqu'elle n'arrive pas mettre en forme d'éventuelles erreurs à la fréquence DC par sa définition actuelle [125].

Pour résoudre ce défaut minime de l'algorithme P-DWA, nous proposons une modification dans sa structure (figure 4.23), il y a deux modalités possibles.

Dans une première option, on fait une division aléatoire. Dans le cas simple, un tel diviseur divise l'entrée $v(n)$ par deux, et puis ajoute le reste de la division, c.-à-d. le bit LSB à une des deux sorties de manière aléatoire. Si cela n'est pas suffisant, on peut forcer le diviseur à considérer, par exemple, les deux bits les plus faibles (LSB et le suivant) comme le reste de l'opération, puis ajouter ces bits à une des sorties d'une manière aléatoire. Cette dernière stratégie permet de blanchir les fréquences indésirables de la DWA plus rapidement sans dégrader le fonctionnement des deux demi DWA.

Cette performance est justifiée, pour ce fait que les deux sous-DWA reçoivent un même signal d'entrée et ainsi, mettent en forme les erreurs à la même vitesse et avec le même cycle d'annulation. Alors que, dans d'autres techniques DWA utilisant deux indicateurs, ce cycle d'annulation est aussi divisé par deux, ce qui diminue la performance de l'algorithme. Le schéma global de cette technique est représenté à la figure 5.1.

Même si la première option résout la difficulté associée à l'algorithme P-DWA conventionnel, nous

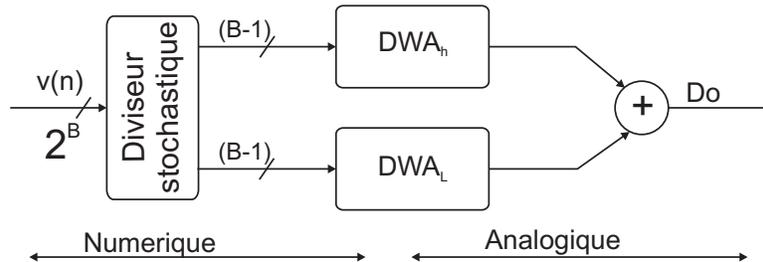


FIG. 5.1 – P-DWA modifiée, premier option qui emploie un diviseur aléatoire

préférons cependant une autre solution. Pour un même coût en complexité que pour l'option précédente, on peut remplacer le diviseur simple de P-DWA par un bloc de commutation de mise en forme

SB_{LP} , comme l'illustre la figure 5.2. Il s'agit d'un bloc de commutation passe-bas, qui est semblable

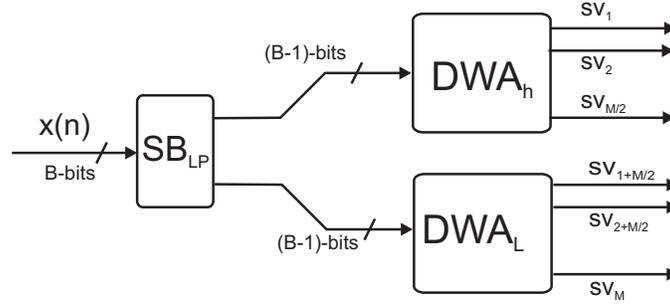


FIG. 5.2 – Topologie générale proposée pour une B -bits MP-DWA

au bloc $SB_{k,r}$ doté d'un algorithme TDEM passe-bas, mais peut fonctionner aussi différemment (voire la figure 5.3). Théoriquement, cette modification conduit à une meilleure performance, car la diffé-

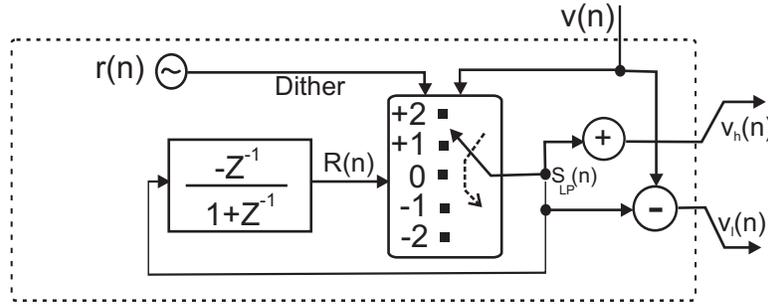


FIG. 5.3 – Bloc de commutation passe-bas qui fonctionne comme un diviseur spécial

rence entre les gains des deux parties, c.-à-d. $\delta_\alpha = \overline{\alpha_h} - \overline{\alpha_l}$, est également mise en forme². Le bloc SB_{LP} divise l'entrée $v(n)$ en deux sorties $v_h(n)$ et $v_l(n)$ par les relations suivantes :

$$\begin{aligned} v_h(n) &= \frac{v(n) + s_{lp}(n)}{2} \\ v_l(n) &= \frac{v(n) - s_{lp}(n)}{2} \end{aligned} \quad (5.1)$$

La séquence $s_{lp}(n)$ est instantanément générée à l'intérieur du bloc à chaque période. La fonction de transfert du bloc SB_{LP} peut être choisie selon les équations 4.26, 4.29 ou 4.30. En fait, chacune des fonctions de transfert donne un très bon résultat, on choisit alors la plus simple (équation 5.2).

$$S_{LP}(n) = \begin{cases} +1 & v(n) \text{ est impair et } R(n) > 0 \\ r(n) & v(n) \text{ est impair et } R(n) = 0 \\ -1 & v(n) \text{ est impair et } R(n) < 0 \\ 0 & \text{dans tous les autres cas} \end{cases} \quad (5.2)$$

2. où, $\overline{\alpha_h} = \sum_{i=1}^{\frac{M}{2}} \alpha_i$ et $\overline{\alpha_l} = \sum_{(i=1+\frac{M}{2})}^M \alpha_i$.

Cependant, dans le cas exceptionnel pour laquelle l'entrée du CNA reste plutôt aux alentours de $\frac{M}{2}$, $\frac{M}{4}, \dots$, il vaut mieux la considérer comme égale à :

$$S_{LP}(n) = \begin{cases} +1 & v(n) & \text{est impair et } R(n) > 0 \\ r(n) & v(n) & \text{est impair et } R(n) = 0 \\ -1 & v(n) & \text{est impair et } R(n) < 0 \\ +2 & 0 < v(n) < M & \text{est pair et } R(n) > 0 \\ -2 & 0 < v(n) < M & \text{est impair et } R(n) < 0 \\ 0 & & \text{dans tous les autres cas} \end{cases} \quad (5.3)$$

où, la séquence $r(n)$ est un bruit imposé ("dither") équivalent à un bit aléatoire traduit comme "-1" pour le cas " $r(n) = 0$ ", ou "1" pour le cas " $r(n) = 1$ ". Ce dither est nécessaire dans toutes les boucles de $\Sigma\Delta$ du premier ordre ou dans les cas similaires comme ceux de DEM. On préfère quand même la fonction de transfert simple donnée par l'équation 5.2, car dans un modulateur sigma delta, la probabilité d'appliquer une entrée constante au niveau du CNA interne est très faible, même dans le cas où l'entrée principale du modulateur est fixe [2]. L'équation 5.3 peut être intéressante si le défaut d'appariement contient des erreurs systématiques (non aléatoire) ce qui peut augmenter la différence entre des gains moyens des deux demis CNAs $\delta_\alpha = \bar{\alpha}_h - \bar{\alpha}_l$. Le brassage des sources obtenu par ces modifications, qui est basé sur la méthode DWA-segmentée, est dénoté MP-DWA dans ce document, ("*Modified partitioned Data Weighted Averaging*").

Afin de mettre en équation le fonctionnement de la MP-DWA, on part des équations 4.7 et 4.9 et du modèle de la cellule normalisée du CNA sans offset, c.-à-d. $do_i(nT) = sv_i(n) \cdot \alpha_i$. Pour la structure générale décrite à la figure 5.2, nous obtenons alors :

$$\begin{aligned} Do_h(n) &= \sum_{i=1}^{\frac{M}{2}} sv_i(n) \cdot (1 + \alpha_i) = \sum_{i=Ptr_h(n-1)}^{Ptr_h(n)-1} sv_i(n) \cdot (1 + \alpha_i) \\ Do_l(n) &= \sum_{i=1+\frac{M}{2}}^M sv_i(n) \cdot (1 + \alpha_i) = \sum_{i=\frac{M}{2}+Ptr_l(n-1)}^{\frac{M}{2}+Ptr_l(n)-1} sv_i(n) \cdot (1 + \alpha_i) \end{aligned}$$

quand :

$$\begin{aligned} Ptr_h(n) &= Ptr_h(n-1) + v_h(n), \quad \text{modulo } \frac{M}{2} \\ Ptr_l(n) &= Ptr_l(n-1) + v_l(n), \quad \text{modulo } \frac{M}{2} \end{aligned} \quad (5.4)$$

où, $sv_h(n) = \{sv_1(n), sv_2(n), \dots, sv_{\frac{M}{2}}(n)\}$ et $sv_l(n) = \{sv_{1+\frac{M}{2}}(n), sv_{2+\frac{M}{2}}(n), \dots, sv_M(n)\}$ sont deux vecteurs sortants, des deux moitiés DWA_h et DWA_l . La différence entre la sortie réelle et la droite idéale adoptée à la sortie de chaque CNA, $e(n) = INL(n) = Do(nT) - v(n) \cdot (1 + \bar{\alpha})$, est donnée par les relations suivantes pour ces demi-CNA :

$$e_h(n) = \sum_{i=Ptr_h(n-1)}^{Ptr_h(n)-1} \alpha_i - \bar{\alpha}_h \cdot v_h(n) = \sum_{i=Ptr_h(n-1)}^{Ptr_h(n)-1} \alpha_i - \bar{\alpha}_h \cdot \left[\frac{v(n) + S_{LP}(n)}{2} \right]$$

$$e_l(n) = \sum_{i=\frac{M}{2}+Ptr_l(n-1)}^{\frac{M}{2}+Ptr_l(n)-1} \alpha_i - \overline{\alpha}_l \cdot v_l(n) = \sum_{i=\frac{M}{2}+Ptr_l(n-1)}^{\frac{M}{2}+Ptr_l(n)-1} \alpha_i - \overline{\alpha}_l \cdot \left[\frac{v(n) - s_{LP}(n)}{2} \right] \quad (5.5)$$

où, les valeurs équivalentes des $v_h(n)$ et $v_l(n)$ sont remplacées à l'aide de l'équation 5.1. Alors, l'erreur de la sortie finale du CNA est obtenue par la somme des deux parties de l'expression 5.5,

c.-à-d. $e(n) = e_h(n) + e_l(n)$:

$$e_{MP-DWA}(n) = \sum_{i=Ptr_h(n-1)}^{Ptr_h(n)-1} \alpha_i + \sum_{i=\frac{M}{2}+Ptr_l(n-1)}^{\frac{M}{2}+Ptr_l(n)-1} \alpha_i - v(n) \left[\frac{\overline{\alpha}_h + \overline{\alpha}_l}{2} \right] - s_{LP}(n) \left[\frac{\overline{\alpha}_h - \overline{\alpha}_l}{2} \right] \quad (5.6)$$

On peut comparer le résultat avec celle du cas DWA conventionnel ayant $\frac{M}{2}$ cellules et une entrée arbitraire, c.-à-d. $e_{\frac{M}{2}}(n) = \sum_{i=Ptr(n-1)}^{Ptr(n)-1} \alpha_i - \overline{\alpha}_{\frac{M}{2}} \cdot x(n)$, dans laquelle le premier terme qui correspond à la somme des α_i est toujours mis en forme par l'évaluation dynamique de l'indicateur $Ptr(n)$ et le deuxième terme $\overline{\alpha}_{\frac{M}{2}}$ peut être toujours vu comme une erreur linéaire du gain ne produisant pas de distorsion.

Par ailleurs, on remarque que le seul terme gênant dans le cas P-DWA est $s_{LP}(n) \left[\frac{\overline{\alpha}_h - \overline{\alpha}_l}{2} \right]$, où la différence entre des gains moyens des deux demi CNA est associée par la séquence $s_{LP}(n)$. En général cette différence est plus petite que α_i si le défaut d'appariement des cellules est aléatoire. Cependant, dans le cas d'erreurs d'appariement systématiques, cette différence moyenne est considérable. En pratique, un CNA réel présente deux types d'erreurs concomitantes, on ne peut donc pas négliger le terme concernant $s_{LP}(n)$. En définitive, pour que le spectre de la sortie d'un P-DWA soit le plus propre possible, le terme $s_{LP}(n)$ doit aussi être mis en forme.

Dans l'algorithme conventionnel de P-DWA [76, 125, 126, 145] ce problème n'est pas résolu, car la séquence $s_{LP}(n)$ est supposée toujours fixe, par exemple, $s_{LP}(n) = 1$. Alors, comme c'est analysé dans [125], la P-DWA possède un spectre non nul à la fréquence nulle, ce qui peut être négligé seulement pour les systèmes destinés aux applications à résolution modérée à la condition que le défaut d'appariement soit aléatoire.

Par ailleurs, avec la première option proposée ci-dessus, $s_{LP}(n)$ est supposé être une séquence aléatoire, ne produisant donc pas de ton parasite. Cependant, son spectre ne s'annule pas à la fréquence nulle. En revanche, dans la deuxième proposition qui a notre préférence, $s_{LP}(n)$ est une séquence qui est mise en forme au moyen d'un filtre passe-bas de premier ordre $(1 - z^{-1})$ et par la fonction de transfert indiquée à l'équation 5.2. Une telle mise en forme est tout à fait compatible avec les fonctions de transfert de chaque demi DWA validées par l'équation 4.9. Cela signifie que l'ensemble des erreurs du CNA est parfaitement mise en forme au premier ordre, comme le montre la figure 5.4, où $OSR=64$, $M=16$, et $\sigma_\alpha = 1\%$.

En ce qui concerne la résolution totale du système, idéalement l'erreur du CNA peut être estimée en utilisant l'équation 4.12 en négligeant les tons parasites, car ceux-ci sont éliminés par l'application de la méthode P-DWA et puis, le défaut de gain entre les deux parties est résolu par la mise en forme de $s_{LP}(n)$ au moyen de l'algorithme MP-DWA. Le coût en est la diminution de la résolution finale d'un demi-bit comme nous le montrerons ci-après.

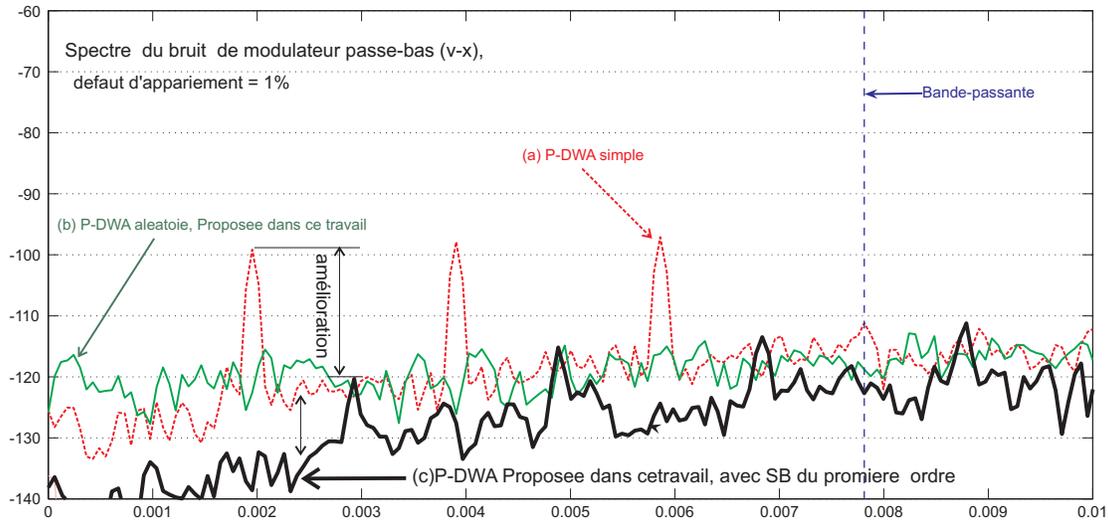


FIG. 5.4 – Spectre du modulateur passe-bas multibit, $M=16$, $\sigma_\alpha = 1\%$: a) méthode P-DWA conventionnelle, b) P-DWA avec un diviseur aléatoire (premier option proposée), c) méthode MP-DWA proposé dans cette section

La résolution finale concernant la MP-DWA peut être exprimée par l'équation -

$ENOB = \text{Log}_2 \left[\frac{M}{2\sqrt{3}N_{total}} \right]$, où l'ensemble des erreurs du CNA est donné par la suite :

$$\begin{aligned}
 e_{CNA}(n) &= e_h(n) + e_l(n) = \sum_{ptr_h(n-1)+1}^{ptr_h(n)} (\alpha_i - \bar{\alpha}_h) + \sum_{\frac{M}{2}+ptr_l(n-1)+1}^{\frac{M}{2}+ptr_l(n)} (\alpha_i - \bar{\alpha}_l) \\
 &= \sum_{ptr_h(n-1)+1}^{ptr_h(n)} \alpha_i + \sum_{\frac{M}{2}+ptr_l(n-1)+1}^{\frac{M}{2}+ptr_l(n)} \alpha_i - v_h(n)\bar{\alpha}_h - v_l(n)\bar{\alpha}_l \\
 &= \sum_{ptr_h(n-1)+1}^{ptr_h(n)} \alpha_i + \sum_{\frac{M}{2}+ptr_l(n-1)+1}^{\frac{M}{2}+ptr_l(n)} \alpha_i - v(n)[\bar{\alpha}_h + \bar{\alpha}_l] - s_{LP}(n) \left[\frac{\bar{\alpha}_h - \bar{\alpha}_l}{2} \right] \quad (5.7)
 \end{aligned}$$

Dans cette équation les deux premiers termes ne sont pas en général indépendants, ainsi que les deux derniers qui interdisent une simplification de l'équation³.

Par chance, le dernier terme est assez petit devant les deux premiers, il a donc peu d'influence sur le résultat final :

- dans notre cas, c.-à-d. MP-DWA, le terme $s_{LP}(n)$ est mis en forme afin d'éliminer sa puissance en bande utile,
- puis, la partie $\frac{\bar{\alpha}_h - \bar{\alpha}_l}{2}$ est \sqrt{M} fois plus petite que σ_α présent dans les deux premiers termes.

Par conséquent, on peut supposer que le bruit imposé dans la bande est la somme statistique du bruit des demi CNA, en ne faisant pas intervenir leur corrélation. Cette estimation optimiste ne

3. Il est possible qu'on cherche une expression pour *auto-correlation* (e_{CNA}) à l'aide de certaines approximations, mais comme elle devient assez complexe sans donner un sens correct, nous préférons, ici, une discussion simple qui est valide aux cas extrêmes. Cela peut être raisonnable, car, les deux extrêmes (max et min) sont peu éloignées.

donne qu'une valeur estimée pour ENOB qui s'exprime par la relation :

$$\begin{aligned}
e_{h,rms}^2 &= \frac{\sigma_\alpha^2 \cdot (1 - \frac{2}{M})^2 \cdot \frac{M}{2}}{12} \\
e_{l,rms}^2 &= \frac{\sigma_\alpha^2 \cdot (1 - \frac{2}{M})^2 \cdot \frac{M}{2}}{12} \\
e_{rms} &= \sqrt{e_{h,rms} + e_{l,rms}} = \sqrt{2} \left[\frac{\sigma_\alpha \cdot (1 - \frac{2}{M}) \cdot \sqrt{\frac{M}{2}}}{12} \right]
\end{aligned} \tag{5.8}$$

où, la valeur efficace de chaque demi algorithme est estimée à la référence [89]. Ensuite, la résolution finale du CNA peut être obtenue en appliquant l'effet de la mise en forme au premier ordre $(1 - z^{-1})$, c.-à-d. $e_{rms,BW} = e_{rms} \cdot \frac{\pi}{\sqrt{3 \cdot OSR^3}}$ [1].

$$\begin{aligned}
ENOB_{MP-DWA,max} &= \text{Log}_2 \left[\frac{M}{2\sqrt{3} \cdot Noise_{total,rms}} \right] = \text{Log}_2 \left[\frac{M \cdot \sqrt{3 \cdot OSR^3}}{2\sqrt{3} \cdot \sqrt{2} \cdot \pi \left(\frac{\sigma_\alpha \cdot (1 - \frac{2}{M}) \cdot \sqrt{\frac{M}{2}}}{\sqrt{12}} \right)} \right] \\
&= ENOB_{DWA} + \text{Log}_2 \left[\frac{(M-1)}{(M-2)} \right]
\end{aligned} \tag{5.9}$$

Cela est intéressant, parce que l'on obtient une résolution légèrement meilleure que celle obtenu au moyen de l'algorithme DWA conventionnel, surtout pour les petits M . Par exemple, pour $M=8$, c.-à-d. deux CNA de 4-cellules, la résolution finale peut être améliorée de 0.22 bit. En plus, chaque demi P-DWA est moins complexe qu'un algorithme DWA conventionnelle.

Ce résultat, bien que simple n'est pas acceptable car les deux parties des erreurs produites par les deux demi CNA semblent être bien corrélées, car leurs entrées sont assez corrélées $\frac{v(n) \pm s_{LP}(n)}{2}$. L'observation pratique confirme bien cette dépendance entre deux indicateurs avec un coefficient de l'ordre de 80%. Par conséquent, la résolution minimale peut être estimée si on utilise la fonction $e_{rms,max} \leq \sqrt{e_{h,rms} + e_{l,rms}} = 2 \cdot e_{h,rms}$, c'est-à-dire complètement corrélée.

$$\begin{aligned}
ENOB_{MP-DWA,min} &\geq \text{Log}_2 \left[\frac{\sqrt{3M \cdot OSR^3}}{\sqrt{2} \cdot \sigma_\alpha \cdot (1 - \frac{2}{M})} \right] \\
&= ENOB_{DWA} + \text{Log}_2 \left[\frac{(M-1)}{(M-2)} \right] - 0.5 \text{ , bit}
\end{aligned} \tag{5.10}$$

Alors, le ENOB est diminué d'environ 0.5 bit soit moins de -3dB, ce qui peut être compensé en partie par le terme $\text{Log}_2 \left[\frac{(M-1)}{(M-2)} \right]$:

$$(ENOB_{DWA,ideal} - 0.5bit) < ENOB_{MP-DWA} \leq ENOB_{DWA,ideal} \tag{5.11}$$

Différentes simulations montrent que le SNDR est diminué d'environ 2dB (soit 0.33 bit) avec l'algorithme MP-DWA par rapport l'algorithme DWA idéal, ce qu prévoit l'équation 5.11. En définitive, les équations 5.10 et 5.11 indiquent que la résolution d'une technique MP-DWA peut être toujours estimée par celle de la DWA conventionnelle.

En revanche, le SFDR est amélioré d'au moins 25dB par rapport à l'algorithme DWA, car les fréquences parasites ne sont pas générées par un algorithme MP-DWA comme le montre la figure 5.4.

De plus, la condition limitant de l'équation 4.14 n'intervient pas, car avec l'algorithme MP-DWA les demi DWA ne voient jamais une entrée périodique. En tous cas, si on la vérifie pour une éventuelle entrée périodique, on voit que la limite d'OSR est bien améliorée d'un facteur 2. C'est une voie prometteuse pour les applications large-bande. Elle s'exprime par les relations suivantes, où $k \in \{1,2,3,\dots\}$ est le nombre d'harmoniques :

$$\begin{aligned} OSR_{DWA,ideal} &> \frac{M(k+1)}{2}, \\ OSR_{MP-DWA} &> \frac{M(k+1)}{4} \end{aligned} \quad (5.12)$$

Par conséquent :

$$Minimum[OSR_{MP-DWA}] = 2 \times Minimum[OSR_{DWA}] \quad (5.13)$$

On peut remarquer que la mise en forme de l'erreur associée à $s_{LP}(n)$ peut aussi être d'ordre supérieur à un, comme c'est possible avec la configuration TDEM. Mais cela n'est pas nécessaire, car le reste du système de MP-DWA fonctionne à l'ordre 1.

5.3 DWA modifiée aux applications de DEM passe-bande

L'étude bibliographique montre que la réalisation d'une méthode DEM passe-bande est peu fréquente [82,83,90,129,146–148]. En plus, les réalisations existantes sont, en général, du second ordre et utilisent majoritairement la méthode SDEM⁴. Il n'y a pas d'algorithmes diversifiés comme pour le cas passe-bas. En gros, les deux méthodes, SDEM et TDEM, sont adaptés aux applications passe-bande, mais ce ne sont pas des solutions simples et rapides. Dans les références [31, 90, 129, 146, 148], un algorithme DWA modifié est proposé pour des applications passe-bande. Peut-être est-ce la méthode de correction passe-bande la plus simple. Cependant par sa définition actuelle, elle est très sensible aux signaux d'entrée périodiques comme des signaux de porteuse qui sont souvent présents dans ce type d'applications.

La figure 5.5 montre la performance d'un CNA 4-bits ayant 1% d'erreur qui emploie une telle méthode, c.-à-d. DWA passe-bande. Les tons parasites ont dégradé la performance, comme dans le cas du DWA passe-bas conventionnel.

Dans cette section, nous détaillons cet algorithme et proposons une modification permettant d'obtenir une technique de DEM au second ordre passe-bande proche de l'idéal.

Afin d'expliquer le fonctionnement d'un algorithme DWA passe-bande, on peut utiliser la transformation linéaire du passe-bas en passe-bande au quart de la fréquence d'échantillonnage, c.-à-d. $z^{-1} \iff -z^{-2}$. En appliquant cette transformation à l'équation 4.9, pour le cas passe-bande on obtient :

$$E_{LP}(z) = (1 - z^{-1})INL[Ptr(z)] \iff E_{BP}(z) = (1 + z^{-2})INL[Ptr(z)]$$

4. Il faut bien noter que un DEM passe-bande centrée au quart de la fréquence d'échantillonnage ($\frac{f_s}{4}$) du second ordre peut être équivalente à une DEM passe-bas du premier ordre.

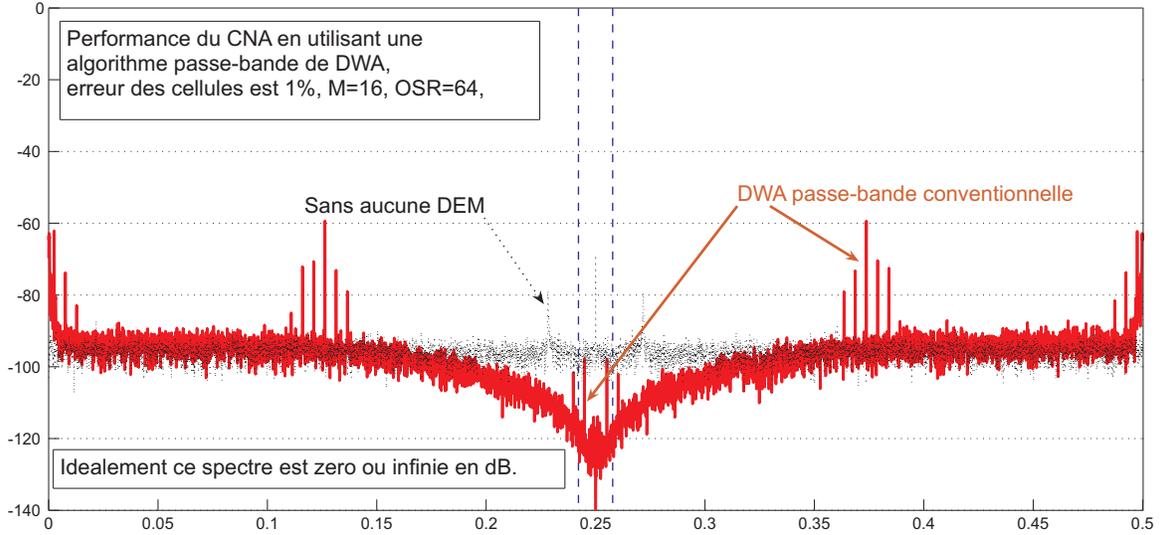


FIG. 5.5 – Performance d'un CNA 4-bits ayant 1% de défaut d'appariement qui est corrigée par une méthode de DWA passe-bande conventionnelle.

En temps discret, la relation suivante s'écrit :

$$e_{LP}(n) = \sum_{i=Ptr_{LP}(n-1)}^{Ptr_{LP}(n)-1} sv_i(n) \cdot (1 + \alpha_i) \iff e_{BP}(n) = \sum_{i=Ptr_{BP}(n)}^{Ptr_{BP}(n)-1} sv_i(n) \cdot (1 + \alpha_i)$$

L'étape suivante consiste à savoir comment l'indicateur $Ptr_{BP}(n)$ doit être évalué pour une application passe-bande. Dans le cas passe-bas, une somme modulo M convient.

$$Ptr_{LP}(n) = [Ptr_{LP}(n-1) + v(n)] , \quad \text{modulo } M \quad (5.14)$$

Une première idée consiste à utiliser une équation similaire en appliquant le sens de la transformation $z^{-1} \iff -z^{-2}$ dans le domaine temporel :

$$Ptr_{BP}(n) = [Ptr_{BP}(n-2) \mp v(n)] , \quad \text{modulo } M \quad (5.15)$$

Il y aura donc deux réponses : une solution complète comprendra deux indicateurs.

Une autre méthode consiste à supposer qu'un résonateur est composée de deux intégrateurs entrelacés en orientation inverse. Par conséquent, on doit utiliser deux indicateurs dont le signe d'intégration est changé à chaque opération correspondante.

Cependant, l'introduction d'une étape intermédiaire facilite la compréhension. Examinons d'abord comment le terme passe-haut $(1 + z^{-1})$ peut être obtenu. Dans un système de traitement passe-haut, les signaux de fréquences élevées sont privilégiées. Alors, afin d'obtenir une DEM passe-haut, il faut que les cellules soient sélectionnées parmi celles qui sont récemment utilisées. Par conséquent, l'indicateur passe-haut peut être évalué comme suit :

$$Ptr_{HP,(mod:M)}(n) = \begin{cases} Ptr_{HP}(n-1) - v(n) & \text{de la gauche à la droite quand } n \text{ est pair} \\ Ptr_{HP}(n-1) + v(n) & \text{de la droite à la gauche quand } n \text{ est impair} \end{cases}$$

Cela signifie que Ptr_{HP} doit reculer à chaque fois pour choisir les cellules récemment sélectionnées. Autrement dit, la direction de la progression de l'indicateur doit être changée à chaque utilisation. Afin de réaliser une DEM passe-bande $(1 + z^{-2})$ une transformation complémentaire de double voies $z^{-1} \implies z^{-2}$ doit être appliquée. La figure 5.6 illustre un modèle de sélection et le passage du passe-haut au passe-bande [129]. Cet algorithme est le plus simple outil permettant la mise en forme du

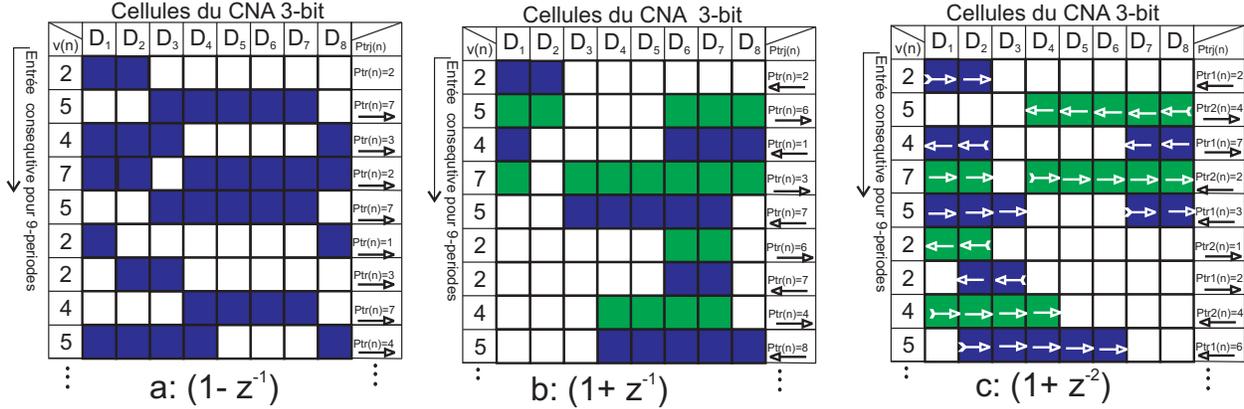


FIG. 5.6 – Sélection des éléments du CNA en passant de passe-bas à passe-bande : a) passe-bas du premier ordre, b) passe-haut du premier ordre, c) passe-bande du deuxième ordre.

défaut d'appariement des cellules du CNA ordre passe-bande du second ordre.

En ce qui concerne la mise en équation, un algorithme P-DWA passe-bande a idéalement la même performance que l'algorithme similaire passe-bas (P-DWA), laquelle est donnée par les équations 5.8-5.13.

Un algorithme DWA passe-bande présente aussi des fréquences parasites. Afin d'éloigner ces fréquences, nous utilisons la méthode de partition des cellules en deux groupes séparés, connue selon le nom de méthode de P-DWA. Cependant, un système P-DWA passe-bande souffre de la présence de la différence entre le gain moyen des ses deux sous-algorithmes $(\bar{\alpha}_h - \bar{\alpha}_l)$, comme dans le cas passe-bas. De plus, un algorithme P-DWA passe-bande simple n'est pas vraiment insensible aux entrées périodiques. La technique de la partition par un diviseur simple ($s(n) = 1$) fait diminuer le comportement vis à vis des fréquences parasites DWA, mais, cela n'est pas encore suffisant, comme le montrent les simulations.

Une P-DWA passe-bande demande donc à être modifiée. Une stratégie analogue à celle du passe-bas est montrée à la figure 5.7.

Maintenant, on doit appliquer une fonction de transfert complexe comme celle de l'équation 5.3 plutôt qu'une fonction de transfert simple. Si l'OSR est faible, on préfère même une mise en forme d'ordre deux pour le diviseur passe-bande ce qui permet d'éloigner simplement les éventuelles fréquences parasites de la bande utile. La figure 5.8 montre un exemple d'application de l'algorithme proposé (MP-DWA passe-bande) quand $OSR=64$, $\sigma_\alpha = 1\%$ et $M=16$.

Si on compare la complexité d'un algorithme P-DWA ou MP-DWA avec un algorithme de TDEM pour les applications passe-bande, on peut remarquer que la TDEM a un degré de complexité moindre

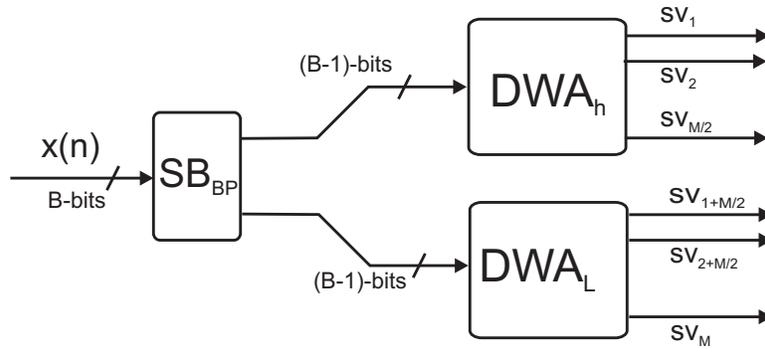


FIG. 5.7 – Structure proposé pour le brassage des sources (DEM) de second ordre passe-bande, MP-DWA

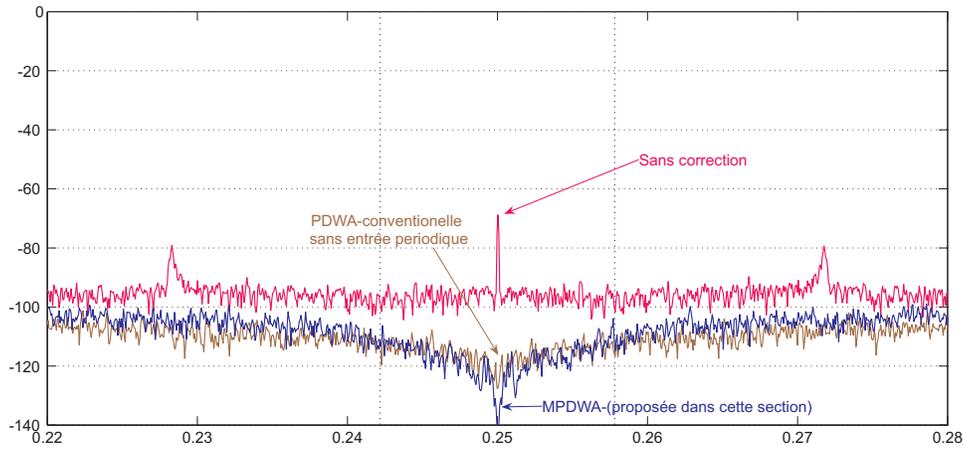


FIG. 5.8 – Performance d’un CNA 4-bits ayant 1% de défaut d’appariement qui est corrigée par une méthode de MP-DWA passe-bande.

que P-DWA ou MP-DWA quand M est petit ($M \leq 4$). Bien sûr, ceci est valable à condition que les deux algorithmes aient la même performance, ce qui n’est pas pratiquement démontré. Dans tous les cas, le bénéfice de la MP-DWA est patent quand le nombre des cellules est grand ($M \geq 8$).

5.4 Brassage des sources en structure d’arbre segmenté, (Mixed structured, MDEM)

Comme on l’a déjà dit, parmi les nombreuses méthodes développées pour le brassage des sources (DEM), seulement deux algorithmes SDEM et TDEM peuvent fonctionner à un ordre supérieur à un. Les autres méthodes, par exemple RDEM, CLA, ILA et DWA, ont beaucoup de mal et ne peuvent même pas être modifiées et utilisées en tant que DEM de deuxième ordre ou supérieur. Cela vient du fait que ces méthodes emploient en réalité un algorithme cyclique qui peut donner lieu parfois à une mise en forme de premier ordre et qu’elles ne contiennent pas de vrai filtre numérique. Par ailleurs, les techniques SDEM et TDEM emploient de vrais filtres numériques, qu’ils soient passe-bas, passe-haut ou passe-bande. On peut, concevoir des filtres avec les caractéristiques souhaitées en terme de bande

passante et d'ordre arbitraire. Une telle possibilité (DEM d'ordre élevé) peut être utile quand :

- le CNA présente un taux élevé de défaut d'appariement entre ses cellules, par exemple à partir de 1 pourcent ($\sigma_\alpha > 1\%$),
- le système a besoin d'une précision élevée, par exemple dans un modulateur $\Sigma\Delta$ multibit ayant une linéarité plus de 16 bits,
- le système a un taux normal d'erreurs du CNA ($\sigma_\alpha < 1\%$) et une résolution moyenne ($12 < ENOB \leq 16$ bit), mais une large bande passante ($OSR < 16$).

De nos jours, les technologies d'intégration des systèmes électroniques ont bien progressé, ce qui entraîne une diminution du taux d'erreur des circuits. Cependant, les deux derniers cas mentionnés ci-dessus sont fortement probables. Ainsi, il est raisonnable de développer certaines méthodes de DEM d'ordre élevé et/ou de bande passante large.

Malheureusement, chacun des deux algorithmes originaux SDEM et TDEM ont des faiblesses, de sorte qu'ils ne sont pas utilisables pour toutes les applications, comme on en a discuté au chapitre précédent. D'une part, la TDEM conventionnelle a l'avantage de la simplicité mais présente une forte instabilité à partir du premier ordre, et il est quasi inutilisable si son algorithme n'est pas modifié. Même si des modifications intéressantes ont été récemment présentées [133], le problème principal qui réside dans les deux premières couches de TDEM n'a jamais été résolu. Le SDEM, qui par ailleurs a une meilleure stabilité, devient trop complexe à mesure que le nombre des cellules du CNA augmente. Nos études montrent qu'il est possible de surmonter ces limites en utilisant les deux algorithmes afin d'en composer un nouveau.

Dans cette section, nous présenterons, un nouvel algorithme qui présente les avantages à la fois de la simplicité de la TDEM et de la stabilité de l'algorithme SDEM [132]. Il est présenté dans les deux prochaines sous-sections. Cet algorithme mixe les deux méthodes précédemment analysées, c.-à-d. TDEM et SDEM [11, 12]. Le résultat est appelé MDEM ("*Mixed structured Dynamic Element Matching*") à cause du mélange de deux algorithmes. De plus, afin on examinera le comportement de MDEM en donnant une expression mathématique unifiée.

5.4.1 Première option de l'algorithme MDEM

Puisque l'instabilité de la TDEM réside principalement dans les première et deuxième couches, nous proposons d'employer un algorithme de tri (SDEM) dans tous les blocs de commutation des deux dernières couches ($SB_{k,r}$, $k=1,2$). Afin d'améliorer la stabilité de l'algorithme dans les couches suivantes ($k \geq 3$), nous proposons d'employer des règles de décision semblables aux équations 4.26 et 4.29, mais sans aucune restriction vis-à-vis de l'entrée du bloc, qui peut être paire ou impaire comme suit :

$$S_{LP}(n) = \begin{cases} +1 & v(n) & \text{est impair et } R(n) > 0 \\ r(n) & v(n) & \text{est impair et } R(n) = 0 \\ -1 & v(n) & \text{est impair et } R(n) < 0 \\ +2 & 0 < v(n) < M & \text{est pair et } R(n) > 0 \\ -2 & 0 < v(n) < M & \text{est impair et } R(n) < 0 \\ 0 & & \text{dans tous les autres cas} \end{cases} \quad (5.16)$$

Il n'y a aucune restriction sérieuse dans la nouvelle définition des règles de décision (5.16). Ainsi, le fonctionnement produit veut être stable pour toutes les couches à l'exception de la première où l'équation 5.16 ne peut pas être appliquée.

Le schéma pratique d'un tel algorithme est montré à la figure 5.9. Ici, l'algorithme de TDEM est

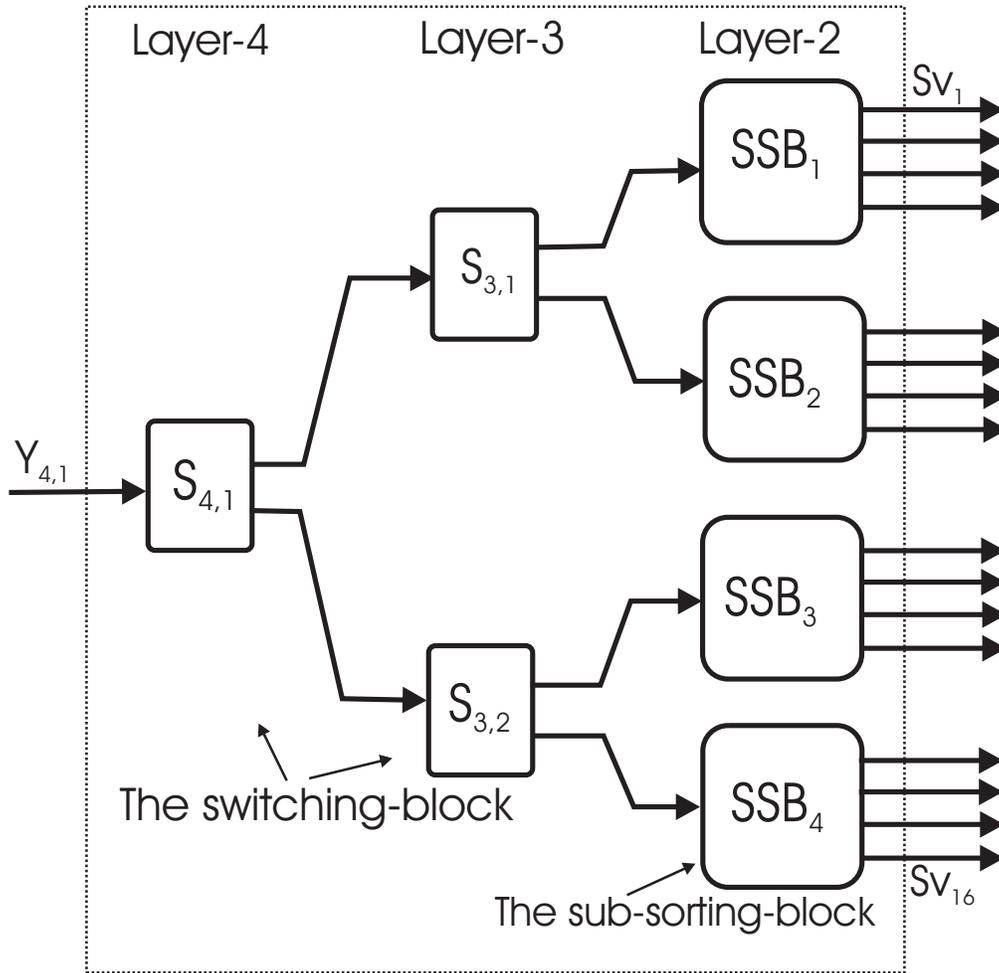


FIG. 5.9 – La structure proposé pour l'algorithme MDEM pratique, $M=16$.

bien employé à condition qu'il n'y ait pas les quatre cellules du CNA à commander. Dans ce cas, nous proposons un sous bloc de tri "Sub-Sorting Block (SSB)", permettant de commander localement chaque groupe de quatre cellules à l'aide de l'algorithme SDEM. En raison de la nature exempte de contraintes de l'algorithme SDEM en terme d'entrées paires ou impaires et en raison de l'application de l'équation 5.16 à la partie arborescente, cet algorithme a toutes les raisons d'être moins sensible aux caractéristiques stochastiques de l'entrée, offrant plus de stabilité et plus de performances.

L'algorithme proposé (MDEM) se compose de deux algorithmes différents: le fonctionnement simultané correct de ces deux algorithmes est critique. Notre but est de montrer comment un bloc SDEM partiel peut être compatible avec un système global TDEM. La structure de base de TDEM est décrite par l'équation 4.27, le mélange des deux algorithmes peut être justifié si les parties remplacées produisent la même fonction que dans la TDEM pure. Autrement dit, chacun des blocs SSB_r , repré-

sentés à la figure 5.9 doit avoir la même fonction qu'avec une TDEM appliqué à ses quatre cellules. A partir de l'équation 4.27 décrivant une TDEM groupant les 4 dernières cellules ($B = 2$), on obtient l'expression suivante :

$$\begin{aligned}
Do_r(n) &= \sum_{i=1}^4 do_i(n) = (1 + \overline{\alpha_r}) y_r(n) + e_{TDEM_r}(n) + \epsilon_r \\
\text{où: } \alpha_r &= \frac{1}{4} \sum_{i=1}^4 \alpha_i \quad , \quad \epsilon_r = \sum_{i=1}^4 \epsilon_i \\
e_{TDEM_r}(n) &= \sum_{k=1}^2 \sum_{r=1}^{2^{2-k}} \Delta_{kr} S_{kr}(n) \\
&= \Delta_{11} S_{11} + \Delta_{12} S_{12} + \Delta_{21} S_{21} \\
\Delta_{11} &= (\alpha_1 - \alpha_2)/2 \quad , \quad \Delta_{12} = (\alpha_3 - \alpha_4)/2 \\
\Delta_{21} &= (\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4)/4 \\
S_{11}(n) &= sv_1(n) - sv_2(n) \quad , \quad S_{12}(n) = sv_3(n) - sv_4(n) \\
S_{21} &= [sv_1(n) + sv_2(n)] - [sv_3(n) + sv_4(n)] \tag{5.17}
\end{aligned}$$

D'autre part, les équations décrivant l'algorithme SDEM pour les 4 cellules (SSB_r), peuvent s'écrire comme suit :

$$\begin{aligned}
Do_r(n) &= (1 + \alpha_1)SV_1(n) + \epsilon_1 + (1 + \alpha_2)SV_2(n) + \epsilon_2 \\
&\quad + (1 + \alpha_3)SV_3(n) + \epsilon_3 + (1 + \alpha_4)SV_4(n) + \epsilon_4 \tag{5.18}
\end{aligned}$$

Après quelques simplifications et en utilisant la même notation que dans l'équation 5.17, nous pouvons réécrire :

$$Do_r(n) = (1 + \alpha_r) y_r(n) + e_{SDEM}(n) + \epsilon_r \tag{5.19}$$

$$\begin{aligned}
e_{SDEM}(n) &= \Delta_{11}(SV_1(n) - SV_2(n)) + \Delta_{12}(SV_3(n) - SV_4(n)) \\
&\quad + \Delta_{21}(SV_1(n) - SV_3(n)) + \Delta_{21}(SV_2(n) - SV_4(n)) \tag{5.20}
\end{aligned}$$

En utilisant l'équation 4.18, le terme d'erreur de non linéarité du CNA ayant 4 éléments e_{SDEM} est donné par :

$$\begin{aligned}
E_{SDEM}(z) &= \Delta_{11}H(z)[Se_1(z) - Se_2(z)] + \Delta_{12}H(z)[Se_3(z) - Se_4(z)] \\
&\quad + \Delta_{21}H(z)[Se_1(z) - Se_3(z)] + \Delta_{21}H(z)[Se_2(z) - Se_4(z)] \tag{5.21}
\end{aligned}$$

Ces relations (5.17 et 5.21) sont comparées sur le tableau 5.1, d'où nous tirons deux observations importantes :

- les erreurs de gain et d'offset (c.-à-d. $\overline{\alpha_r}$ et ϵ_r , qui ne sont pas mis en forme, ni par la TDEM et ni par la SDEM), sont identiques.
- toutes les erreurs de non linéarité modulées par l'entrée du bloc sont bien mises en forme par des filtres $H(z)$ ou par les séquences $S_{k,r}(z)$ du même ordre.

Par conséquent, l'algorithme mélangé MDEM proposé à la figure 5.9 fonctionne aussi bien que la TDEM, car il présente la même erreur de gain et la même erreur d'offset. En outre, on peut remplacer

TAB. 5.1 – Comparaison des erreurs produites par une TDEM et celles d'une SDEM similaire, pour 4 éléments

Termes	Composition de $Do_r(n)$	TDEM pure	SSB_r , SDEM partielle
1	l'erreur de gain, sont pareilles	$\overline{\alpha_r} y_r(n)$	$\frac{\alpha_1+\alpha_2+\alpha_3+\alpha_4}{4} y_r(n)$
2	l'erreur d'offset, sont pareilles	ϵ_r	$\epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4$
3-1	Les erreur modulées avec l'entrée (sont similaires au terme de l'ordre de la mise en forme et des coefficients, pour les TDEM et MDEM)	$\Delta_{11} S_{11}$	$\frac{\alpha_1-\alpha_2}{2} H(z)[Se_1(z) - Se_2(z)]$
3-2		$\Delta_{12} S_{12}$	$\frac{\alpha_1-\alpha_2}{2} H(z)[Se_3(z) - Se_4(z)]$
3-3		$\Delta_{21} S_{21}$	$\frac{(\alpha_1+\alpha_2-\alpha_3-\alpha_4)}{4} H(z) \times$ $[(Se_1 - Se_3) + (Se_2 - Se_4)]$

n'importe quel groupe de trois blocs SB dans les deux dernières couches de TDEM par un bloc SDEM local (SSB).

Cependant, il peut y avoir une différence entre leurs performances de mise en forme. Par exemple, on doit se poser la question "quelle est la différence entre les trois couples $\{3 - 1\}$, $\{3 - 2\}$ et $\{3 - 3\}$ représentées au tableau 5.1, pour les deux cas TDEM et SDEM partielle. Théoriquement, comme on l'a dit auparavant, les deux semblent identiques, car les coefficients Δ_{kr} sont tous les mêmes, et les termes $S_{kr}(z)$ et $Se(z).H(z)$ mettent en forme le Δ_{kr} par la même $H(z)$ si deux séquences $s_{kr}(n)$ et $se(n)$ sont bornées. Malgré ce fonctionnement compatible, $s_{kr}(n)$ n'arrive pas à suivre réellement la mise en forme de l'entrée $y_r(n)$ effectuée par le filtrage dans la boucle, car le fonctionnement effectué par le filtrage dans la boucle est fortement non linéaire à cause de son propre quantificateur imposé par la règle de conservation de l'entrée (voir la section 4.9.1). En revanche, l'évaluation de la séquence $se(n)$ n'est pas sensible au type d'entrée $y_r(n)$ (paire et impaire). Cette stratégie de tri permet de stabiliser la boucle incluant $H(z)$ et de suivre la mise en forme des erreurs associées Δ_{kr} sauf dans le cas limite, c.-à-d. $y_r(n) \in \{0,4\}$, ce qui est logique. Afin de minimiser les éventuels cas appliqués aux blocs SB_r s, nous avons proposé l'utilisation de l'équation 5.16 qui permet de distribuer l'entrée principale du système sans contraintes sur les différents SB_r .

Ainsi, l'algorithme MDEM est plus stable qu'une TDEM et il est presque indépendant des propriétés stochastiques de l'entrée du CNA. En pratique, parmi différents essais, la MDEM n'a jamais montré d'instabilité, de la même façon que la SDEM pur du même ordre qui est stable dans les mêmes conditions. La technique MDEM permet de réaliser une mise en forme des défauts d'appariement du CNA multibit à l'ordre 1, 2 et 3 pour les applications passe-bas ainsi qu'à l'ordre 2, 4 et 6 pour les applications passe-bande. La figure 5.10 montre le spectre de $s_{21}(n)$ dans le cas d'une TDEM conventionnelle et de son homologue $(sv_1(n) + sv_1(n) - sv_3(n) - sv_4(n))$ dans le cas de l'algorithme de MDEM du quatrième ordre passe-bande avec une entrée sinusoïdale en milieu d'échelle. L'amélioration de la performance de la TDEM par l'algorithme MDEM est évidente dans cet exemple. Pour presque tous les niveaux d'amplitude en entrée, on observe que peu de différences. En général, la diminution du bruit dans la bande est considérable, puisqu'elle peut atteindre jusqu'à 40 dB.

En raison de sa nature mixte, nous appelons cette nouvelle méthode MDEM. Elle fournit une solution pratique à un problème non résolu de stabilité lié depuis de longue date à l'algorithme TDEM. La

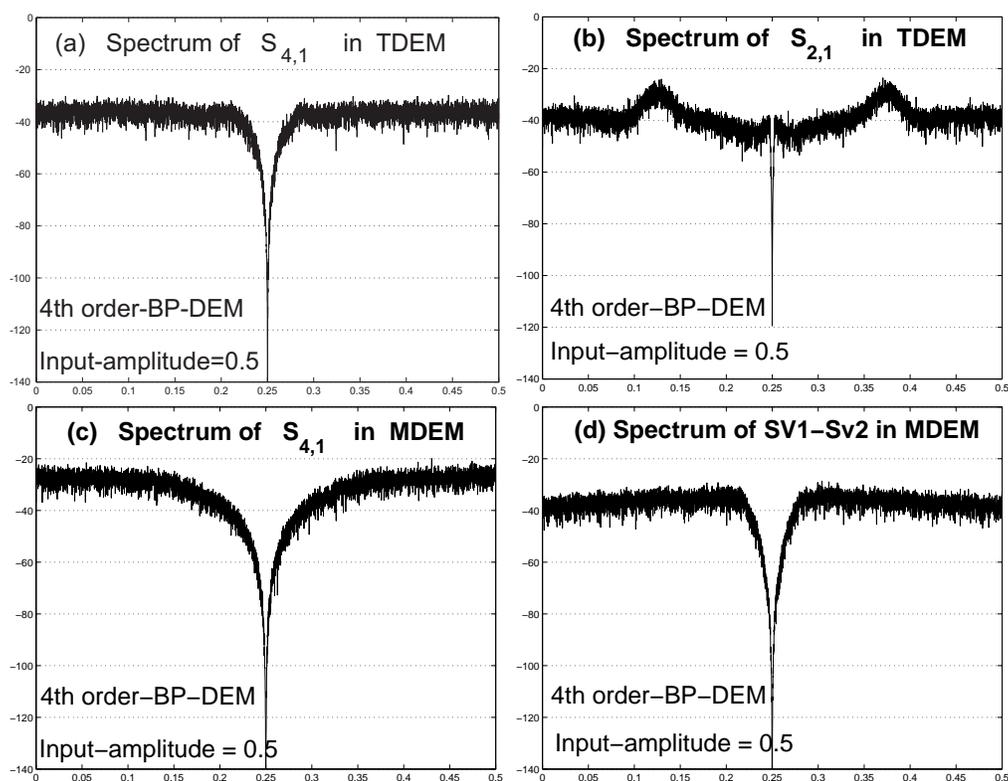


FIG. 5.10 – Comparaison de la fonctionnent des blocs de commutation : a,b) s_{41} et s_{21} dans une TDEM conventionnel, c) s_{41} dans l'MDEM-algorithme proposé, d) $sv_1(n) + sv_1(n) - sv_3(n) - sv_4(n)$ à SSB1 dans le MDEM proposé

réalisation de l'algorithme MDEM pour un CNA quatre bits ($M=16$) exige 19 filtres numériques, 24 comparateurs numériques et 4 décodeurs (6-à-4), alors que l'algorithme d'une SDEM pur semblable a besoin de 16 filtres numériques, 120 comparateurs numériques, 3 comparateurs de signe et d'un décodeur plus complexe (120-à-16). Dans la comparaison, une TDEM pure a besoin de 15 filtres numériques, de 15 comparateurs de signe et de 8 blocs supplémentaires de décision, mais elle est potentiellement instable donc inutilisable sauf pour le premier ordre.

5.4.2 Algorithme MDEM généralisé, deuxième option

Bien que le remplacement partiel de l'algorithme SDEM dans la deuxième couche d'une TDEM global puisse considérablement améliorer la stabilité du système, le remplacement peut être développé pour les autres couches, aussi bien que pour l'autre nombre des cellules dans chaque SSB. Le nombre des cellules appartenant à chaque bloc SSB n'a pas à être nécessairement un multiple de 2, mais ici le nombre est considéré comme puissance multiple de 2. Dans ce cas, l'expression générale d'erreur de non linéarité $e(n)$ se compose de deux parties, une allant avec l'algorithme de tri (SDEM) et l'autre allant avec la structure arborescente (TDEM). La partie de TDEM de l'équation peut être directement dérivée de l'expression 4.27. La partie de SDEM est développée de la même manière et avec la même méthode que celle présentée dans l'annexe de la référence [11].

Supposons que les couches B à $m+1$ soient décodées par la partie TDEM, et le reste de l'algorithme soit une SDEM-partielle avec 2^m cellules dans chacun des 2^{B-m} blocs SSB_r , donc, la sortie d'un CNA de $M = 2^B$ -bits peut s'exprimer comme suit :

$$Do_{MDEM}(n) = (1 + \bar{\alpha}) y(n) + e_{TDEM}(n) + e_{SDEM}(n) + \epsilon \quad (5.22)$$

où :

$$\begin{aligned} e_{TDEM}(n) &= \sum_{k=m+1}^B \sum_{r=1}^{2^{B-k}} \Delta_{kr} s_{kr}(n) \\ e_{SDEM}(n) &= \sum_{j=1}^{2^{B-m}} e_{SDEM_j}(n) \\ e_{SDEM_j}(n) &= \sum_{k=1}^m \sum_{r=1}^{2^{m-k}} \Delta_{j,k,r} sv_{j,k,r} \\ \Delta_{j,k,r} &= \frac{1}{2^k} \sum_{i=(j-1)2^m+(r-1)2^k+1}^{(j-1)2^m+(r-1)2^k+2^{k-1}} (\alpha_i - \alpha_{i+2^{k-1}}) \\ sv_{j,k,r} &= \sum_{i=(j-1)2^m+(r-1)2^k+1}^{(j-1)2^m+(r-1)2^k+2^{k-1}} (sv_i - sv_{i+2^{w-1}}) \\ \bar{\alpha} &= \frac{1}{M} \sum_{i=1}^M \alpha_i \quad , \quad \epsilon = \sum_{i=1}^M \epsilon_i \end{aligned} \quad (5.23)$$

et ϵ , α_i , Δ_{kr} et $s_{kr}(n)$ ont les mêmes définitions que dans l'équation 4.27. Puisqu'il n'y a aucune différence entre les vecteurs $sv_i(n)$ produits par les différents blocs SSB_r , toutes les combinaisons de $(sv_i - sv_j)$ pour $i, j \in \{1,2,3,4\}$ et $i \neq j$ ont les mêmes propriétés stochastiques et spectrales

à condition que l'algorithme fonctionne correctement (c.-à-d. que la distribution soit uniforme). En général, le vecteur $(sv_i(n) - sv_j(n))$ ou également $[H(z)(Se_i(z) - Se_j(z))]$ est semblable à un bruit blanc mis en forme par un filtre de même ordre. Alors, quand ils apparaissent dans le terme $e(n)$ à la sortie du CNA en multiplication avec $\Delta_{j,k,r}$, les défauts d'appariement des cellules seront écartés de la bande utile sans connaître les valeurs exactes des erreurs. Tous ces raisonnements sont valables à condition que les sorties des filtres soient bornées, autrement dit que les boucles soient stables.

Si ces équations sont développées avec $B = 4$ et $m = 2$, l'équation relative à la MDEM pratique proposée dans la section précédente sera obtenue.

Afin d'estimer le bruit dans la bande du terme $e_{MDEM}(n)$, d'autres simplifications peuvent être apportées à l'équation 5.22. Il convient de noter que les mêmes précautions que dans [12] doivent être prises :

- Les erreurs d'appariement des cellules du CNA, ayant l'écart type σ_α , sont blancs, non-corrélés les uns avec les autres et fixes durant l'opération de la correction. A cette condition l'erreur liée à chaque $\Delta_{j,k,r}$ à 5.22 ou pareillement $\Delta_{k,r}$ à 4.27 peut s'exprimer comme :

$$var(\Delta_k) = \frac{1}{2^k} \sigma_\alpha^2 \quad (5.24)$$

- Les erreurs du vecteur produit par différents filtres se_j sont également blanches et non corrélées avec une valeur efficace *rms* dénotée σ_{se_r} pour SSB_r .⁵
- La valeur moyenne du vecteur se_j dans chaque bloc SSB_r peut être ajustée de sorte que leurs valeurs aux moyennes carrées soient les mêmes sans affecter le rendement de SSB_j , c.-à-d. $\overline{se_j} \equiv 0$.
- Les filtres incorporés dans chacun des SSB_r sont tous identiques. Le terme σ_H dénote la valeur efficace dans la bande utile. La valeur efficace du filtre est en général définie comme suit :

$$\sigma_H = \sqrt{\frac{\pi^m}{(2m+1)OSR^{(2m+1)}}} \quad (5.25)$$

pour un filtre d'ordre $L = m$ passe-bas et d'ordre $L = 2m$ passe-bande ayant une bande passante de $\frac{f_e}{2OSR}$, où f_e est la fréquence d'échantillonnage [1].

- Tous les séquences $s_{kr}(n)$ dans la partie TDEM sont non corrélées les unes avec les autres et avec $se_r(n)$. Elles ont une moyenne nulle et σ_{s_k} évalue leurs valeurs efficaces dans la bande utile pour le nombre de couches k .

De ces hypothèses, une expression simplifiée pour la puissance du bruit dans la bande du CNA qui emploie un algorithme de SDEM partielle peut être déduite de l'équation 5.22, où var dénote l'opérateur de variance :

$$var(e_{SDEM}) = (2^m - 1) (\sigma_H \sigma_\alpha)^2 \sum_{j=1}^{2^{B-m}} \sigma_{Se_j}^2 \quad (5.26)$$

⁵. Il faut noter que cette condition n'est pas complètement vraie, mais peut néanmoins être employés pour simplifier les équations.

Nous pouvons supposer que tous les blocs de sous-tri SSB_r possèdent les mêmes caractéristiques stochastiques de sorte que les $\sigma_{se_r}^2$ sont identiques et non corrélés. Ainsi, nous obtenons :

$$var(e_{SDEM}) = (2^m - 1)2^{B-m}(\sigma_H \sigma_\alpha \sigma_{se})^2 \quad (5.27)$$

Par exemple quand $B = m = 4$, (c.-à-d. tous dans un bloc de SDEM conventionnelle), cette équation fournit la même expression que celle présentée dans [12] à quoi s'ajoute l'erreur du gain $\bar{\alpha} = \sum_{i=1}^{2^m} se_i$ de son côté.

La partie de l'erreur liée à l'algorithme TDEM ne peut pas être simplifiée facilement puisque les différents s_{kr} ne sont pas vraiment identiques. Cependant, si nous supposons que tous les s_{kr} dans la couche de niveau k sont similaires et non corrélés, alors, le bruit dans la bande de la partie TDEM est donné par :

$$var(e_{TDEM}) = 2^b \sigma_\alpha^2 \sum_{k=m+1}^b \frac{1}{2^{2k}} \sigma_{Sk}^2 \quad (5.28)$$

En raison du coefficient $\frac{1}{2^{2k}}$, on remarque que l'erreur qui est demeurée dans les blocs de commutation dans la couche inférieure affecte la sortie du CNA davantage que celle produite par les couches suivantes. Par exemple l'effet de la première couche ($k = 1$) est 64 fois plus grand que celui de la quatrième couche ($k = 4$). C'est tout à fait logique, car les couches les plus éloignées des cellules du CNA travaillent avec une erreur moyenne des erreurs des cellules respectives. Alors, comme on a dit auparavant, la performance des deux premières couches est dominante, cependant, ils ont plus de mal à fonctionner correctement avec la TDEM conventionnelle. De ce fait, en tenant compte du meilleur fonctionnement de l'algorithme SDEM, un autre avantage de l'algorithme mixte proposé devient évident.

Les valeurs de σ_{se} et de $\sigma_{s_{kr}}$ peuvent être estimées par des simulations. Leurs valeurs normales varient entre 0,1 et 0,8 pour différents m et pour différents ordres du filtre [12]. De plus, l'autre avantage de MDEM par rapport à TDEM conventionnelle est sa latence moindre sur le chemin du signal. La raison est que le signal d'entrée doit passer par une série de blocs de commutation $SB_{k,r}$. Leur nombre dans un MDEM de 16 cellules ($M=16$), est deux fois plus petit que dans le cas TDEM pur. Et puis il convient de noter que, comme dans n'importe quel SDEM, la partie principale du traitement des signaux dans chaque bloc SSB_r local, peut avoir lieu pendant la période d'horloge et pas nécessairement juste au moment où l'entrée arrive. En effet, les lignes de sortie qui doivent être activées peuvent être sélectionnées rapidement par les portes logiques juste après le changement d'entrée, alors que le tri des sorties des filtres pour la prochaine entrée s'est fait durant le reste de la période. Réduire la latence sur l'ensemble de l'algorithme est important si le modulateur est extrêmement sensible au retard sur le chemin de rétroaction, (par exemple, dans les applications haute fréquence [99]).

Comme le montre la figure 5.11, les résultats de la simulation confirment bien l'amélioration prévue de la stabilité et de l'abaissement du bruit dans la bande apportées par l'algorithme proposé. La méthode MDEM s'est bien comportée vis à vis du cas idéal et de la méthode SDEM, alors que le comportement de TDEM n'est pas acceptable du tout sauf pour le premier ordre. Le taux de défaut

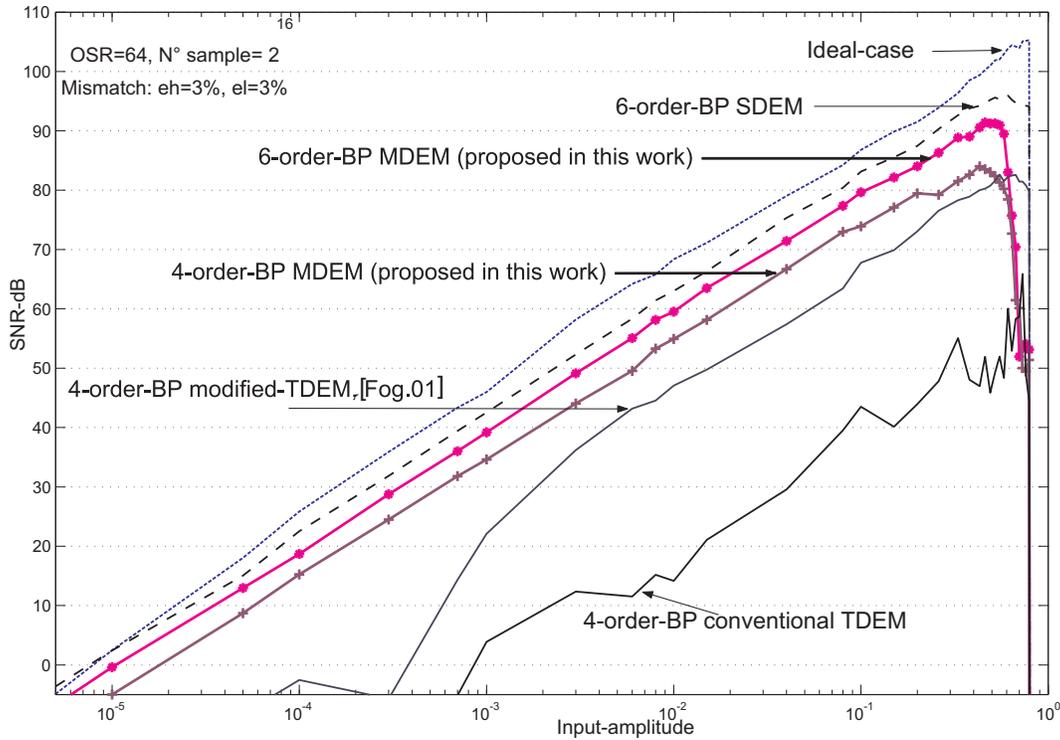


FIG. 5.11 – Comparaison les différentes techniques du brassage des sources, TDEM, MDEM, SDEM, ayant 3% défaut d'appariement, $OSR=64$, $M=16$

des cellules est supposé plus élevé que la normale afin de mettre en évidence la différence entre les différents méthodes.

À nouveau, compte tenu des technologies actuelles des circuits intégrés, l'utilisation d'un algorithme de DEM du second ordre passe-bas, ou également un algorithme du quatrième ordre passe-bande, suffit pour la plupart des applications. Nous avons cependant conçu et simulé un algorithme MDEM du sixième-ordre passe-bande, afin d'examiner son aptitude à s'intégrer dans un algorithme d'ordre supérieur. Sachant qu'on ne peut le concevoir par d'autres méthodes sauf par la SDEM.

Dans la pratique, afin de mettre en forme une erreur d'un niveau d'une valeur courante de 1%, l'utilisation d'une DEM passe-bande du quatrième ordre est suffisant. Les filtres $H(z)$ peuvent être différent dans leurs structures et caractéristiques, c.-à-d. être de type *FIR* simple ou à coefficients optimisés. Le dernier filtre doit être réalisé au moyenne de circuits plus complexes, alors que le premier peut être conçu en utilisant simplement des registres et des additionneurs sans multiplicateur. La structure d'une MDEM du quatrième ordre passe-bande simple est représentée sur la figure 5.12. Dans ce filtre simple, le filtre passe-bande est le suivant : $H(z) - 1 = 2z^{-2} + z^{-4}$. La simulation montre que la taille nécessaire des registres est au maximum de 7 bit au dernier étage du filtre dans la partie SDEM ainsi que dans la partie TDEM. Cette faible taille de registre peut être comparée à la taille habituelle des registres dans une TDEM conventionnelle qui peut être supérieur à 10 bits. La figure 5.13 montre le SNDR d'un modulateur passe-bande qui emploie une MDEM ayant la structure donnée à la figure 5.12. La simulation est effectuée dans le cas d'un CNA 4-bits ayant un défaut

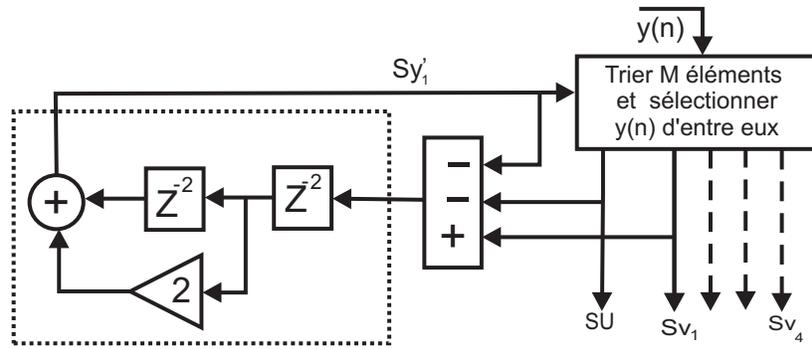


FIG. 5.12 – Structure d’une SSB simple conçue pour une MDEM passe-bande d’ordre quatre

d’appariement de 1% et un $OSR=64$. Le résultat d’une TDEM similaire bénéficiant de la dernière modification proposée dans [133] donne lieu à une comparaison dans les mêmes conditions.

Pour terminer, une remarque intéressante est que l’algorithme généralisé de la MDEM du premier

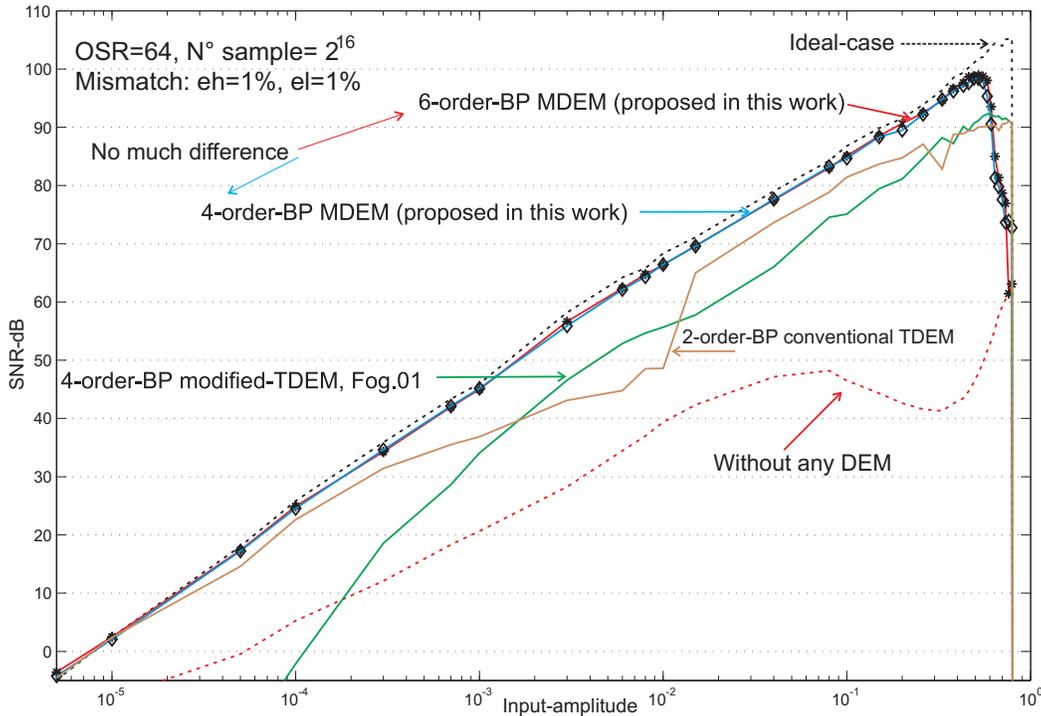


FIG. 5.13 – Comparaison de différentes techniques de DEM avec l’erreur de disparité de 1%.

ordre englobe aussi l’algorithme MP-DWA (présentée dans la section précédente), si $m = B - 1$, ce qui assure la continuité des méthodes proposées. La raison en est que le fonctionnement de l’algorithme SDEM du premier ordre est pratiquement équivalent à un algorithme de DWA, y compris en terme de résolution [130].

5.5 Brassage de source en structure arborescence tronquée, (*Shortened Tree structured, STDEM*)

Dans la section précédente, un algorithme associant les deux algorithmes de la technique TDEM et de la technique SDEM a été présenté. Le bon fonctionnement de cet algorithme (MDEM) a été validé par vérification analytique et par simulations. De plus, nous constatons que la réalisation de la MDEM a besoin d'un nombre de circuits assez réduit par rapport aux besoins de la SDEM, surtout au fur et à mesure que le nombre de cellules du CNA augmente. Cependant, la complexité des circuits par rapport à celle de la TDEM n'est pas négligeable. Même si la complexité est un peu supérieure à celle des circuits de TDEM similaire justifiée par le résultat obtenu avec MDEM, il est préférable chercher un algorithme plus économique en taille de la logique.

Dans cette section, nous présentons un autre algorithme qui utilise tout les recommandations suivies jusqu'ici. Autrement dit, ce nouvel algorithme doit remplir les conditions suivantes :

- un algorithme d'ordre supérieur est parfois nécessaire pour certaines applications de haute performance, mais, la complexité doit rester raisonnable par rapport aux autres techniques comparables.
- parmi les différentes techniques qui supportent une DEM d'ordre élevé, l'algorithme TDEM est le plus simple et aussi le plus économique, surtout pour $M > 8$, à la condition qu'elle soit stable.
- l'instabilité de l'algorithme de TDEM conventionnel peut être améliorée en utilisant une fonction de transfert modifiée, par exemple celle de l'équation 5.16, sauf dans les deux premières couches proches des cellules du CNA ($k = 1, 2$).
- le problème de l'instabilité qui réside dans les deux premières couches de la structure TDEM peut être résolu par un changement local dans sa structure par laquelle la sélection des cellules actives sollicitées par l'entrée arbitraire peut être effectuée autrement que dans la méthode TDEM normale, par exemple la sélection par l'algorithme de tri dans le SSB de la structure MDEM.
- un tel changement local doit obligatoirement respecter l'évolution du gain et de l'offset de la partie remplacée. Ceci permet d'avoir une mise en forme parfaite pour l'ensemble de l'algorithme modifié donné à l'expression 4.27.

Puisque l'instabilité de l'algorithme TDEM réside principalement dans la première couche, nous proposons une TDEM raccourcie dans laquelle on a retiré la première couche. Le concept principal de TDEM n'est cependant pas vraiment abandonnée dans les couches remplacées [149, 150]. En raison de sa structure tronquée "*Shortened*", nous appelons ce nouvel algorithme STDEM. Le schéma fonctionnel général d'un tel algorithme est montré dans la figure 5.14. L'algorithme de l'arborescence est bien respecté même s'il n'y a plus que les éléments du CNA à commander. Un bloc de commutation terminal appelé ESB_r , "*Ending Switching Block*" distribue son entrée ($0 \leq y_r(n) \leq 4$) aux quatre cellules. Chaque ESB_r se compose de trois filtres numériques indépendants et d'une logique de décision qui commande chaque groupe de 4 éléments du CNA. L'idée de cet algorithme découle de

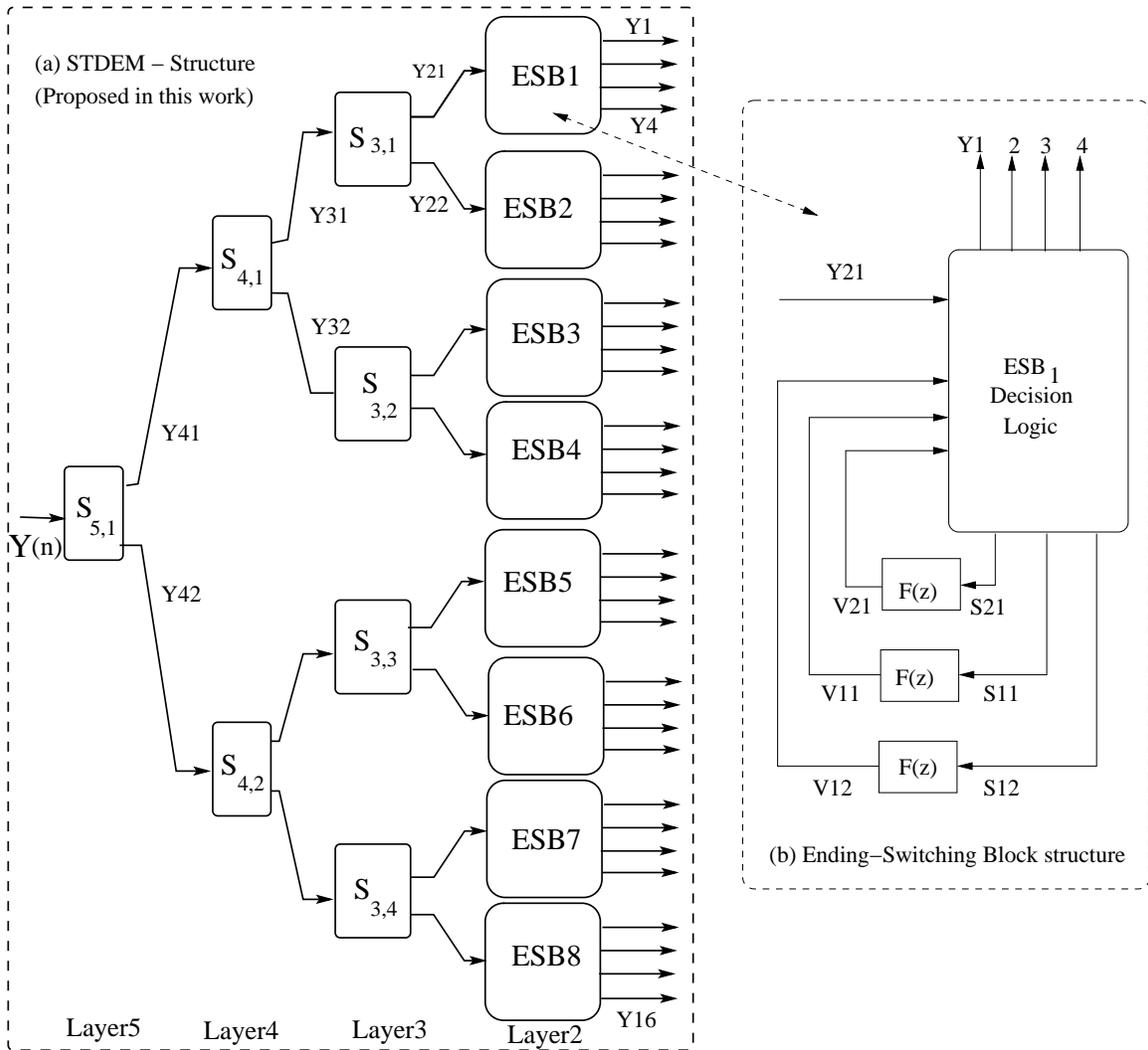


FIG. 5.14 – a) L’algorithme Proposé en structure arborescence tronquée (STDEM), b) le schéma fonctionnel d’une ESB

l’algorithme de la MDEM dans lequel la sensibilité au type d’entrée, paire et impaire est réduite ou éliminée. Autrement dit, il faut que chaque ESB puisse librement sélectionner les cellules commandées parmi les 4 éléments possibles afin de minimiser la moyenne d’énergie de l’ensemble des sélections dans la bande utile. Ceci est réalisé par les filtres inclus, leurs sorties étant maintenues à un niveau bas.

Afin de mieux suivre l’idée directrice de l’algorithme proposé, on prend un exemple avec les quatre cellules. nous comparons dans deux cas, la TDEM normale et la TDEM tronquée, sur le tableau 5.2. On voit bien que pour un groupe de quatre cellules, la différence principale entre un bloc de TDEM conventionnelle et un algorithme idéal est essentiellement liée à l’entrée $y_r(n)$ paire. Avec $y_r(n) = 2$, la TDEM conventionnelle offre quatre cas différents, alors qu’un algorithme idéal peut offrir six cas possibles. La modification proposée à la référence [133] n’y change rien, car le problème réside dans la deuxième couche où, en général, l’entrée ne peut pas être divisible par quatre. De plus, la modification

TAB. 5.2 – Distribution de l'entrée aux cellules du CNA, TDEM normale et TDEM- tronquée

$Y_r(n)$	Possibles cas, TDEM $\in \{sv_1, sv_2, sv_3, sv_4\}$	Possibles cas, STDEM $\in \{sv_1, sv_2, sv_3, sv_4\}$	remarque :
0	$\{\}$	$\{\}$	Similaire
1	l'une des cellules	l'une des cellules	similaire
2 ***	$\{sv_1, sv_3\}, \{sv_1, sv_4\}$ $\{sv_2, sv_3\}, \{sv_2, sv_4\}$	$\{sv_1, sv_2\}, \{sv_3, sv_4\}, \{sv_1, sv_3\}$ $\{sv_1, sv_4\}, \{sv_2, sv_3\}, \{sv_2, sv_4\}$	STDEM est plus libre que TDEM
3	3 cellules entre 4	3 cellules entre 4	similaire
4	Les 4 cellules	Les 4 cellules	similaire

proposée de la fonction de transfert 5.16 donne lieu aux deux possibilités $\{sv_1, sv_2\}$ ou $\{sv_3, sv_4\}$, qui sont différentes de celles proposées par la TDEM.

L'idée est donc que la meilleure méthode de sélection doit avoir les six cas possibles pour une entrée qui vaut deux. Cela est l'une des raisons pour laquelle la technique MDEM a été proposée pour résoudre la défaillance de l'algorithme TDEM. Nous essayons maintenant d'implanter cette idée mais dans une structure plus simple.

La priorité des ces différentes possibilités est de toute façon imposée par les filtres inclus qui font mettre en forme le défaut d'appariement des cellules. Avec la méthode MDEM le nombre des filtres est maximum, c'est-à-dire qu'il y a quatre filtres pour quatre cellules avec une stratégie de sélection par tri. En technique TDEM, il n'y a que trois filtres, mais, la stratégie n'est pas optimale. Par conséquent, on cherche une méthode de décision optimale pour la nouvelle technique STDEM sans ajouter de filtres. Afin d'optimiser une telle stratégie de sélection qui doit obligatoirement être compatible avec une structure globale de TDEM, il faut que les circonstances de décision soient identifiées.

Puisque l'algorithme proposé se compose de deux sortes de bloc de commutation, il est essentiel qu'ils fonctionnent correctement ensemble, comme nous avons déjà vu au cours de la présentation de l'algorithme MDEM. Le fonctionnement de la TDEM conventionnelle est exprimé par l'équation 4.27. Si nous le développons pour le cas $B = 2$ on obtient l'expression 5.17. Chaque ESB_r représenté sur la figure 5.14-a, doit avoir la même fonctionnement que l'équation 5.17. On doit noter les deux points importants suivants :

- l'erreur de gain doit être égale à $\overline{\alpha_r}$ et l'offset reste toujours la même $\epsilon = \sum_{i=1}^M \epsilon_i$
- toutes les erreurs modulées par une partie de l'entrée (par exemple, s_{11}, s_{12}, s_{21}) doivent être mis en forme par un filtre de même ordre L .

Une solution simple est présentée dans la figure 5.14-b, où l'état des trois filtres définissent quelle combinaison des cellules du CNA doit être sélectionnée vis à vis de l'entrée $y_r(n)$. Les logiques de décision sont récapitulées dans le tableau 5.3, où les niveaux des seuil " t_1 " et " t_2 " dépendent de l'ordre et de la structure des filtres; ils peuvent être optimisés par simulation ou analytiquement par l'estimation.

En fait, chaque branche comprenant trois blocs de commutation dans les première et deuxième couches de la TDEM conventionnelles sont unifiées pour recomposer un nouveau bloc ESB. Par exemple, les blocs $S_{2,1}$, $S_{1,1}$ et $S_{1,2}$ (dans la figure 4.32) sont remplacés par le bloc ESB_1 . La différence

TAB. 5.3 – Logique de décision employé dans l'ESB de l'algorithme STDEM

y_r	Conditions sur : R_{21}, R_{11}, R_{12}	$s_{21}s_{11}, s_{12}$	sv_{1-4}
0	Don't-care	0, 0, 0	0000
1	$R_{21} \geq -t_2, R_{11} \geq t_1$	1, 1, 0	1000
"	$R_{21} \geq 0, R_{11} \geq 0, -t_1 \leq R_{12} < t_1$	1, 1, 0	1000
"	$R_{21} < t_2, R_{11} < -t_1$	1,-1, 0	0100
"	$R_{21} \geq 0, R_{11} < 0, -t_1 \leq R_{12} < t_1$	1,-1, 0	0100
"	$-t_1 \leq R_{11} < t_1, R_{12} \geq t_1$	-1, 0, 1	0010
"	$R_{21} < 0, -t_1 \leq R_{11} < t_1, R_{12} \geq 0$	-1, 0, 1	0010
"	$-t_1 \leq R_{11} < t_1, R_{12} < -t_1$	-1, 0,-1	0001
"	$R_{21} < 0, -t_1 \leq R_{11} < t_1, R_{12} < 0$	-1, 0,-1	0001
2	$R_{21} \geq t_2$	2, 0, 0	1100
"	$R_{21} < -t_2$	-2, 0, 0	0011
"	$-t_2 \leq R_{21} < t_2, R_{11} \geq 0, R_{12} \geq 0$	0, 1, 1	1010
"	$-t_2 \leq R_{21} < t_2, R_{11} \geq 0, R_{12} < 0$	0, 1,-1	1001
"	$-t_2 \leq R_{21} < t_2, R_{11} < 0, R_{12} \geq 0$	0,-1, 1	0110
"	$-t_2 \leq R_{21} < t_2, R_{11} < 0, R_{12} < 0$	0,-1,-1	0101
3	$-t_1 \leq R_{11} < t_1, R_{12} \geq t_1$	1, 0, 1	1110
"	$R_{21} \geq 0, -t_1 \leq R_{11} < t_1, R_{12} \geq 0$	1, 0, 1	1110
"	$-t_1 \leq R_{11} < t_1, R_{12} < -t_1$	1, 0,-1	1101
"	$R_{21} \geq 0, -t_1 \leq R_{11} < t_1, R_{12} < 0$	1, 0,-1	1101
"	$R_{11} \geq t_1$	-1, 1, 0	1011
"	$R_{21} < 0, R_{11} \geq 0, -t_1 \leq R_{12} < t_1$	-1, 1, 0	1011
"	$R_{11} < -t_1$	-1,-1, 0	0111
"	$R_{21} < 0, R_{11} < 0, -t_1 \leq R_{12} < t_1$	-1,-1, 0	0111
4	Don't-care	0, 0, 0	1111

importante est l'utilisation d'une décision dynamique pour stabiliser l'algorithme. Cela est fait en considérant non seulement la valeur d'entrée mais également toutes les sorties des filtres de ESB à chaque période. Afin de satisfaire les conditions générales de TDEM (c.-à-d. l'équation 5.17), le quantificateur spécial dans chaque ESB doit agir en satisfaisant les conditions suivantes :

$$\begin{aligned} s_{11} &= sv_1 - sv_2 & \text{and} & & s_{21} &= sv_1 + sv_2 - (sv_3 + sv_4) \\ s_{12} &= sv_3 - sv_4 & \text{and} & & y_{21} &= sv_1 + sv_2 + sv_3 + sv_4 \end{aligned} \quad (5.29)$$

De plus, pour neutraliser la dépendance au type d'entrée dans la reste de l'algorithme, la fonction de transfert du quantificateur spécial du bloc $SB_{k,r}$ est remplacée par celle de 5.16 dans les couches suivantes ($k > 2$). Par conséquent, l'algorithme proposé dans la figure 5.14 a le même comportement fonctionnel que TDEM conventionnelle mais son algorithme n'est pas sensible à l'entrée paire ou impaire de manière critique.

A notre connaissance, il n'y a aucune méthode algébrique permettant d'examiner la stabilité de l'algorithme DEM d'ordre supérieur. On ne traite ce problème que par les simulations [2, 11, 12, 79, 83, 129, 133, 136, 137], comme nous le ferons ici.

La STDEM fournit enfin une solution très simple et sans coût supplémentaire à un problème non résolu de stabilité lié à l'algorithme TDEM. La réalisation d'algorithme de STDEM n'exige aucun circuit supplémentaire. Selon la nouvelle fonction de transfert 5.16 et le tableau 5.3, quelques changements nécessaires des quantificateurs spéciaux dans différentes couches peuvent être faits dans l'étape de conception de la logique. L'expression globale pour les erreurs du CNA est la même qu'en TDEM conventionnelle 4.27 avec seulement des considérations nécessaires sur les indices. Cependant, la différence réelle est associée à la qualité de la mise en forme par les différentes couches. Un autre avantage de STDEM est son temps de réponse plus rapide, car l'ESB n'a plus besoin d'additionneur ou de multiplicateur.

Les simulations confirment l'efficacité des modifications proposées ci-dessus, si bien que maintenant la STDEM peut même utiliser des filtres d'ordre trois passe-bas ou d'ordre six passe-bande. Afin de vérifier la différence entre ces algorithmes, il vaut mieux réaliser les simulations avec un taux d'erreur légèrement plus élevé de l'ordre de $\sigma_\alpha = 3\%$. La figure 5.15 montre le spectre de la sortie du modulateur passe-bande $v(n)$, produit avec un CNA de 4-bit : idéal (a), sans aucune DEM (b), d'utilisant une TDEM conventionnelle d'ordre quatre (c), et l'utilisant d'une STDEM d'ordre quatre (d). Dans la comparaison, le fonctionnement prévu de STDEM est justifié dans la mesure où elle supprime des erreurs de disparité d'environ 35dB.

Afin d'effectuer de plus nombreuses comparaisons, dans le deuxième groupe de simulations nous avons simulé les différents algorithmes de DEM avec un même modulateur en entrée variant entre 10^{-6} et sa valeur maximale (voire la figure 5.16). Le système utilise :

- la TDEM d'ordre deux passe-bande conventionnelle,
- la TDEM d'ordre quatre passe-bande conventionnelle,
- la TDEM modifiée d'ordre quatre passe-bande présentée par [133],
- la STDEM d'ordre quatre passe-bande proposée dans ce travail

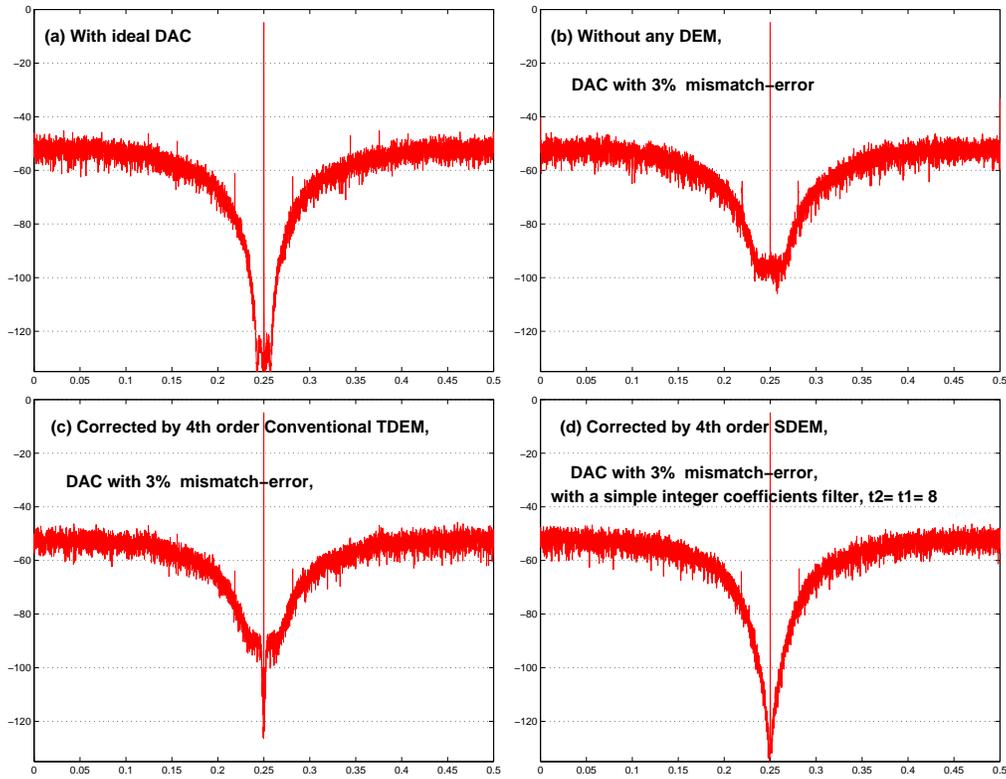


FIG. 5.15 – Spectre de la sortie du modulateur sigma delta passe-bande d'ordre 8 : a) avec CNA idéal, b) sans aucune correction mais $\sigma_\alpha = 3\%$, c) en utilisant une TDEM conventionnelle d'ordre 4 passe-bande, d) en utilisant une STDEM d'ordre 4 proposée dans cette section.

– la STDEM d'ordre six passe-bande proposée dans ce travail

Comme représenté sur la figure 5.16, pour une erreur de disparité de 3% dans le modulateur de base, le système ne peut pas atteindre un SNR_{max} de plus que 58dB, alors que le cas idéal est à l'ordre de 105dB. Bien que la TDEM de second ordre conventionnel ait une correction d'erreur considérable, l'utilisation d'une TDEM du quatrième ordre passe-bande conventionnelle n'augmente pas la précision du système et sa performance peut même être inférieure à celle d'un système sans correction. Cependant, moyennant les quelques modifications qui ont été présentées dans [133], le quatrième ordre passe-bande TDEM possède une bonne linéarité mais il est loin du cas idéal. Son spectre non régulier signifie que cet algorithme dépend toujours des propriétés stochastiques de l'entrée. Il convient de noter que ces modifications (celles dans [133]) ne peuvent pas s'appliquer à un algorithme du sixième ordre.

Les résultats de la simulation suivent correctement les aspects théoriques de l'algorithme proposé STDEM dans laquelle les seuils t_1 et t_2 sont optimisés par la simulation. Les résultats semblent réguliers pour les deux systèmes d'ordre quatre et d'ordre six, comme décrit à la figure 5.16.

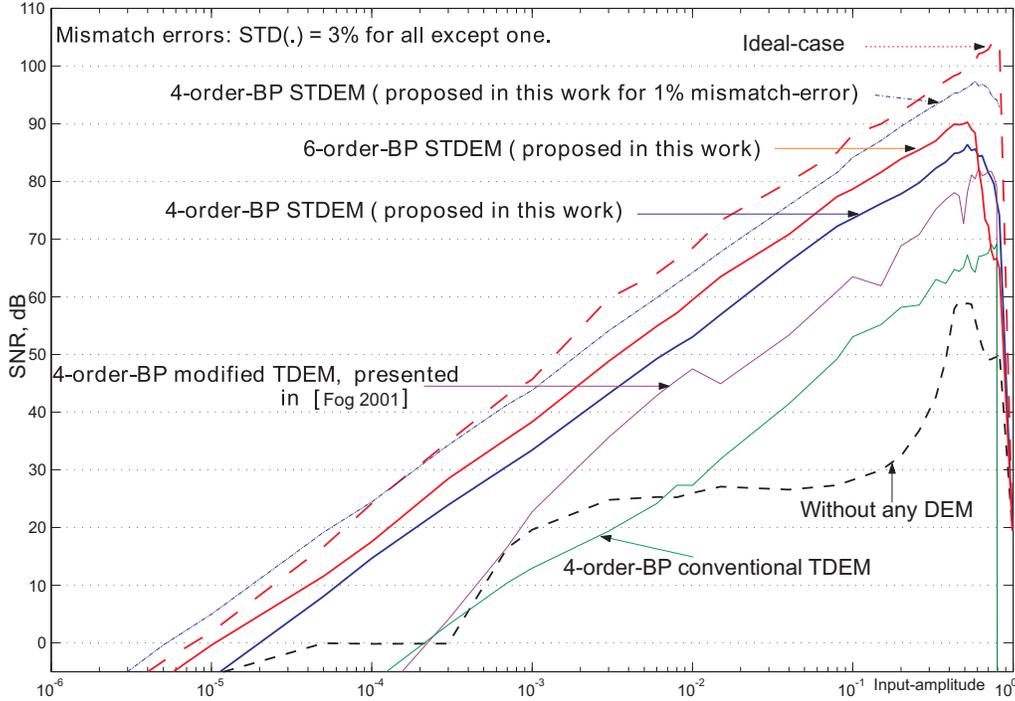


FIG. 5.16 — Comparaison de la performance des différentes techniques de DEM dans un modulateur passe-bande

5.5.1 Description des circuits de STDEM et estimation des seuils

L'écart type sur les circuits est de l'ordre de 1%, (habituellement $0.1\% \leq \sigma_\alpha \leq 1\%$). Ce niveau de disparité provoque une distorsion harmonique indésirable dans la bande utile. Dans un modulateur multibit, à partir d'une résolution de 10 bit, si on veut utiliser un modulateur multibit il faut appliquer au moins une méthode de correction. Tant que l'OSR reste normal, par exemple $OSR > 16$, l'une des techniques d'appariement dynamique (DEM) est celle qui convient le mieux. Cependant, une méthode DEM de premier ordre n'est pas suffisante pour une résolution plus de 15 bits quand l'OSR est inférieur à 100. Il faut envisager une DEM de second ordre qui peut autoriser une résolution allant jusqu'à 20 bits. Dans cette section, nous détaillerons l'algorithme proposé ci-dessus (STDEM) ayant une structure d'ordre quatre passe-bande qui convienne à une résolution entre 16 et 20 bit en technologies CMOS actuelles. Afin de parvenir à une résolution de 24 bits, il faut augmenter l'OSR ainsi que l'ordre de DEM.

La figure 5.17 montre le bloc de commutation $SB_{k,r}$, $k > 2$, conçu pour être utilisé dans une STDEM simple d'ordre quatre passe-bande. On le réalise complètement dans le domaine numérique; tous les coefficients sont des nombres entiers, ce qui simplifie les circuits. Puisque le quantificateur spécial fonctionnant dans la boucle est modifié, la sortie $s_{kr}(n)$ peut accepter une valeur différente de zéro pour des entrées (y_{kr}) toutes différentes de zéro. Alors, la résolution maximale du registre à l'ordre de 5-bits est tout à fait raisonnable (c.-à-d. $Max[V_{kr}] \leq 15$, ou 4 bits).

La conception actuelle de ESB_r passe-bande d'ordre quatre est représentée sur la figure 5.18, où un groupe de règles de décision suffisamment flexible est basé sur les équations 5.29 et la table 5.3. Afin

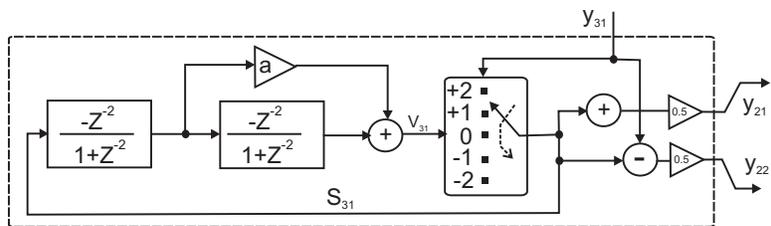


FIG. 5.17 – $SB_{k,r}$ modifié pour STDEM, utilisé dans toutes les couches avant le dernier, $k \geq 3$.

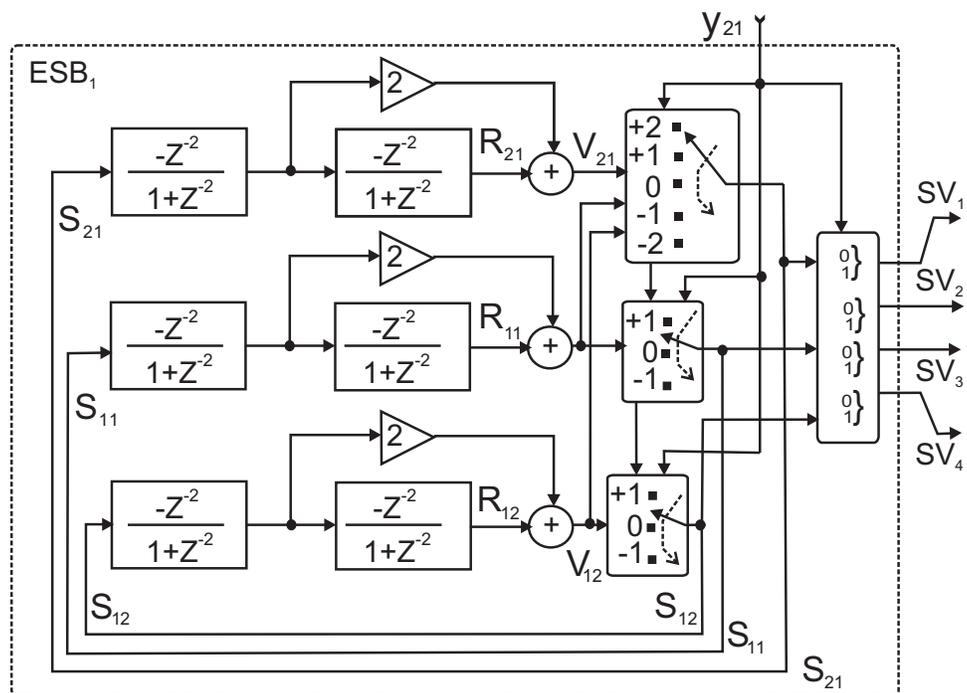


FIG. 5.18 – Schéma d'un bloc de ESB_r , utilisé dans la dernière couche de STDEM

d'estimer des niveaux de seuil t_1 et t_2 , on doit s'intéresser aux niveaux critiques des valeurs numériques des sorties des filtres. Pour la boucle, qui par exemple contient S_{21} , R_{21} et V_{21} , une situation critique commence si la boucle ne peut pas être commandée par un s_{21} non zéro pendant au moins deux périodes consécutives alors que les intégrateurs de la boucle (ou les résonateurs) ont déjà pris des valeurs différentes de zéro. Par conséquent, le deuxième intégrateur peut partir en débordement en quelques périodes, car le nombre de bits de ses registres est limité dans la pratique. Par exemple, quand les sorties initiales des intégrateurs sont supposées égales à 4, alors, le premier intégrateur reste inchangé, tandis que la sortie du deuxième intégrateur (ici : R_{21}) devient égale à 8, puis $V_{21} = 16$, tandis que la valeur de rebouclage (s_{21}) est encore nulle pendant deux périodes consécutives. Par conséquent, la valeur du seuil t_1 doit détecter la tendance à l'augmentation d'une telle sortie. Ceci peut simplement être réalisé en considérant la valeur actuelle de t_{21} comme l'indique le tableau 5.3. Le niveau de seuil estimé avec cet exemple simple est dans l'ordre de 16 et de -16, respectivement pour les sorties positives et négatives du résonateur.

Cependant, nous sommes intéressés par sa valeur minimale afin de suivre correctement le fonctionnement de la boucle. Ainsi, avec la même façon de raisonner pendant une période, le niveau le plus bas de seuil peut être estimé à environ 8. Dans différentes simulations, ces valeurs estimées ont pour conséquence la mise en forme parfaite du défaut d'appariement des cellules du CNA. La valeur optimale obtenue par la simulation est comprise entre 8 et 12, valeurs théoriquement prévues dans les discussions menées ci-dessus. Nous pouvons également estimer la valeur du deuxième niveau de seuil $t_{2,opt}$ de la même manière. La valeur t_2 estimée est 8. Sa valeur optimisée vaut 6. Il n'y a cependant pas beaucoup de différence si on la considère pratiquement identique à t_1 , soit $t_{2,opt} = 8$.

Cette méthode peut également être employée pour estimer la résolution maximale des registres. Si nous supposons qu'une situation autant critique peut être encore présente pendant environ quatre périodes, le deuxième registre ($R_{21,out}$) peut aller jusqu'à la valeur de 28, qui nécessite une résolution de 5-bits. Ceci est bien confirmé par les différentes simulations, et de fait il excède rarement plus de 20 (voire la figure 5.19). Afin de donner une meilleure assise à ce jugement, la figure 5.20 montre les niveaux maximaux des registres pour différents niveaux d'entrée; la même sortie d'un algorithme de TDEM conventionnelle est montrée à des fins de comparaisons.

De plus, même en acceptant l'utilisation d'une résolution élevée pour le registre, l'opération de mise en forme ne se produit pas correctement dans la boucle de TDEM conventionnelle. Ce phénomène dépend de l'OSR du système de sorte que l'erreur du CNA ne peut être constituée que si la boucle relative est commandée plus rapidement que la transmission de signal. Dans le domaine de fréquence numérique, il est représenté par une bande très étroite autour de la fréquence centrale comme le décrit la figure 5.15-c pour une TDEM conventionnelle, améliorée à la figure 5.15-d par la méthode STDEM.

De la figure 5.20, on voit bien que le niveau maximal du registre dans un ESB ainsi que dans un SB-modifié est en général inférieur à celui d'un SB dans sa définition conventionnelle. C'est parce que dans le SB-modifié, le quantificateur spécial basé sur l'équation 5.16 n'est pas limité par les types d'entrée pairs ou impairs. De plus, les SB ont plus de chances de recevoir une entrée différente de

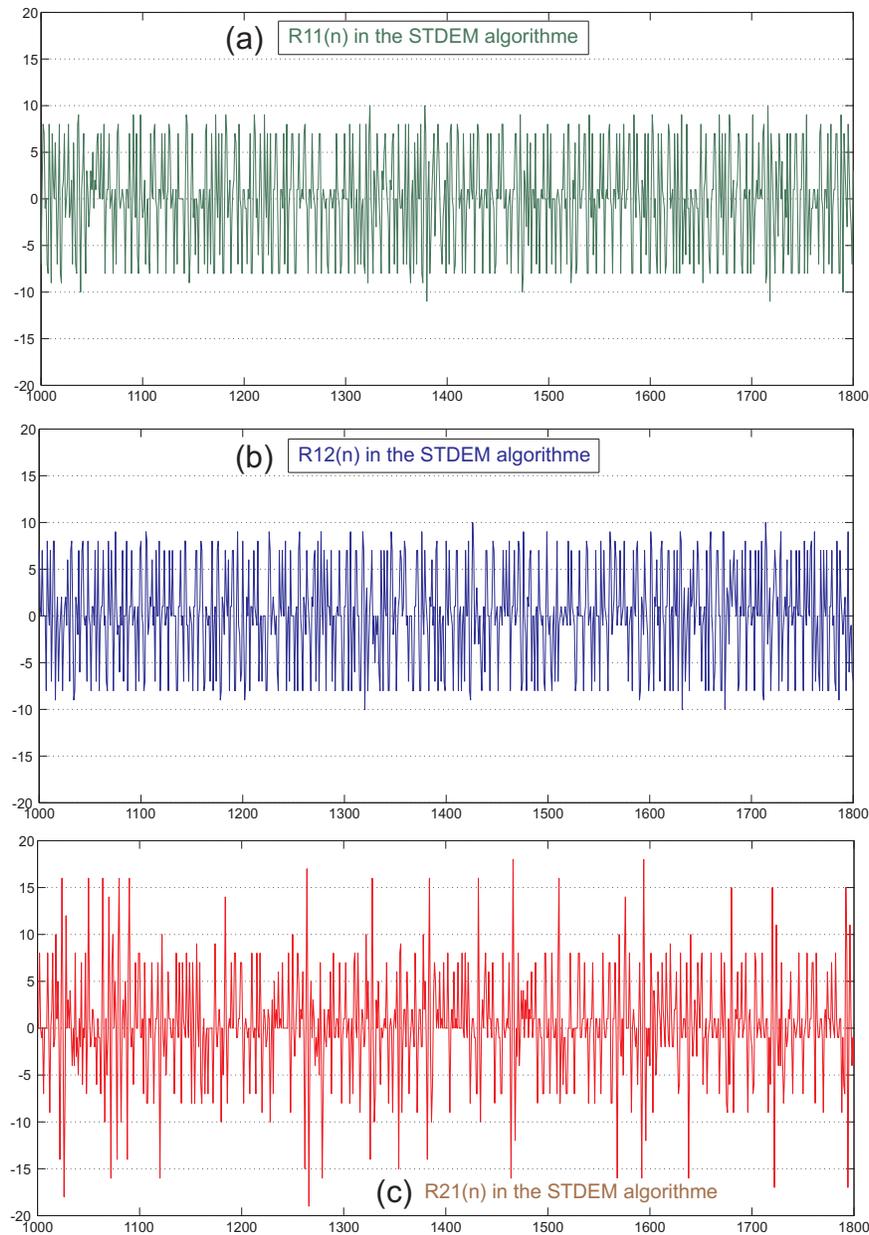


FIG. 5.19 – Sortie du deuxième résonateur utilisé dans la STDEM qui dépasse rarement au dessus du 16, a) $R11$, b) $R12$ c) $R21$; expliquées au tableau 5.3.

zéro puisqu'ils traitent d'abord toutes les entrées. Les circuits conçus avec STDEM sont simulés avec une précision fixe du registre. Comme discuté précédemment, les résolutions de tous les registres sont limitées à 5-bits et sont signées. Ensuite, la STDEM mentionnée a été mise en application dans l'environnement *Cadence* et *Verilog* et puis optimisée au moyens des outils *Ambit*. La figure 5.21 montre le spectre de la sortie du modulateur, dans le cas idéal, sans aucune correction et lorsque mise en forme par une STDEM passe-bande d'ordre quatre ayant des seuils $t_1 = t_2 = 8$. Le bruit de disparité est diminué environ $35dB$ dans la bande d'utile pour $OSR = 100$. Ce circuit peut fonctionner avec une fréquence d'échantillonnage maximale de $330MHz$ conçu en technologie $0.35\mu m$ -CMOS, comme le montre la simulation après le routage final ("post layout"). Afin d'utiliser la STDEM présentée à

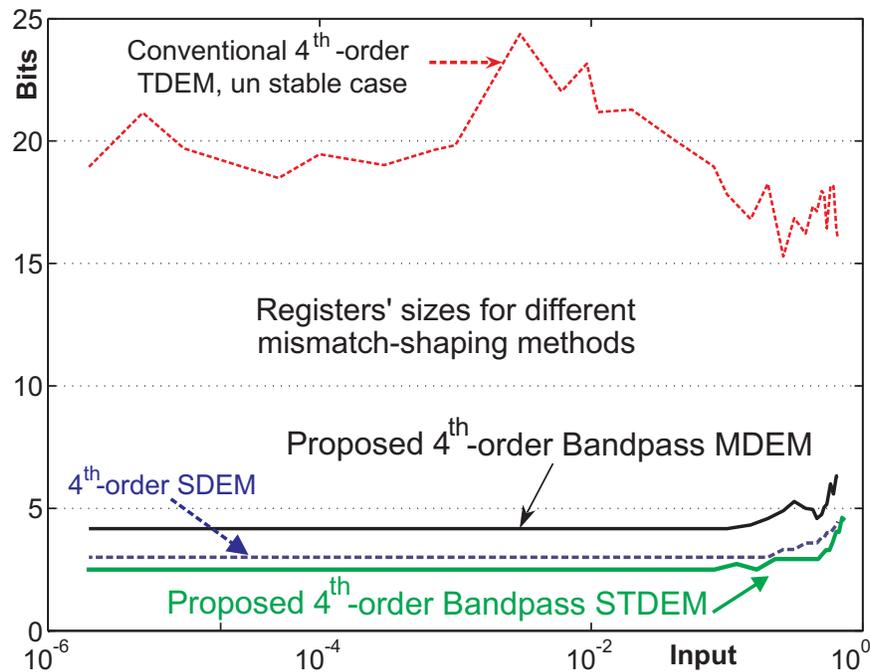


FIG. 5.20 – Niveaux maximum de la sortie des registres, pour un ESB usées dans la STDEM et pour un SB utilisé dans la TDEM conventionnel

cette vitesse si élevée, une période de retard doit être envisagée et compensée au niveau système [99]. Pour 8 cellules, cette STDEM a besoin d'une surface de $0.22\mu m^2$ et contient environ 3,000 portes comme cela est rapporté par Ambit sur la figure 5.22. Dans la comparaison, une SDEM semblable a besoin d'une surface de $0.6\mu m^2$ et ne peut fonctionner que deux fois plus lentement qu'une STDEM semblable.

Comparaison des différents techniques de brassage des sources

La performance des quatre algorithmes du brassage des sources proposées dans ce chapitre peuvent être comparées à l'intérieur des quatre catégories définies comme suit :

- du premier ordre passe-bas, (d'ordre zéro incluse),
par exemple : RDEM, CLA-DEM, DWA, SDEM, TDEM, MDEM et STDEM.
- du second ordre passe-bande,
par exemple : BP-DWA, TDEM, SDEM, MDEM et STDEM.
- du second ordre passe-bas,
par exemple : SDEM, TDEM, MDEM et STDEM.
- du quatrième ordre passe-bande,
par exemple : TDEM, SDEM, MDEM et STDEM.

Chaque comparaison s'effectue selon différents critères et différents paramètres qui varient de la manière suivante :

1. le SNDR vis-à-vis de l'amplitude d'entrée,
2. le SFDR vis-à-vis de l'amplitude d'entrée unique ou multiple,

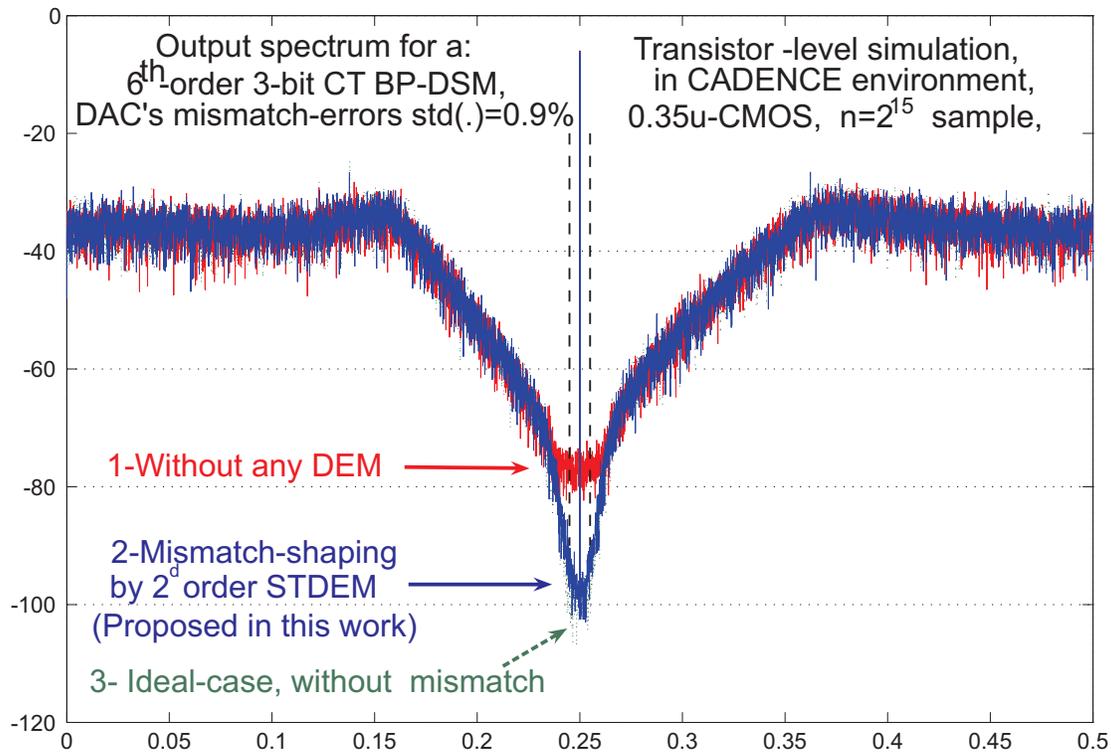


FIG. 5.21 – Simulation effectuée au niveau Transistor qui est implantée à l'environnement Cadence, pour un modulateur multibit qui emploie une algorithm STDEM d'ordre quatre passe-bande

Summary Area Report, Module: 'STDEM_ambit'

Block report for module	Current	Cumulative
Number of combinational instances	370	1079
Number of noncombinational instances	278	278
Number of hierarchical instances	8	8
Total number of instances	656	1365
Area of combinational cells	1396.00	4524.00
Area of non-combinational cells	4544.00	4544.00
Total cell area	5940.00	9068.00
Number of nets	715	1524
Total area	5940.00	9068.00

FIG. 5.22 – Rapporte de l'Ambit, pour une STDEM ayant 8-cellules

3. le SNDR vis-à-vis d'OSR,
4. le SFDR vis-à-vis d'OSR,
5. le SNDR vis-à-vis du défaut d'appariement des cellules du CNA,
6. le SFDR vis-à-vis du défaut d'appariement des cellules du CNA,
7. la résolution finale vis-à-vis du coût et de surface des circuits nécessaires,
8. la résolution finale vis-à-vis de la consommation des circuits,
9. la fréquence maximale de fonctionnement par rapport au nombre des cellules.

Certaines comparaisons ne sont pas envisagées dans ce travail, par exemple celles de la consommation, d'autres ont déjà été présentées dans les différentes références (nos propres communications incluses) [82,149–153]. Dans la suite, nous donnons aussi une autre comparaison entre le SNDR de l'ensemble des méthodes de DEM passe-bande d'ordre quatre qui sont possible à la figure 5.23, où les méthodes considérées sont : la TDEM-conventionnelle, la TDEM-modifiée dans [133], la SDEM, la MDEM et la STDEM. La bande utile est supposée ayant un OSR valant 64 et le défaut d'appariement est de 1%. On note que le meilleur résultat est obtenu par application de la méthode SDEM, suivi par MDEM, puis STDEM. Pour l'aspect économique en circuits, la technique de STDEM est préférable. Il sera ainsi possible de réaliser un modulateur $\Sigma\Delta$ passe-bande ou passe-bas plus rapide qu'avec une

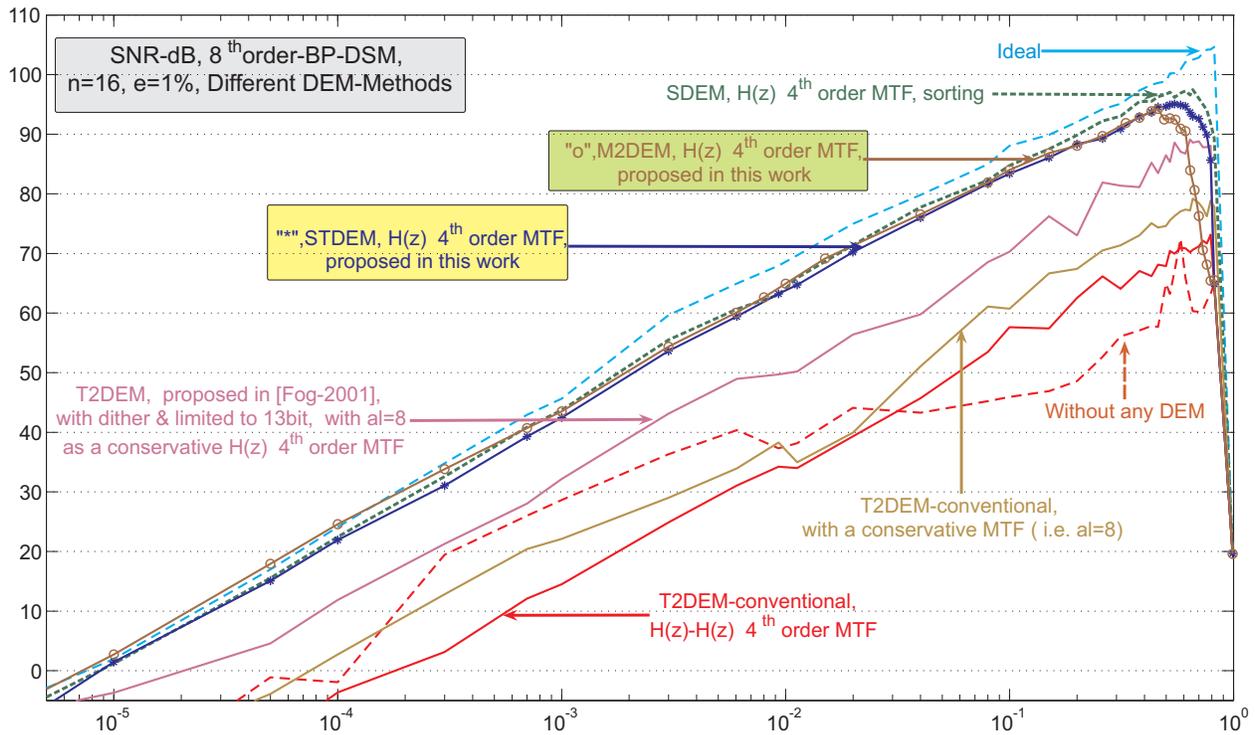


FIG. 5.23 – Comparaison le SNDR des différentes techniques de DEM du quatrième ordre passe-bande

quelconque autre technique.

5.6 Conclusion

Après avoir étudié les différents algorithmes existants de DEM au chapitre précédent, nous avons présenté dans ce chapitre quatre nouveaux algorithmes. Les deux premiers sont en fait obtenus au moyen de certaines modifications permettant de rendre la technique DWA plus performante dans la pratique, l'une visant l'application passe-bas et l'autre l'application passe-bande. La vérification analytique ainsi que la comparaison par des simulations numériques montrent l'efficacité de ces deux propositions, dénotées MP-DWA. Elle rend l'élaboration des modulateurs multibits possible dans les technologies actuelles destinées aux résolutions atteignant 10 à 16 bits qu'ils soient passe-bas ou

passer-bande. L'avantage majeur de cet algorithme est d'éloigner les tons indésirables de la bande utile d'une manière plus efficace qu'auparavant avec très peu de circuits supplémentaires, surtout pour l'application passer-bande.

Puis, deux autres algorithmes ont été présentés et généralisés aux applications passer-bas et passer-bande sans que leur soient imposées des conditions particulières. Ces deux algorithmes utilisent de vrais filtres numériques qui peuvent avoir des caractéristiques arbitraires d'ordre et de bande passante. La première méthode nommée MDEM est un bon compromis entre les deux algorithmes TDEM conventionnelles et SDEM en terme de circuits nécessaires avec un degré de performance proche du SDEM. La dernière méthode proposée est une TDEM tronquée dénotée STDEM. Elle fournit une très belle performance sans aucun circuits supplémentaires par rapport à la TDEM conventionnelle. Elle peut atteindre un plus grand ordre et se prêter ainsi à différents types de filtre sans être gêné par le phénomène d'instabilité. Les méthodes MDEM et STDEM ont été implantées aux ordres 2 et 3 dans le cas passer-bas et aux ordres 4 et 6 dans le cas passer-bande, ce qui permet d'avoir des modulateurs multibit de très haute résolution, jusqu'à 16 bits voire même 20 bits. De plus, une résolution de 24 bits peut aussi s'obtenir si on suppose un modulateur ayant un CNA interne de 6 bits avec un défaut d'appariement modéré, par exemple 0.1%. Un avantage de la méthode STDEM est qu'elle peut fonctionner avec une fréquence assez élevée. Cela a été vérifiée par les simulation au niveau des transistors jusqu'au routage final dans l'environnement *Cadence*.

Chapitre 6

Implémentation de l'étage CAN/CNA dans un modulateur $\Sigma\Delta$ multibit

6.1 Introduction

Nous avons examiné au chapitre 3 les principales contraintes pesant sur chacun des blocs du modulateur $\Sigma\Delta$. Nous avons vu que les erreurs du modulateur peuvent être classées en deux catégories : les erreurs statiques et les erreurs dynamiques. Des erreurs statiques, c'est le défaut d'appariement des cellules du CNA multibit qui a l'influence la plus importante, aussi bien dans le cas discret que dans le cas continu. Afin d'obtenir un modulateur multibit ayant des caractéristiques statiques s'approchant le plus possible du modulateur idéal, certaines techniques ont ensuite été présentées aux chapitres 4 et 5. Dans le cadre d'une correction passe-bande d'ordre 4 des défauts d'appariement du CNA interne, l'implantation des deux méthodes proposées au chapitre précédent (MDEM et STDEM) sera suivie dans ce chapitre jusqu'à l'étape finale (routage). Une comparaison sera également faite avec la méthode DEM conventionnelle.

Par ailleurs, les erreurs dynamiques les plus importantes pour un modulateur multibit sont, d'après la conclusion du chapitre 3 :

– **soit liées au CNA:**

1. gigue d'horloge,
2. *glitch*,
3. injection de charge *CFT*,
4. temps de transition non équilibrés,

– **soit liées au CAN:**

1. hystérésis du quantificateur,
2. erreur de métastabilité du quantificateur.

Nous présenterons au cours de ce chapitre quelques techniques permettant de compenser les erreurs dynamiques du CNA multibit ainsi que leur implémentation au niveau transistor, en technologie *CMOS*. Ensuite, une architecture entrelacée pouvant limiter les erreurs dynamiques du quantifica-

teur pour les applications haute fréquence sera présentée¹.

Nous présenterons les circuits conçus et les résultats de simulations obtenus au niveau système et au niveau transistor. Les simulations au niveau système (présentées dans ce chapitre) ont été réalisées à l'aide d'un outil de conception de haut niveau développé dans notre département [45], et les simulations au niveau transistor à l'aide des modèles tirés de la technologie CMOS $0.35\mu m$ fournis par la société *Austria Micro System (AMS)*. Les circuits ont été simulés à l'aide du logiciel *Cadence*, du simulateur *Spectre* pour la partie analogique, et de *Verilog* pour la partie numérique. Afin de déterminer la fréquence maximale de fonctionnement et la surface occupée par les différents blocs, la partie numérique a été synthétisée et optimisée à l'aide de l'outil de synthèse logique "Ambit".

6.2 Fonctionnement idéal du modulateur

Afin d'étudier le fonctionnement idéal du modulateur, nous considérerons, dans un premier temps, deux exemples : un modulateur multibit passe-bas et un modulateur multibit passe-bande avec les paramètres suivants : $OSR=64$, $f_e = 800MHz$, $m=3$ ($L_{LP} = m = 3$, et $L_{BP} = 2m = 6$), $B=4$ ou $M=16$, dont les zéros sont placés pour obtenir une NTF optimale (voir la section 2.3.11). Théoriquement, un tel modulateur peut atteindre une résolution de l'ordre de 19 bits. Cependant, afin de montrer l'effet des différentes sources d'erreurs, nous avons fixé une résolution objective d'environ 16 bits.

À l'aide de l'outil MATLAB "ADOC" [45] ou de l'outil "MATLAB-Schreier", voire en utilisant directement la méthode analytique (dans le cas simple), nous pouvons obtenir le modulateur souhaité. Nous nous sommes intéressés aux structures à un seul rebouclage, du fait de leur simplicité de réalisation. Soulignons néanmoins que les méthodes de correction d'erreur développées dans ce travail ne sont pas, en général, limitées à une structure particulière : elles peuvent donc être appliquées aux autres types de modulateurs.

La figure 6.1 montre, par exemple, quatre structures qui sont idéalement équivalentes :

1. un modulateur discret, passe-bas, à filtre simple
2. un modulateur discret, passe-bas, à rebouclage distribué
3. un modulateur discret, passe-bas MSCL
4. un modulateur à temps continu, passe-bas à rebouclage multiple

Le signal de sortie du modulateur, le spectre de sortie, et la densité spectrale de bruit sortant (i.e. la sortie moins l'entrée) sont représentés sur la figure 6.2 (pour le cas passe-bas à filtre simple). Le SNR et le SFDR du système sont aussi tracés à la figure 6.3 en fonction de l'amplitude de l'entrée. La résolution statique $ENOB_s = \frac{SNDR_{max}-1.76}{6.02}$ et la résolution dynamique $ENOB_d = \frac{SFDR}{6.02}$ peuvent atteindre une valeur maximale de 102dB, soit une résolution théorique de 16.5bit.

Un modulateur passe-bande équivalent centré au quart de la fréquence d'échantillonnage peut être obtenu en remplaçant "z" par " $-z^2$ " dans la fonction de transfert passe-bas discrète. Alors, autant

1. Certains des autres blocs du modulateur à temps continu ont été récemment présentés dans un autre travail de thèse dans notre laboratoire [24]

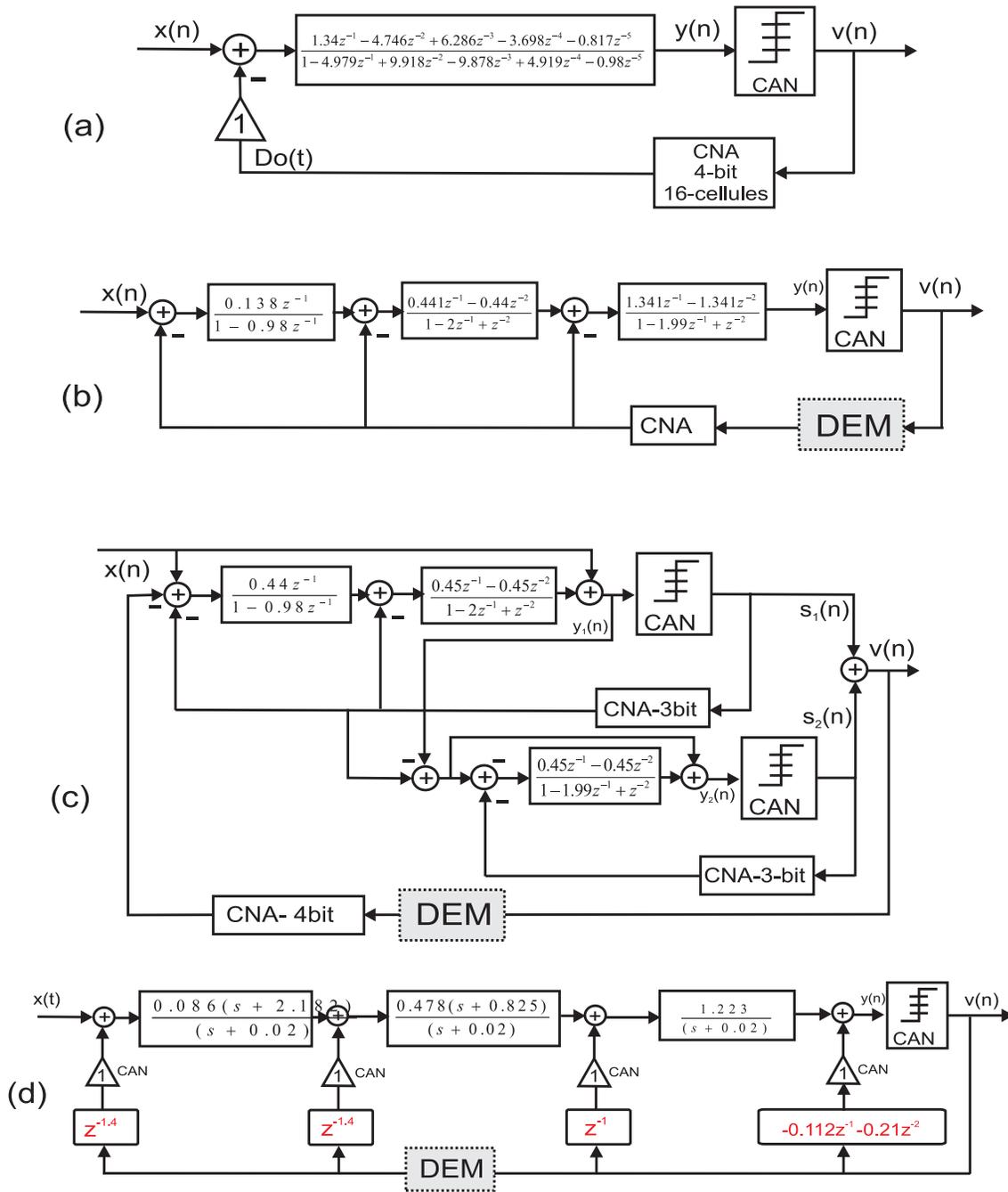


FIG. 6.1 – Structure d'un modulateur passe-bas du troisième ordre cas : a) à filtre simple, b) à reboilage multiple "CIFB", c) MSCL, d) à reboilage multiple à temps continu

de structures différentes de modulateur passe-bande sont attendus dans le cas discret ainsi que dans le cas à temps continu (à chacune des structures passe-bas présentées précédemment correspond donc une structure passe-bande équivalente). Cependant, à part ces structures équivalentes, il existe deux possibilités supplémentaires dans le cas passe-bande, comme le montre la figure 6.4. Dans le cas (a), une architecture de type *Résonateurs-Cascadés à Forme de Reboilage distribué*, "CRFB" est modifiée de sorte que les résonateurs du filtre du chemin direct soient des résonateurs purs ($\frac{z^{-1}}{(1+z^{-2})}$) et

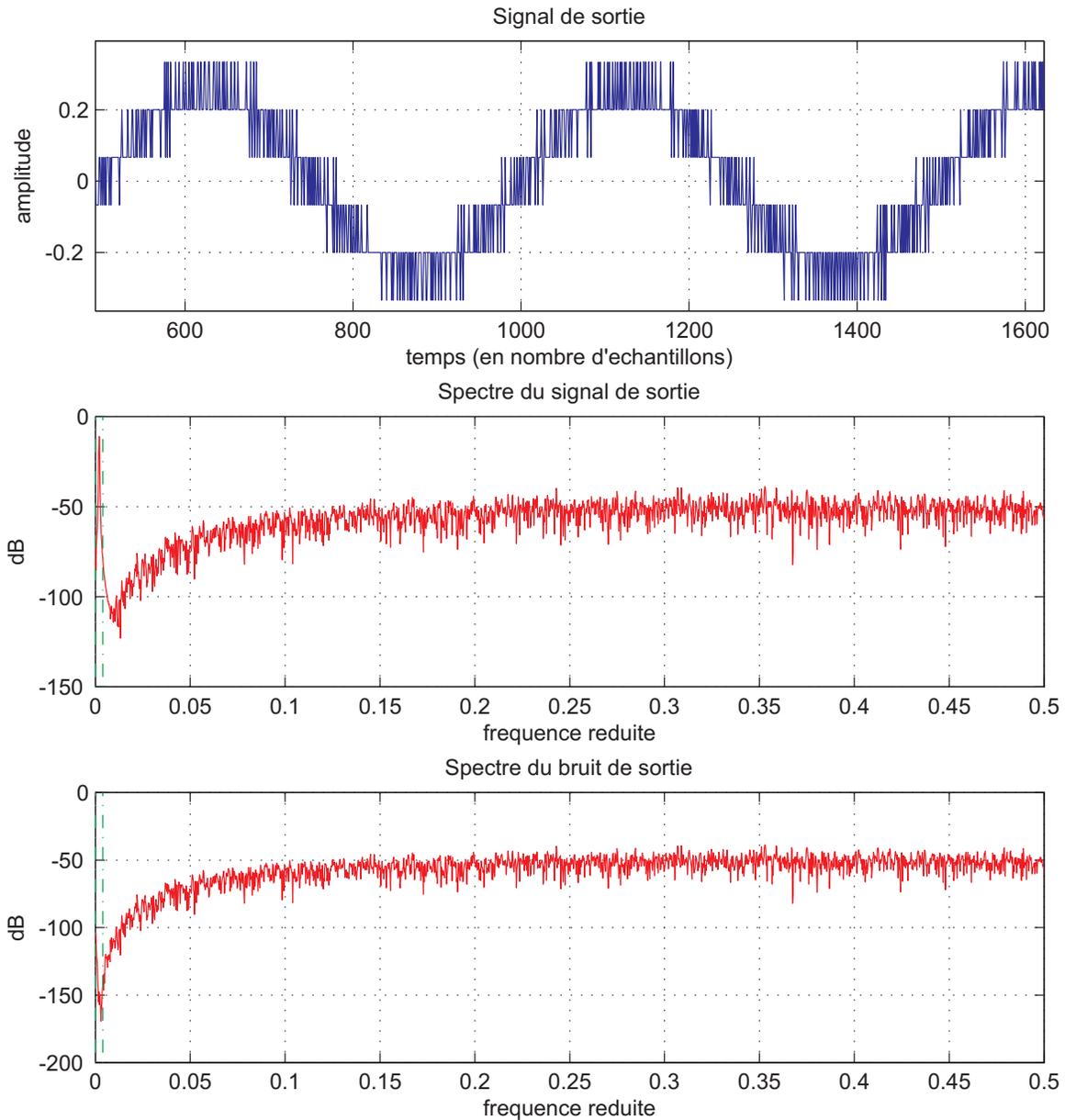


FIG. 6.2 — *Modulateur passe-bas idéal: a) signal de sortie $v(n)$, b) spectre de la sortie $PSD[v(n)]$, c) densité spectrale du bruit en sortie $PSD[v(n) - x(n)]$*

que les différents rebouclages aient des retards différents. Cette architecture, qui sera appelée dans ce travail CRFB-retard-multiple, permet de placer un bloc de DEM en amont des deux CNA les plus proches de l'entrée sans avoir à tenir compte du retard introduit dans la boucle par le processus de DEM. Par conséquent, le temps nécessaire au traitement du signal dans le bloc DEM peut durer une période sans avoir de conséquence négative. Il faut comparer cela au cas CRFB simple, où un tel traitement doit être effectué dans un temps relativement faible par rapport à la moitié de la période d'échantillonnage. En outre, la complexité de la DEM n'impose pas de limite stricte sur la fréquence maximale d'échantillonnage dans l'architecture CRFB-retard-multiple. Bien entendu, le retard maximal du bloc DEM ne doit pas, en général, dépasser la durée d'une période d'échantillon-

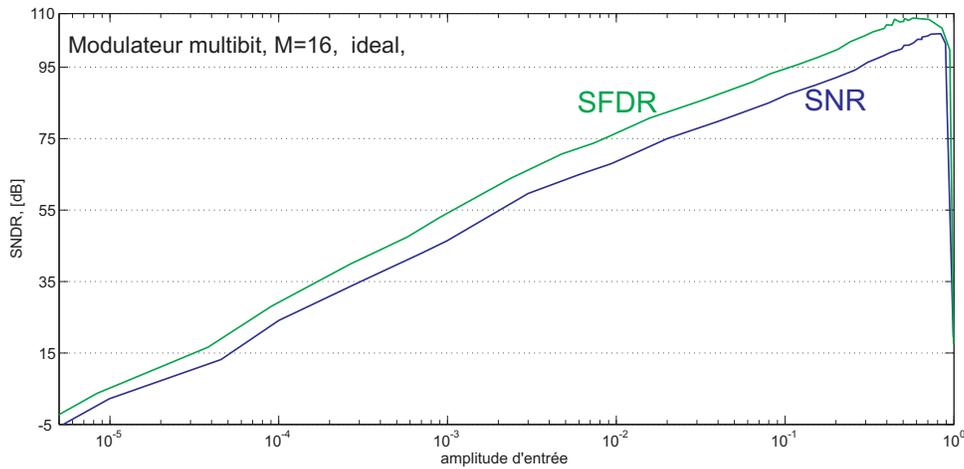


FIG. 6.3 – *SNR et SFDR du modulateur passe-bas idéal en fonction de l'amplitude de l'entrée, $OSR=64$, $M=16$*

nage. La raison en est que chaque nouvelle décision, dans toutes les méthodes de DEM à l'exception de RDEM, dépend de la décision précédente et de la valeur de l'entrée.

La structure CRFB-retard-multiple peut être particulièrement utile dans un modulateur passe-bande multibit à temps continu. Elle ouvre en effet une voie prometteuse pour le développement de modulateurs multibit à temps continu adaptés à des applications pratiques. Les figures 6.4-b et c montrent deux manières de réaliser un modulateur à temps continu passe-bande d'ordre trois. Dans le cas (b), les retards sont entiers et les résonateurs sont supposés être des passe-bande purs. Une DEM avec un retard inférieur à une période d'horloge peut être employée en amont des deux premiers CNA comme cela a été expliqué précédemment. Cependant, les résonateurs réels ne sont, en général, pas purs; il subsiste souvent un terme passe-bas associé de la forme $\frac{(a_i s + b_i)}{(s^2 + \omega_o^2)}$ même après compensation du facteur de qualité par une résistance négative [24]. La figure 6.4-c, montre un exemple de modulateur passe-bande à temps continu, où les retards des branches de rebouclage sont optimisés en tenant compte des imperfections des résonateurs ainsi que du retard de quantificateur.

La dernière architecture de modulateur passe-bande étudiée se compose d'une structure parallèle comme le montre la figure 6.4-d. Cette structure permet de réalisation des modulateurs utilisant des filtres LC intégrés [24, 49].

La figure 6.5 montre les résultats de simulation d'un modulateur $\Sigma\Delta$ passe-bande idéal :

- (a) spectre de sortie dans le cas discret décrit à la figure 6.4-a
- (b) spectre de sortie dans le cas continu décrit à la figure 6.4-b
- (c) spectre de sortie dans le cas continu décrit à la figure 6.4-c.

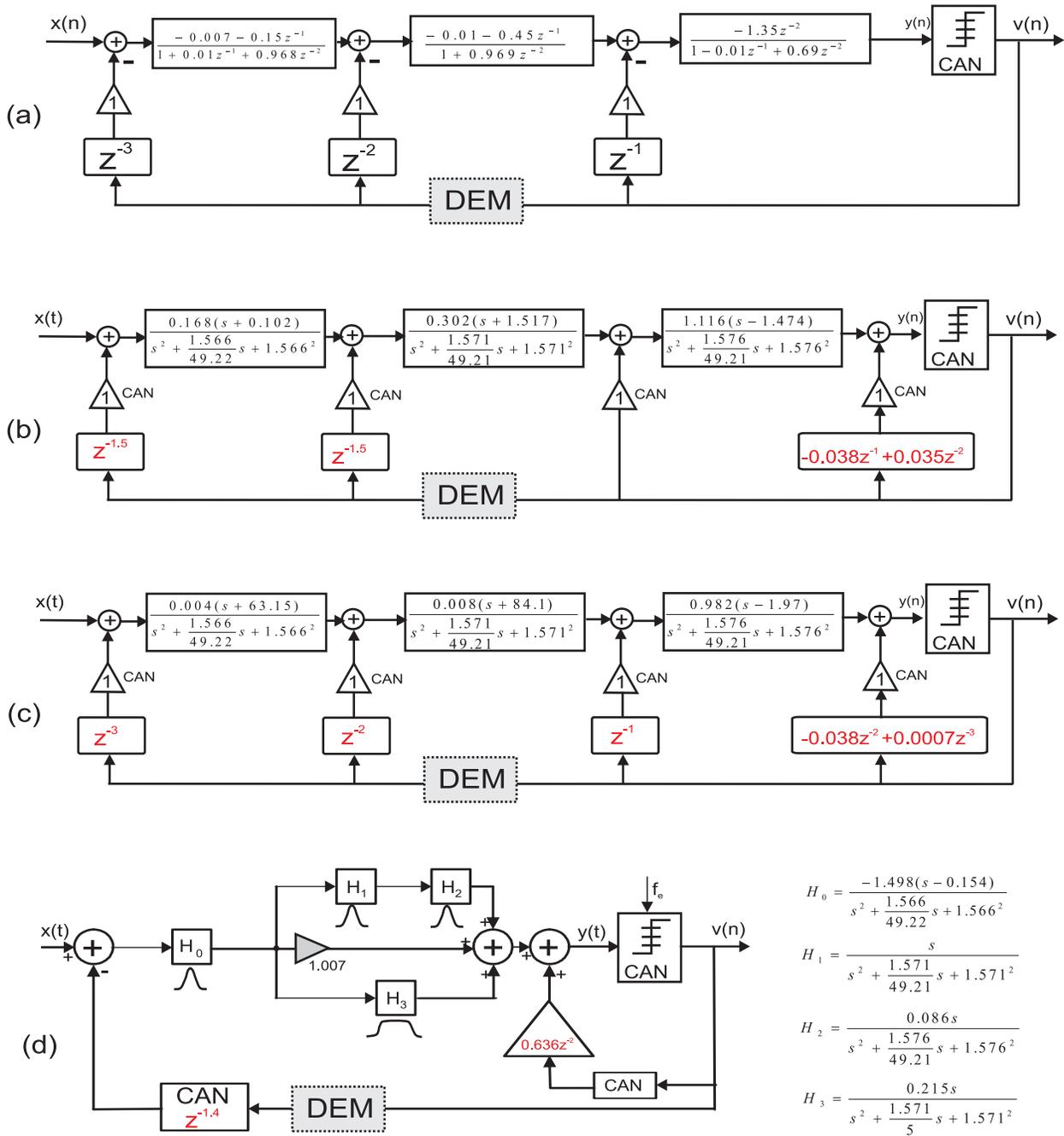


FIG. 6.4 – Structures passe-bande intéressantes : a) Cas discret, CRFB-retard-multiple, b) Cas continu, CRFB-retard-multiple avec des retards entiers, c) Cas continu, CRFB-retard-multiple avec des retard optimisés pour éliminer l'effet des imperfections des résonateurs d) Cas continu, rebouclage simple architecture parallèle

6.3 Effet des erreurs du CNA (défaut d'appariement, glitch, temps de transition, et gigue)

Cette section présente quelques techniques de compensation des erreurs des CNA multibit employées dans les modulateurs delta sigma.

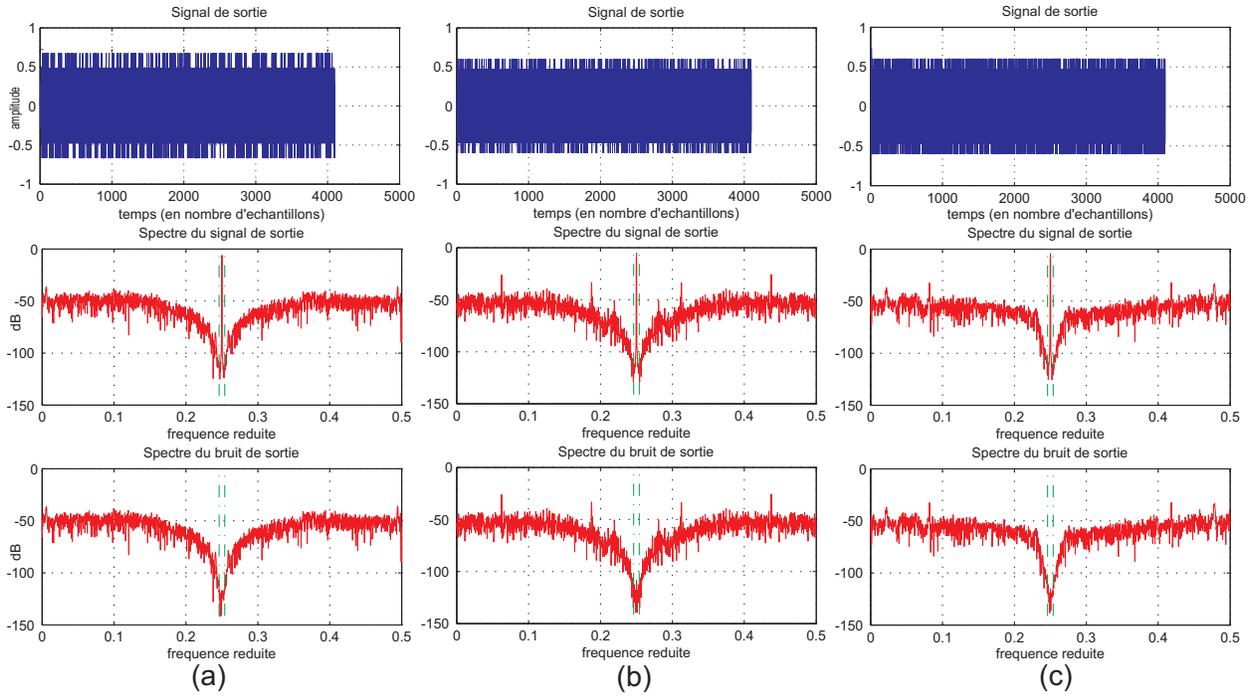


FIG. 6.5 – *Modulateur passe-bande*: a) sortie du modulateur dans le cas CRFB-retard-multiple discret, b) sortie dans le cas CRFB-retard-multiple à temps continu c) sortie dans le cas rebouclage simple architecture parallèle.

6.3.1 Compensation des erreurs statiques d'un CNA multibit

Comme nous l'avons détaillé aux chapitres 3 et 4, les défauts d'appariement entre les cellules d'un CNA limitent considérablement la qualité des modulateurs $\Sigma\Delta$ multibit. Selon la résolution du système et le taux d'appariement des cellules du CNA, on peut diminuer l'effet des erreurs d'appariement en utilisant une des méthodes de DEM. Nous avons développé quatre algorithmes de DEM au chapitre 5 pour différentes applications passe-bas et passe-bande. Les technologies actuelles de circuit intégré CMOS standard offrent une précision relative de l'ordre de 0.2 à 1 % pour des cellules à courants commutés et une précision de l'ordre de 0.1 à 0.5 % [91] pour des cellules à capacités commutées. Nous nous sommes limités au cas des CNA à sources de courant qui sont communément utilisés dans les modulateurs delta sigma à haute fréquence en particulier dans le cas continu [2]. Un modulateur $\Sigma\Delta$ multibit, avec un OSR de 64 et une résolution finale de 12 à 15 bit, nécessite au moins une DEM du premier ordre. La technique de DWA est en principe la meilleure méthode dans ce contexte (voir le chapitre 4). Mais, en raison de la présence de tons parasites, nous en avons proposé une modification simple au chapitre 5, appelée PDWA-modifiée. Une PDWA-modifiée a également été développée pour des applications passe-bande. Ces deux algorithmes ont été analysés et simulés en expliquant leurs avantages et inconvénients au même chapitre.

Cependant, afin d'obtenir une résolution finale de l'ordre de 16 bits ou plus, pour les mêmes OSR et défauts d'appariement, une méthode de correction DEM d'ordre supérieur à 1 est indispensable. Les algorithmes de MDEM et de STDEM introduits au chapitre 5, et la méthode SDEM conventionnelle,

peuvent être employés avec des filtres passe-bas, du deuxième ou du troisième ordre, ou passe-bande, du quatrième ou du sixième ordre [132, 149]. La MDEM améliore la correction de quelques dB , mais elle est plus complexe que la STDEM, ce qui n'est peut-être pas nécessaire dans notre cas ($ENOB \geq 16$, $OSR = 64$, et $M = 16$). Nous choisissons ici la STDEM qui est, selon nos connaissances, la plus simple méthode pour réaliser une DEM d'ordre supérieur [150].

La figure 6.6 représente la spectre de sortie et le SNR d'un modulateur passe-bande avec un CNA de 16 cellules ayant un défaut d'appariement de 1% avec et sans correction par les méthodes MDEM et STDEM. On constate que l'utilisation de la STDEM permet d'atteindre la résolution prévue.

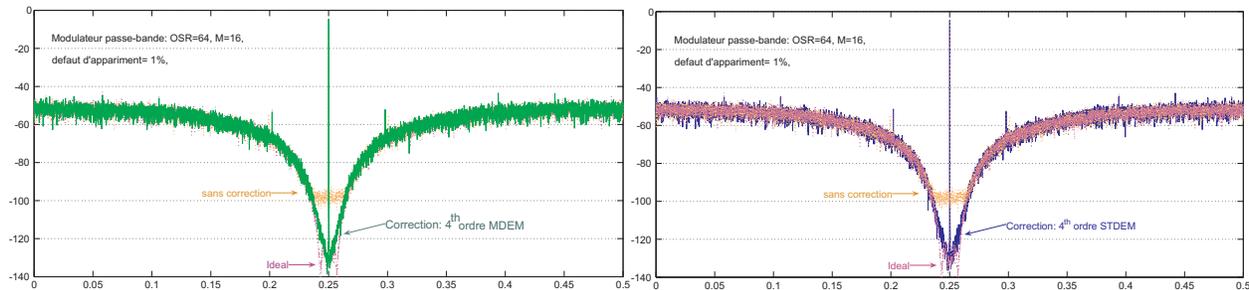


FIG. 6.6 — Spectre de la sortie d'un modulateur multibit passe-bande avec un CNA désapparié de 1%, a) corrigée par la MDEM, b) corrigée par la STDEM.

6.3.2 Implantation de méthode DEM au niveau circuit

La MDEM et la STDEM ont été implantées au niveau circuit (standard cell). De plus, afin de pouvoir faire une comparaison objective, la méthode SDEM a aussi été implantée dans le même environnement. Cette implantation permet de mettre en évidence les différences pratiques entre ces algorithmes qui sont, à ce jour, les trois seuls algorithmes capables de réaliser une DEM d'ordre supérieur. La comparaison sera limitée aux trois caractéristiques suivantes :

1. la surface occupée par chaque algorithme,
2. la fréquence maximale de fonctionnement,
3. la performance de la correction des erreurs des cellules.

Le passage d'un bloc numérique complexe du niveau système au niveau circuit consiste en ces cinq étapes :

1. Détermination de la taille des registres,
2. Détermination d'une stratégie d'implantation pour les fonctions qui ne sont pas directement réalisables au niveau portes logiques, par exemple celles de tri,
3. Ecriture d'un code RTL (Verilog ou VHDL),
4. Conversion du code RTL en portes numériques et bascules,
5. Routage des portes et des bascules.

Dans un premier temps, il faut simuler ces algorithmes avec des registres entiers au niveau système, ce qui permet d'estimer la largeur de chacun des registres en nombre de bits. Comme on l'a dit

auparavant, les deux principaux avantages de ces trois algorithmes sont qu'ils nécessitent des registres de faible taille et qu'ils peuvent être implantés sous forme de filtres numériques à coefficients entiers ne nécessitant aucune multiplication. Pour un filtre de DEM passe-bande d'ordre 4, la fonction de transfert de défaut d'appariement ("*Mismatch Transfer Function, MTF*") est choisie dans le cas le plus simple comme suit :

$$MTF(z) = H(z) = 1 + 2z^{-2} + z^{-4} \tag{6.1}$$

La MTF synthétisée est représentée à la figure 6.7, pour un bloc de tri SSB et pour un bloc de type ESB ou SB. Cette même MTF est utilisée pour les trois algorithmes afin que leurs performances

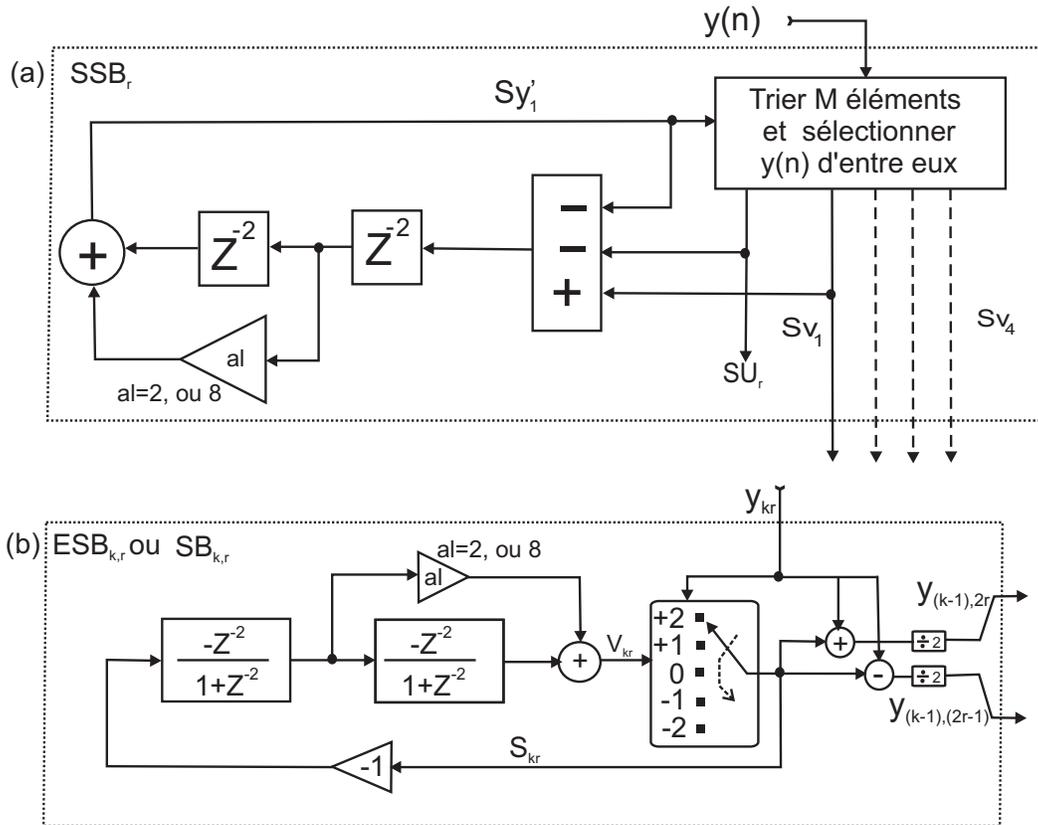


FIG. 6.7 – Schéma bloc de la MTF: a) SSB employé dans la SDEM ou dans la MDEM, b) ESB ou SB employé dans la STDEM ou dans les premières couches de la MDEM

soient comparables. D'après l'estimation théorique donnée au chapitre 5, confirmée par plusieurs calculs réalisés avec MATLAB en précision fixe, la taille nécessaire des registres de chaque algorithme est rappelée au tableau 6.1. On constate que les largeurs des registre nécessaires pour la SDEM et pour la STDEM sont les plus faibles, et que la largeur estimée pour le cas MDEM est assez modérée. En fait, la complexité des circuits dépend de deux autres aspects des algorithmes :

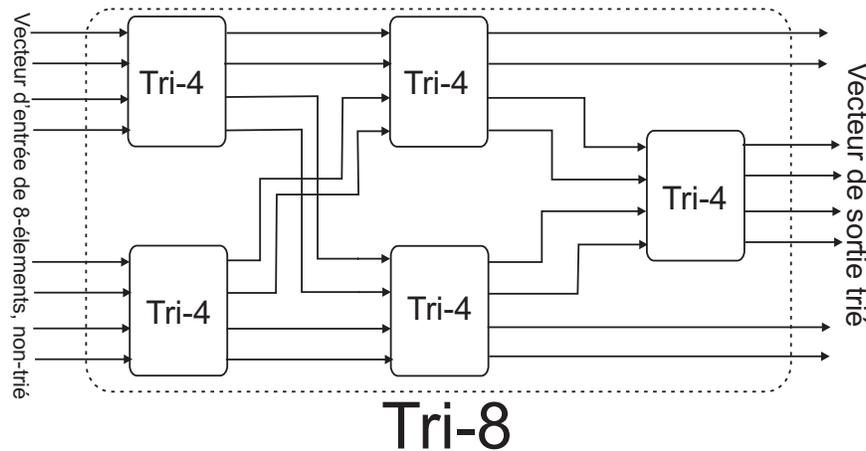
- le type de comparaison ou de quantification des sorties des filtres,
- le nombre de sorties des filtres qui doivent être comparés.

TAB. 6.1 – Taille nécessaire des registres des différents algorithmes d'ordre 4 passe-bande :

Méthode	taille de registre
SDEM	6 bit
STDEM	6 bit
MDEM	8 bit

Pour la STDEM, il s'agit d'une simple comparaison de signe et de deux bits de poids fort, alors que, dans le cas SDEM et pour la dernière couche de MDEM, il s'agit d'une comparaison numérique complète des sorties des filtres. Une comparaison numérique complète peut considérablement augmenter la complexité ainsi qu'augmenter le temps de traitement. Cela peut être encore plus limitant si le nombre d'éléments à comparer est grand. C'est par exemple le cas pour une SDEM pure où M éléments doivent être comparés. Dans le cas de la MDEM, le nombre des éléments qui doivent être comparés est faible, de l'ordre de 4 (voir la figure 5.9).

Il existe plusieurs algorithmes permettant de trier un vecteur de données. La plupart peuvent être utilisés au niveau software. Dans le cas d'une implantation au niveau circuit, la meilleure stratégie est de diviser le vecteur de données en différents groupes de sorte que chaque groupe se compose de 4 éléments [130]. La figure 6.8 montre une stratégie de tri d'un vecteur à 8 éléments. Chaque bloc de tri

**FIG. 6.8** – Schéma bloc du tri d'un vecteur de 8 éléments n'utilisant que des blocs de tri de 4 éléments.

(SDEM) se compose de 5 blocs de tri à 4 éléments, notés Tri-4. Chaque bloc Tri-4 utilisé dans la figure 6.8 contient 6 comparateurs numériques et une partie logique combinatoire qui est chargée de placer des éléments dans l'ordre souhaité en fonction des 6 résultats de comparaison. Chaque comparateur compare deux nombres binaires. Chaque nombre a normalement une largeur de 6 bits pour le cas SDEM ou 8 bit pour le cas MDEM. Notons qu'un algorithme d'appariement dynamique des 8 cellules a seulement besoin de 2 Tri-4 distincts pour le cas MDEM pratique, alors qu'il a besoin de 5 Tri-4 en cascade pour le cas SDEM.

La méthode STDEM n'a besoin ni du bloc de Tri-4 ni du comparateur numérique, ce qui est un avantage très important par rapport à la SDEM et à la MDEM. Le nombre de blocs Tri-4 varie d'une façon linéaire dans l'algorithme de MDEM, et d'une façon exponentielle dans l'algorithme de SDEM,

ce qui met en évidence leur différence de complexité. Par exemple, la figure 6.9 illustre une méthode de tri d'un vecteur à 16 éléments : elle nécessite 25 blocs de Tri-4 en cascade soit 6.25 fois plus que pour la MDEM. Dans la méthode MDEM, la complexité de la partie TDEM est négligeable devant

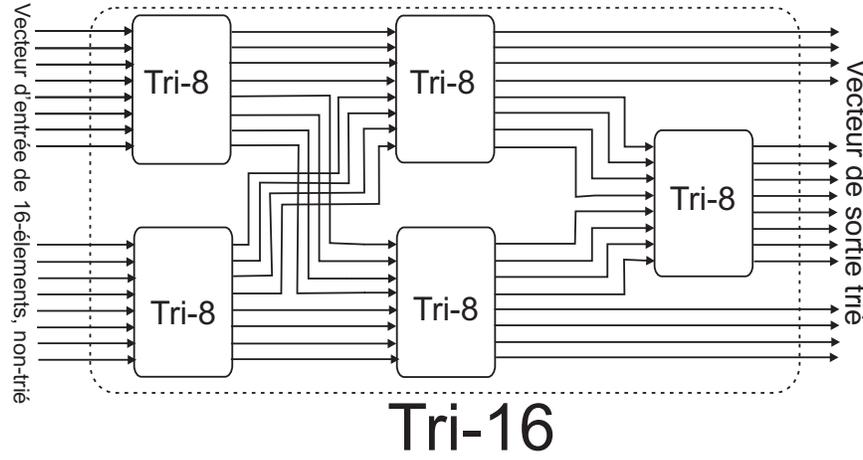


FIG. 6.9 – Tri d'un vecteur de 16 éléments en utilisant des blocs de tri de 8 éléments.

celle de la partie SDEM. Ainsi, la complexité des deux méthodes d'appariement dynamiques du CNA peut être ramenée au nombre des blocs Tri-4 nécessaires, comme résumé dans le tableau 6.2. On

TAB. 6.2 – Nombres de blocs Tri-4 nécessaires aux méthodes MDEM et SDEM en fonction du nombre de cellules du CNA (M)

M, cellules	dans la MDEM	dans la SDEM	complexité relative
4	1	1	1
8	2	5	2.5
16	4	25	6.25
32	8	125	15.63
64	16	625	39.1

remarque que l'avantage, en matière de complexité, de la méthode MDEM développée dans ce travail est considérable, en particulier, pour un nombre de cellules élevé. Il faut noter que la complexité de la méthode STDEM, qui est plus faible que de celle de MDEM, reste toutefois assez inférieur à la complexité d'une méthode SDEM conventionnelle.

Une fois la taille des registres déterminée et la stratégie optimale de passage au niveau circuit choisie, les trois algorithmes ont été codés en langage Verilog et ont été appliqués à un modulateur $\Sigma\Delta$ contenant un CNA désapparié, simulé sous spectre (Cadence). Ceci doit permettre d'évaluer l'impact de la réalisation de chacune des méthodes avec de vrais registres ayant une résolution (taille en bit) limitée. En vérifiant leur rendements maximums, la prévision théorique de la résolution des registres (tableau 6.1) est bien confirmée dans tous les cas. De plus, l'augmentation de la résolution indiquée ci-dessus n'apporte pas d'amélioration visible à la performance des algorithmes. Par contre, la diminution de l'ordre de 1 bit des résolutions prévues n'a qu'une conséquence négligeable. Par exemple, la méthode STDEM peut très bien fonctionner avec des registres de 5 bits à la place de ceux de 6.

Afin d'effectuer les simulation dans un cadre réaliste, nous avons appliqué les algorithmes proposés à un CNA 3 bit. Celui-ci est utilisé dans un modulateur à temps continu à architecture parallèle (figure 6.10) qui a été développé dans notre département [24]. La densité spectrale de puissance du signal de

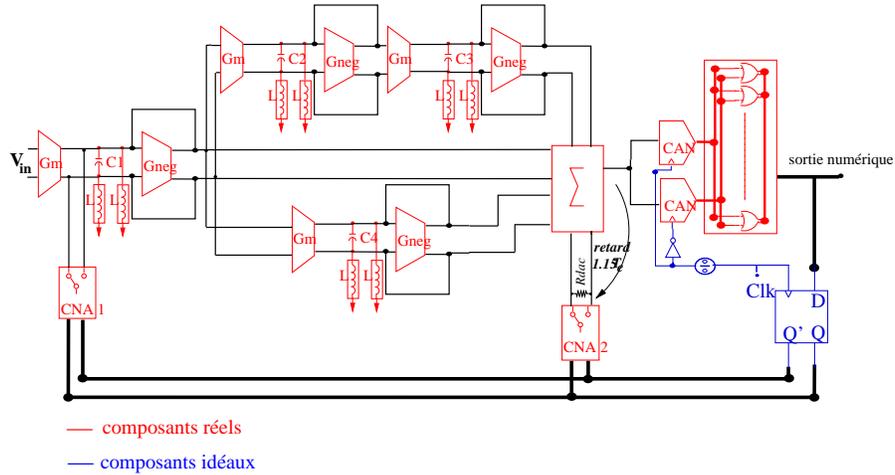


FIG. 6.10 – Schéma bloc du modulateur $\Sigma\Delta$ multibit passe-bande du 6^{ième} ordre à filtre LC.

sortie du modulateur est représentée à la figure 6.11 dans le cas d'un CNA idéal et dans celui d'un CNA désapparié avec ou sans correction par les méthodes MDEM et STDEM. On remarque que la mise en forme du bruit de défaut d'appariement du CNA est bien suivie dans les deux cas.

Une fois que le fonctionnement des algorithmes au niveau RTL est confirmé, on peut remplacer le bloc de haut niveau par de vraies portes logiques. Nous avons utilisé l'outil de synthèse logique Ambient. Ensuite chaque bloc de DEM est optimisé en vue de déterminer sa fréquence maximale de fonctionnement et sa surface. La figure 6.12 donne les surfaces obtenues pour les différentes méthodes, alimentations incluses. Les routages finaux de ces circuits y compris les fils d'alimentations et les pins sont aussi représentés aux figures 6.13, 6.14-a et b, à la même échelle identique.

En tenant compte des effets parasites de routage, la vitesse maximale de chaque bloc de DEM peut être estimée. Les chemins les plus longs des trois méthodes pour un CNA de 3 bit sont donnés à la figure 6.15.

Enfin les performances de ces méthodes sont résumées dans le tableau 6.3. En fait pour le modu-

TAB. 6.3 – Comparaison des méthodes STDEM, MDEM et SDEM pour la correction des défauts d'appariement d'un CNA 3 bit.

Méthode utilisée	surface minimale	fréquence maximale	correction de 1% d'erreur d'appariement	correction de 3% d'erreur (simulée au niveau système)
STDEM	0.24 (mm) ²	330MHz	> 23 dB	33 dB
MDEM	0.523 (mm) ²	125MHz	> 23 dB	35 dB
SDEM	0.69 (mm) ²	250MHz	> 23 dB	35 dB

lateur que nous avons utilisé, un taux d'erreur d'appariement de 1% ne dégrade pas beaucoup sa performance, car sa résolution idéale est limitée. Ainsi, nous présenterons la performance des trois

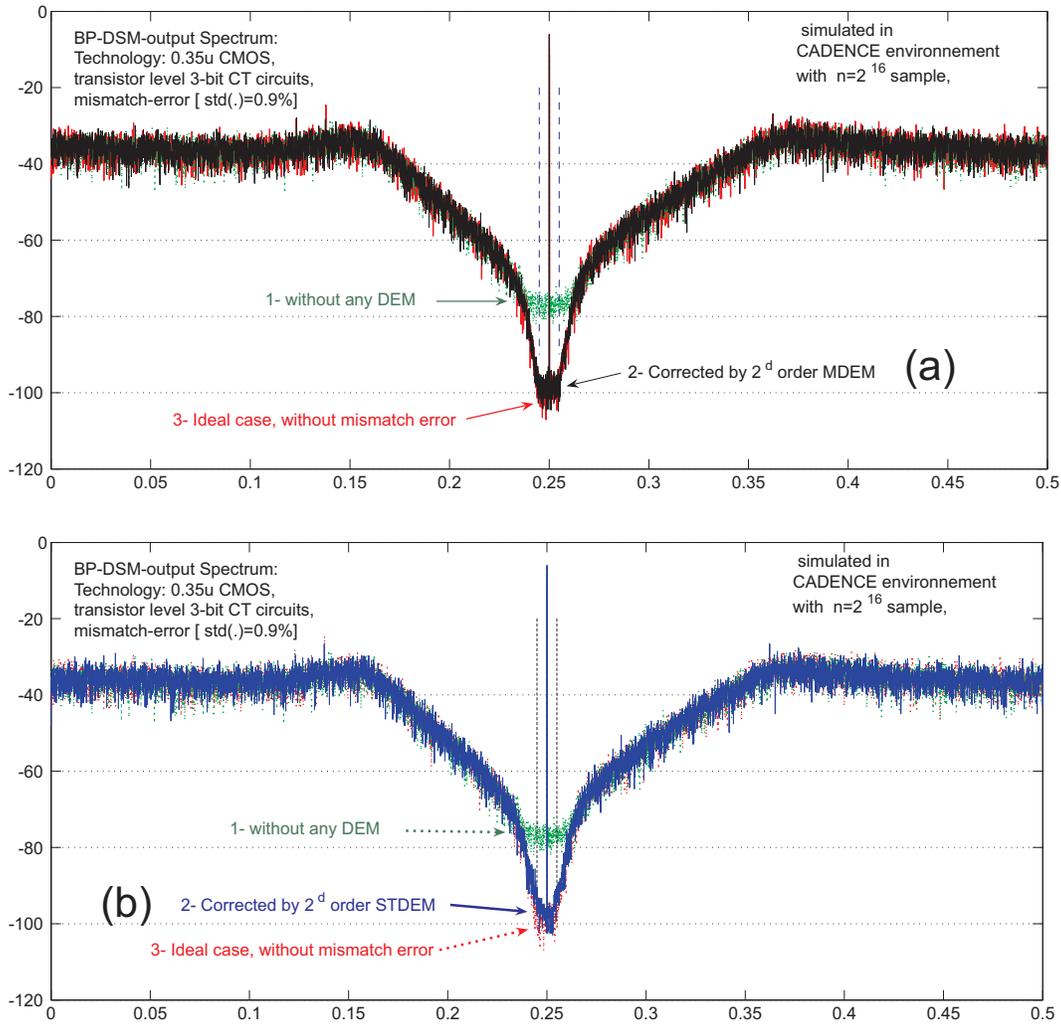


FIG. 6.11 – Résultat de simulation au niveau Transistor dans l’environnement Cadence, pour un modulateur multibit en utilisant un algorithme passe-bande d’ordre quatre a) MDEM b) STDEM

Summary Area Report , Module : 'SDEM_arbit'			Summary Area Report , Module : 'MDEM_arbit'			Summary Area Report , Module : 'STDEM_arbit'		
Block report for module	Current	Cumulative	Block report for module	Current	Cumulative	Block report for module	Current	Cumulative
Number of combinational instances	1893	4519	Number of combinational instances	374	2693	Number of combinational instances	370	1079
Number of noncombinational instances	490	490	Number of noncombinational instances	432	432	Number of noncombinational instances	278	278
Number of hierarchical instances	43	43	Number of hierarchical instances	36	36	Number of hierarchical instances	8	8
Total number of instances	2426	5052	Total number of instances	842	3161	Total number of instances	656	1365
Area of combinational cells	7346.00	16728.00	Area of combinational cells	1374.00	10153.00	Area of combinational cells	1396.00	4524.00
Area of non-combinational cells	7369.00	7369.00	Area of non-combinational cells	7250.00	7250.00	Area of non-combinational cells	4544.00	4544.00
Total cell area	14715.00	24097.00	Total cell area	8624.00	17403.00	Total cell area	5940.00	9068.00
Number of nets	2516	5154	Number of nets	862	3358	Number of nets	715	1524
Total area	14715.00	24097.00	Total area	8624.00	17403.00	Total area	5940.00	9068.00

FIG. 6.12 – Compte-rendu d’Ambit, donnant la surface minimale nécessaire à l’implantation d’un algorithme de DEM pour un CNA 3 bit : a) STDEM b) MDEM c) SDEM

algorithmes dans d’autres cas simulés au niveau système au tableau 6.3. On constate que la méthode STDEM est le meilleur choix en ce qu’elle, avec une plus faible complexité, offre à peu près la même performance qu’une méthode comme la SDEM. Cette différence sera encore plus marquée à mesure que la résolution du CNA augmente, comme le prévoit le tableau 6.2.

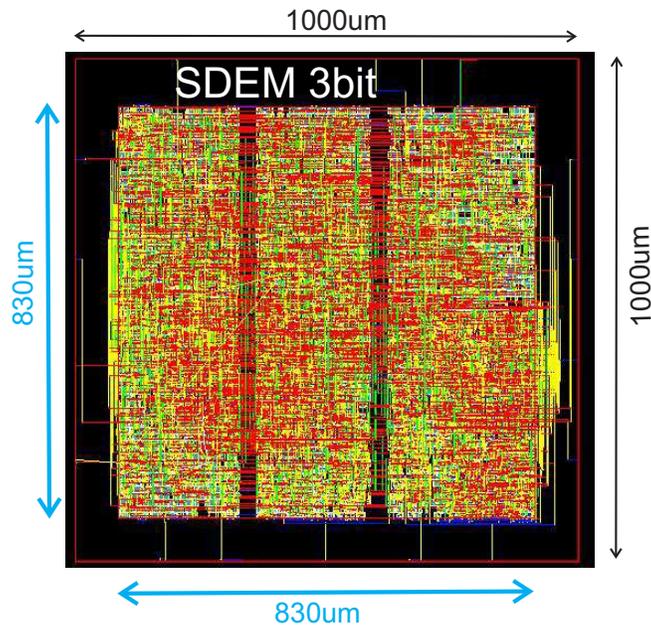


FIG. 6.13 – Routage final pour les circuits conçus de la technique SDEM dans le cas $M=8$, (area = 0.69mm^2).

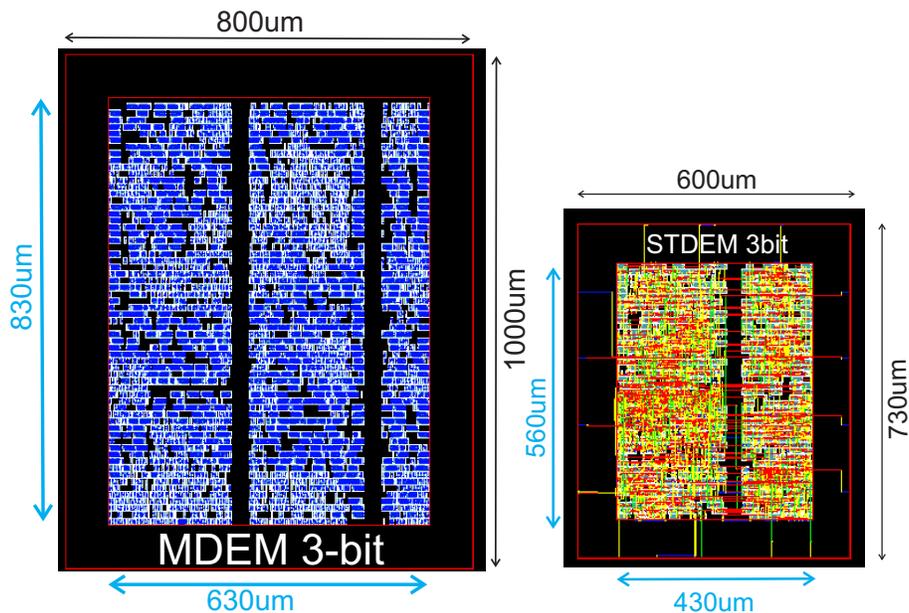


FIG. 6.14 – a) Surface final pour les circuits conçus de la technique MDEM dans le cas $M=8$, (area = 0.53mm^2), b) routage final pour les circuits conçus de la technique STDEM dans le cas $M=8$, (area = 0.24mm^2).

6.3.3 Compensation des erreurs dynamiques d'un CNA multibit

Le schéma d'un CNA 3 bits thermométrique est donné à la figure 6.16. Ce CNA est constitué de 8 sources de courant identiques (U_1 à U_8). Les sources de courant sont commandées par le code numérique (SV_1 à SV_8) délivré par la sortie du quantificateur qui est réarrangée dans le bloc d'une

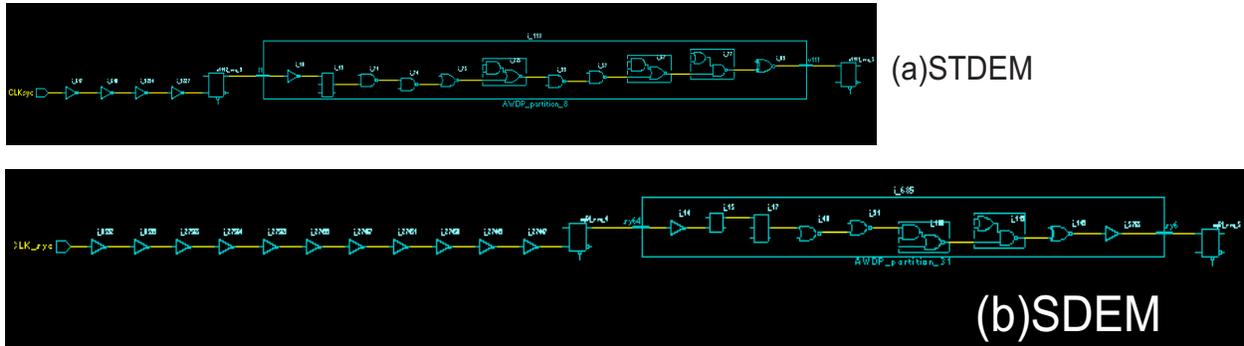


FIG. 6.15 – Compte-rendu d’Ambit pour les chemins les plus longs : a)STDEM b)SDEM

éventuelle DEM. Comme cela a été expliqué précédemment, chacune des sources de courant présente une erreur statique d’appariement qui peut être compensée par une méthode de DEM. Par souci de simplicité, nous supposons que cette erreur statique est nulle pour nous concentrer sur la compensation des erreurs dynamiques liées à la commutation des sources.

Pour le modulateur représenté à la figure 6.10, chaque source de courant unitaire D_i délivre un

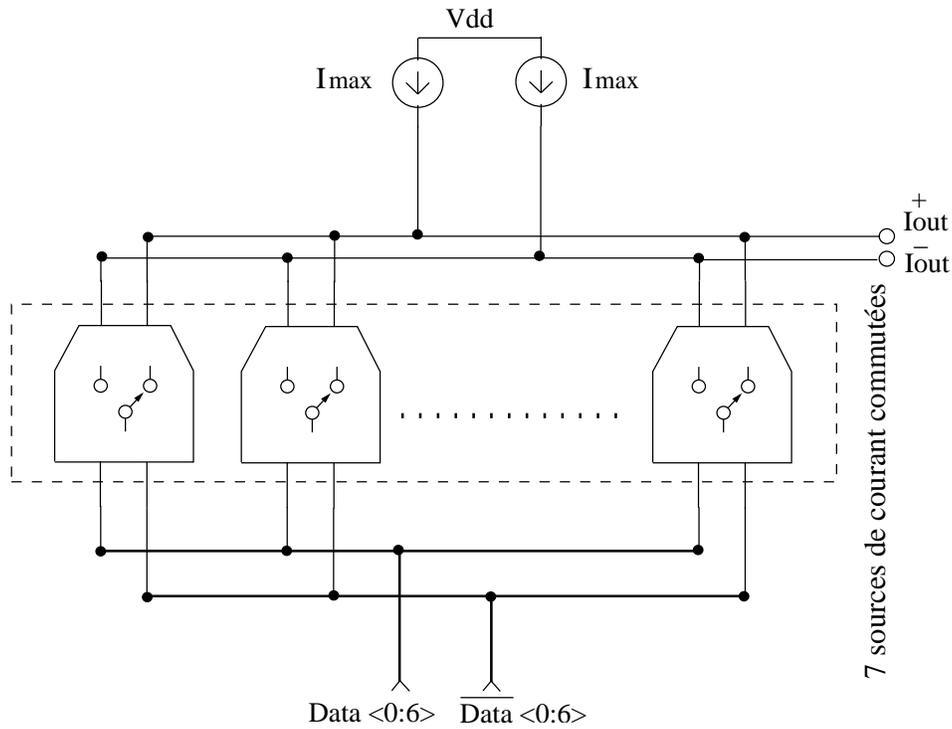


FIG. 6.16 – Circuit CNA 3-bit à code thermométrique.

courant dont la valeur peut être estimée par l’équation suivante :

$$I_u = \frac{G_m \Delta V_{in}}{2.M} \tag{6.2}$$

avec

G_m : transconductance de l'étage se trouvant entre l'entrée du modulateur et le CNA,
 ΔV_{in} : dynamique d'entrée,
 M : nombre de cellules du CAN et du CNA.

En remplaçant par les valeurs obtenues et les spécifications fixées à l'étape de conception au niveau système souhaité, le courant I_u est égal à $\pm 31\mu A$ et le courant de sortie globale du CNA varie entre $\pm 248\mu A$.

Le noyau de chaque cellule se compose d'une paire différentielle (M_1 et M_2) fonctionnant en commutateur, et d'une source de courant fixe (M_0). Dans un CNA précis et rapide cependant, la présence d'une cellule de courant commutée introduit des erreurs dynamiques. Certaines modifications sont rendues nécessaires. Ainsi que le montre la figure 6.17. Les trois modifications proposées sont :

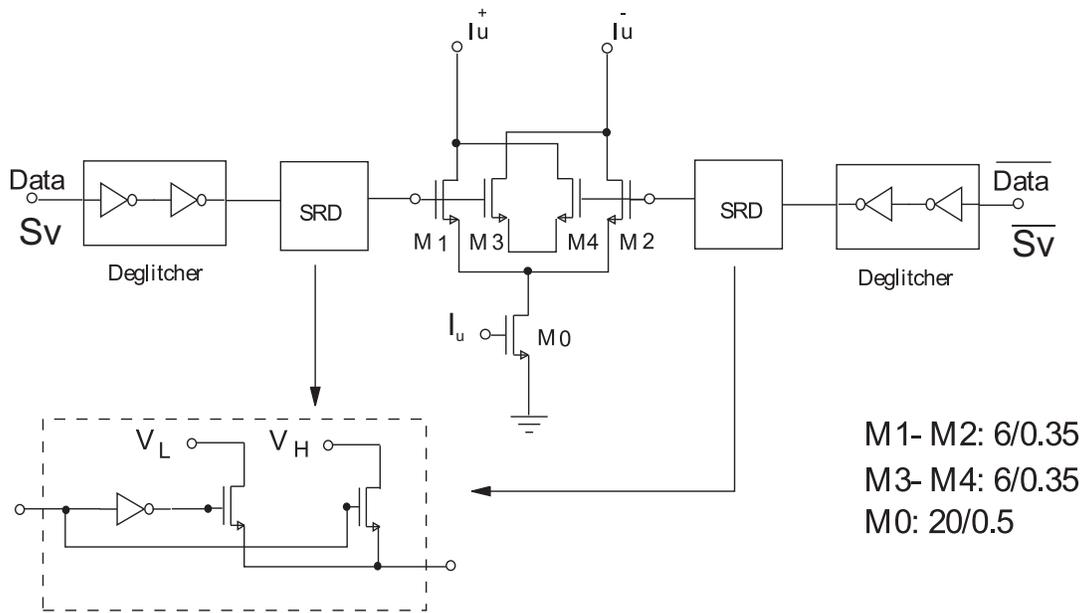


FIG. 6.17 – Source de courant commutée.

- paire différentielle fantôme
- anti glitch, "Deglitcher"
- SRD ou "Swing-Reduced-Driver"

Paire différentielle fantôme

Une source de courant commutée doit être dimensionnée de façon à avoir un temps d'établissement optimal et une grande impédance de sortie [103, 105]. Cette optimisation n'est pas suffisante dans beaucoup de cas, il faut donc compenser le temps nécessaire de transition des charge des canaux des transistors interrupteurs, par exemple, avec une paire différentielle fantôme composé de deux transistors de compensation $M3$ et $M4$ (*Dummy transistor*). Les dimensions des transistors fantômes et des commutateurs sont supposées identiques mais cette paire fantôme est excitée en sens opposé à la paire principale. Par conséquent, les charges des canaux aux instants de commutation s'écoule

sans provoquer de retard de transition des sorties [91]. La paire fantôme dispose en général d'une très grande résistance de sortie mais présente une capacité parasite de l'ordre de celle de la paire principale. Ainsi, comme on l'a analysé au chapitre 3, la paire fantôme introduit un rapprochement de partie imaginaire des impédances de sortie des cellules pour les deux états "off" et "on", ce qui contribue à améliorer la précision du CNA [31].

Déglitcher

Le phénomène de "glitch" est normalement réduit dans un CNA thermométrique simple, mais il devient extrêmement important si on utilise le même CNA avec une méthode de DEM, comme cela a été expliqué au chapitre 3. Il y a donc deux stratégies de compensation : utiliser un bloc d'ajustement aux instants de transition des cellules, bloc communément appelé "déglitcher", ou utiliser un CNA de type RZ. Le "degitcher" permet d'obtenir des temps de montée et de descente indépendants et par conséquent de réduire les impulsions transitoires (glitch) d'où son nom. Il est composé de deux inverseurs successifs dont les dimensions (les rapports WL de transistor PMOS et de transistor NMOS) sont ajustées à cet effet. Le degitcher ajoute un petit retard dans toutes les transition de l'état "on" à l'état "off" des cellules afin d'obtenir une transition douce de la branche opposés de l'état "off" à l'état "on". De ce fait, le tension de drain du transistor M_0 ne varie pas beaucoup et en particulier la source de courant de transistor M_0 ne sort jamais de sa région active (saturation). Ce dernier effet est essentiel à la minimisation des erreurs de glitch et des impulsions transitoires. La figure 6.18-(a et c) montre le détail des circuits du "degitcher" et des signaux entrants et sortants.

Comme nous l'avons dit auparavant, une autre technique, dans laquelle le phénomène de glitch n'a pas un caractère limitant, consiste à utiliser un CNA de type RZ. Une telle structure RZ n'a pas en général besoin un bloc de degitcher, car il n'y a pas de glitch du fait de la remise à zéro. Cependant, un degitcher peut améliorer la vitesse de commutation et réduire les impulsions transitoires. En revanche, un CNA RZ possède peu d'intérêt pour des modulateurs de haute fréquence car il est sujet à la gigue d'horloge. Nous allons introduire par la suite un genre de CNA semi-RZ après examen d'un bloc SRD de SRD.

SRD

Dans le cas d'un CNA simple, les niveaux numériques (0 et V_{dd}) sont acheminés jusqu'à la grille du transistor interrupteur. Par conséquent, les grilles des transistors, par exemple, M_1 et M_2 à la figure 6.17, doivent supporter une forte variation de tension lors des instants de commutation alors qu'ils ont besoin de moins d'un volt. L'impulsion transitoire, en particulier l'effet d'injection de charge CFT, augmente proportionnellement à la variation de la tension de grille, alors qu'une variation d'amplitude modérée permet de une commutation propre. Il existe plusieurs méthodes de compensation notamment en utilisant un bloc SRD (*Swing Reduced Driver*) [154–156]. Nous proposons un nouveau type de SRD, qui permet de réduire la dynamique des signaux appliqués à l'entrée de la paire différentielle de la source de courant commutée. Chaque SRD est composé de deux interrupteurs préalables chargés de transmettre deux niveaux de tension V_L ou V_H sur

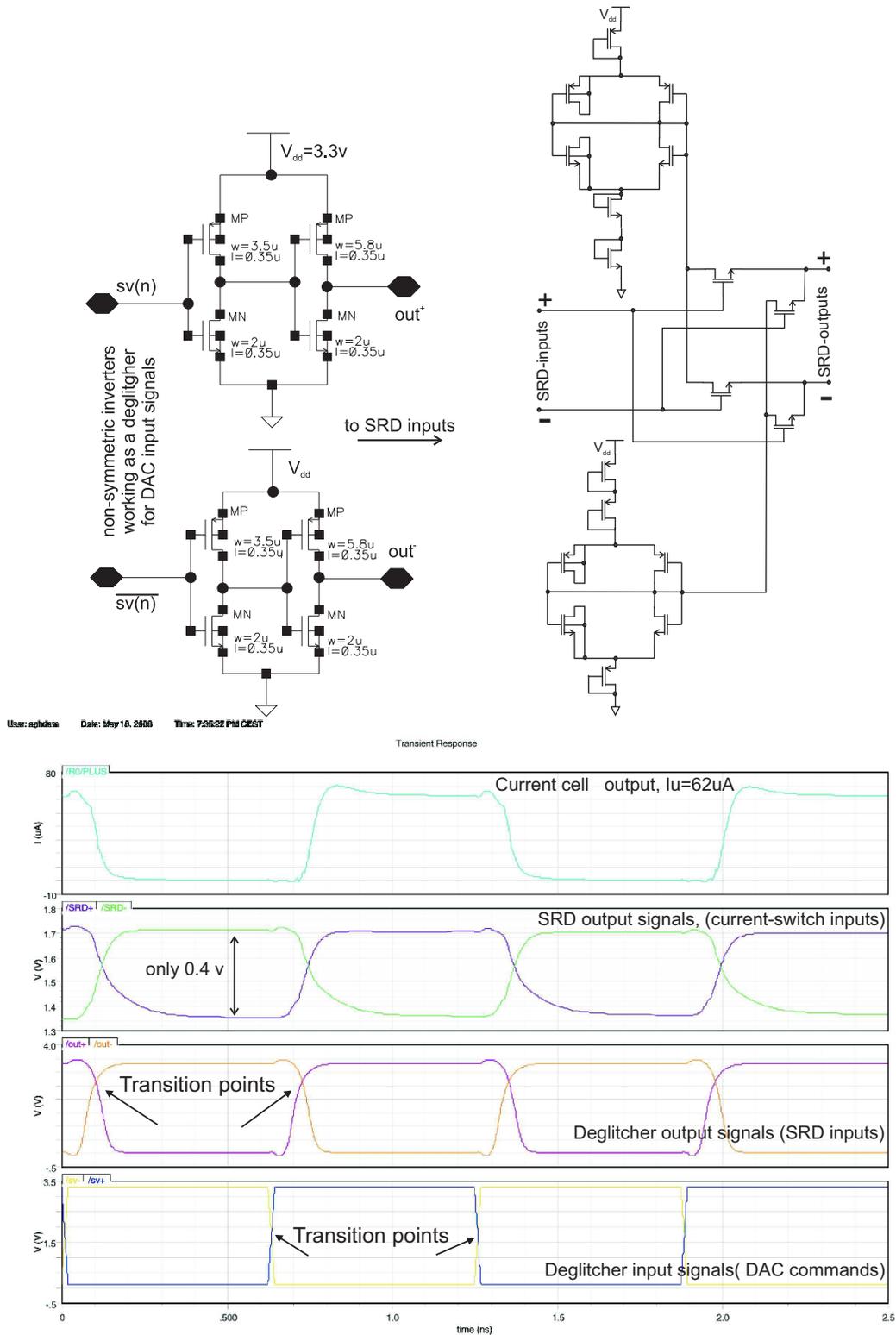


FIG. 6.18 – a) Circuit d'un Déglitcher simple, b) circuit d'un SRD simple, c) les signaux d'entrants et les signaux sortants du deglitcher et du SRD, et de la sortie d'un cellule courant commuté

les grilles des interrupteurs principaux comme le montre la figure 6.18-b et c. Chaque interrupteur préalable fonctionne en amplificateur de type drain commun (ou *Source Follower*) pendant la durée de

transition, avec la rapidité requise. Chacune des tensions fixe V_L ou V_H est précédée par un inverseur CMOS qui se trouve au milieu de certains transistors en série. Chaque transistor série (une connexion courte entre son drain et sa grille), joue un rôle de décalage de tension ("*Level-Shifter*") à l'instar d'un diode. L'inverseur CMOS a aussi une connexion courte entre son entrée et sa sortie qui joue un rôle de diviseur de tension [91]. On peut y ajouter des capacités (*bypass*) en utilisant des transistors MOS toujours en région "off" qui ne consomment pas. Les dimension de l'inverseur et des transistors séries sont aussi optimisées afin de limiter leur consommation. Il faut noter que le SRD n'a besoin de valeur précise ni pour V_L ni pour V_H . En général une précision relative de l'ordre de 10% et une consommation négligeable suffisent à un SRD souhaitable. Dans notre cas les niveaux de tension commutant sont alors optimisés entre 1.4V (V_L) et 1.8V (V_H). Le SRD proposé permet de réduire les effets d'injection de charges et de la durée de la de commutation de sorte que la cellule du CNA obtenue peut commuter facilement même avec une horloge de fréquence supérieure à 1.5 GHz. La réponse temporelle du CNA est représentée sur la figure 6.19 dans le cas d'une fréquence d'horloge égale à 1.2GHz. Comme on peut le constater, la réponse du CNA sur une période est proche d'une réponse indicielle dans le cas idéal, le temps de propagation est égal à 160ps pour l'ensemble "CNA", "deglitcher", et "SRD".

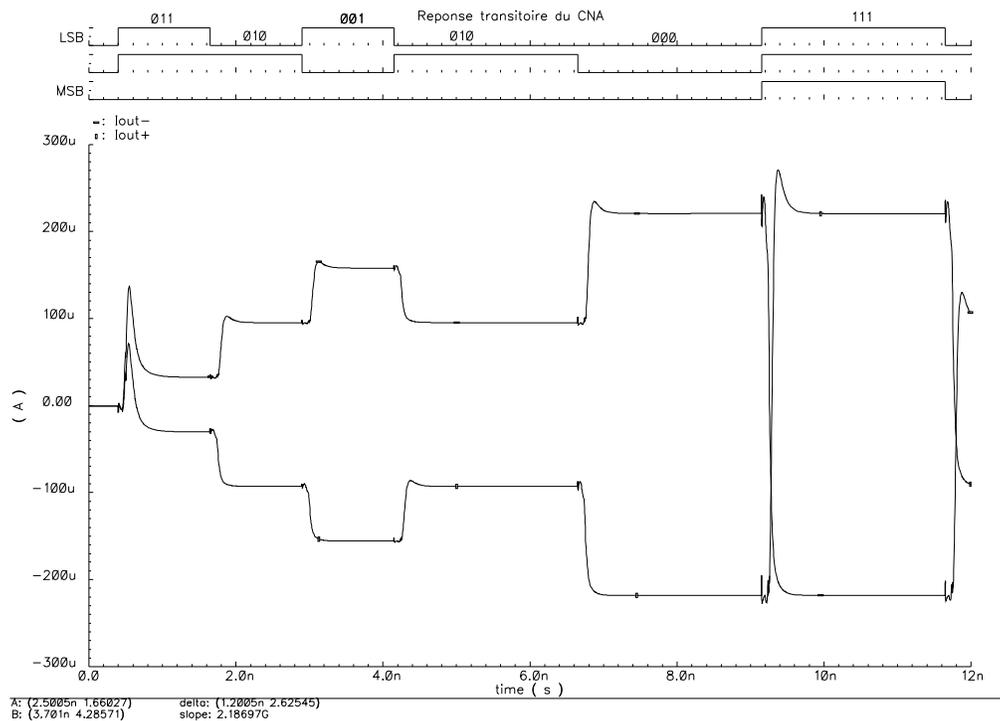


FIG. 6.19 – Réponse temporelle du CNA.

Architecture Semi-RZ

La performance d'un CNA NRZ est fortement altérée par la présence d'éventuels "glitches" et la différence entre le temps de transition $t_{on} - t_{off}$. Ces erreurs sont particulièrement importantes dans

les applications haute fréquence d'un CNA thermométrique avec une DEM, comme on l'a expliqué au chapitre 3. Cependant, un CNA de type RZ n'est pas sensible à ces deux effets par conception. En revanche, le CNA RZ est d'avantage sujet à la gigue d'horloge, car deux impulsions de gigue apparaissent à chaque période d'amplitude plus élevée que pour le CNA NRZ (l'erreur augmente 4 fois plus). De plus, les éléments actifs d'un modulateur basé sur l'utilisation d'un CNA RZ, posent plus de problèmes que dans le cas d'un modulateur NRZ, notamment en ce qui concerne la linéarité, la largeur de la bande de fréquence, et la vitesse de balayage SR. Nous présentons donc un CNA semi-RZ dans lequel la gigue d'horloge ne survient qu'une fois par période comme dans le NRZ, l'amplitude de sortie étant très proche à celle de CNA NRZ (à 5% près par exemple), on n'observe pas d'impulsion transitoire (glitch) et d'erreur causée par la différence entre $t_{on} - t_{off}$. La cellule de CNA semi-RZ est semblable à celle de CNA NRZ à la différence près que chaque signal d'entrée doit obligatoirement mettre sa sortie à zéro puis activer si cela est nécessaire consécutivement. La figure 6.20 illustre le schéma de cellule proposée; le signal est appliqué sur les grilles des transistors commutateurs. La commande de remise à zéro qui dure une petite partie de la période (par exemple

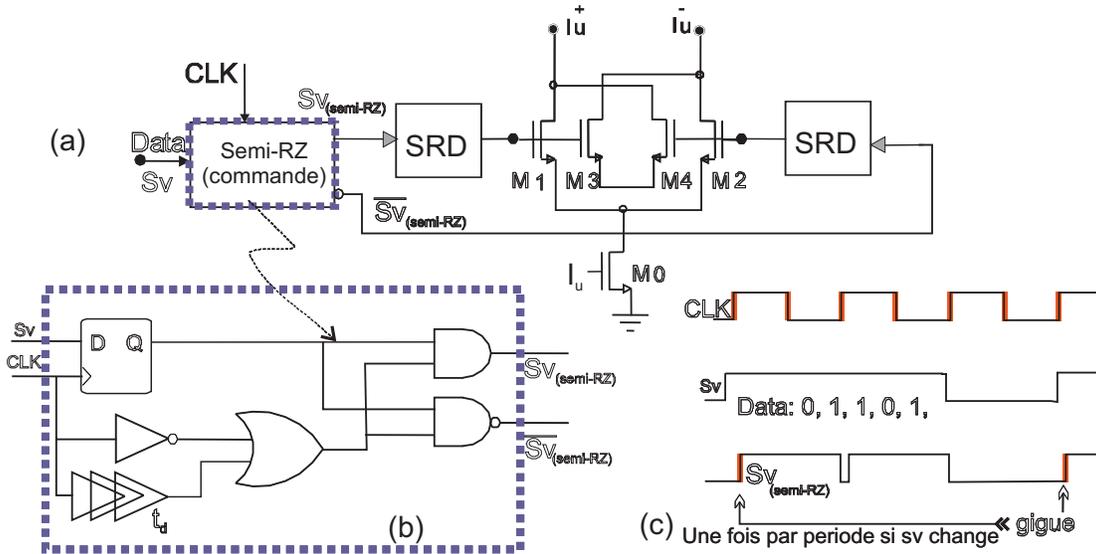


FIG. 6.20 – Cellule proposée pour un CNA semi-RZ a, b) circuit, c) les signaux

$0.02T_e < t_d < 0.05T_e$) est fournie par le même front d'horloge commandant l'application du signal d'entrée. Par conséquent, la gigue d'horloge ne peut survenir qu'une seule fois dans chaque période, ce qui est une caractéristique essentielle du CNA semi-RZ.

Afin de limiter l'effet de gigue d'horloge, le CNA semi-RZ est avantageux surtout dans le cas multibit de type passe-bas. Comme analysé au chapitre 3 pour le CNA composé de M cellules de type NRZ, l'effet de gigue d'horloge peut en général être de M fois inférieur à celle d'un CNA monobit (l'équation 3.78). Alors que le nombre des cellules importe peu pour un CNA RZ (l'équation 3.79), ce qui est la perte la plus importante quand on utilise un CNA RZ au lieu de son concurrent NRZ [4,32,108,109]. Cependant, le CNA multibit de type semi-RZ proposé ici se comporte vis à vis de la gigue d'horloge exactement de la même façon un CNA multibit NRZ, ce qui veut dire que l'effet de gigue sera

inversement proportionnel au nombre des cellules M . Analytiquement, l'erreur du charge sortant du CNA multibit dépend de la différence entre deux entrées consécutives affectée par le coefficient de gigue d'un front d'horloge, exprimée comme suit :

$$e_{j,semi-RZ}(n) = [v(n) - v(n-1)] \frac{\delta\tau}{(T_e - t_d)} \quad (6.3)$$

où $v(n)$ est l'entrée du CNA et $\delta\tau$ est l'erreur d'horloge à l'instant "n" (gigue). Si t_d est supposé petit devant T_e , alors, l'équation ci-dessus sera la même que celle du CNA NRZ donnée à l'équation 3.77. Comme le cas NRZ analysé au chapitre 3, la puissance d'erreur de gigue du CNA multibit semi-RZ peut être alors estimée par la relation suivante :

$$\begin{aligned} N_{j,semi-RZ}^2 &= var[e_{j,semi-RZ}] \times \frac{1}{OSR} = \sigma_{\Delta_v}^2 \times \frac{\sigma_\tau^2}{(T_e - t_d)^2} \times \frac{1}{OSR} \\ &= \frac{\sigma_\tau^2}{(T_e - t_d)^2} \times \frac{1}{M \times OSR} \simeq \frac{\sigma_\tau^2}{M \times OSR \times (T_e - t_d)^2} \end{aligned} \quad (6.4)$$

Remarque importante :

Il faut noter que cette relation, que ce soit avec un CNA NRZ ou avec un CNA semi-RZ, donne un résultat tout à fait raisonnable pour un modulateur passe-bas. Cependant, la précision est moindre pour le cas passe-bande, car la variation du signal d'entrée de CNA σ_{Δ_v} est habituellement supérieure au LSB. La figure 6.21 montre le signal de sortie du modulateur (l'entrée du CNA) dans les deux cas passe-bas et passe-bande. Cette observation n'est pas conforme à certaines hypothèses données aux références [48,109], parce que celles-ci s'appuient sur les résultats du cas passe-bas. Un tel écart montre que l'effet de gigue d'horloge d'un modulateur passe-bande doit être étudié plus précautionneusement. La variation du signal sortant ($\Delta_v = [v(n) - v(n-1)]$) d'un modulateur passe-bas est en général à l'ordre d'un bit, indépendamment de l'amplitude d'entrée dans les deux cas monobit et multibit. Le modulateur passe-bande monobit a naturellement le même comportement en sortie. Cependant, un modulateur passe-bande multibit varie en sortie de Δ_v qui dépend fortement de l'amplitude du signal d'entrée. Le Δ_v du modulateur passe-bande centré sur le quart de la fréquence d'échantillonnage est en général limité entre 1 LSB pour de petits signaux d'entrée et $\frac{M}{4}$ pour de grands signaux d'entrée, ($LSB \leq \Delta_v \leq \frac{M}{4}$). Une telle dépendance s'explique si on considère que les informations passe-bande sont toujours modulées sur une fréquence porteuse. Le signal de porteuse passe-bande doit balayer toute la plage d'amplitude du signal modulé une fois toutes les quatre périodes si sa fréquence vaut le quart de la fréquence d'échantillonnage. Afin d'estimer la variation de sortie d'un modulateur passe-bande pour une entrée arbitraire, on peut utiliser l'équation 2.10 qui suit :

$$V(z) = X(z) \times STF(z) + Q(z) \times NTF(z) \simeq X(z) + Q(z) \times NTF(z) \quad (6.5)$$

Pour une entrée sinusoïdale $x(t) = A \sin(\omega_{in}t)$, si A est supposée assez grande par rapport au bruit de quantification, on obtient :

$$\Delta_v(n) = [v(n) - v(n-1)] \simeq x(t) - x(t-nT) \simeq \frac{\partial x(t)}{\partial t} \Delta t = A\omega_{in} \cos(\omega_{in}t) \times \Delta t \quad (6.6)$$

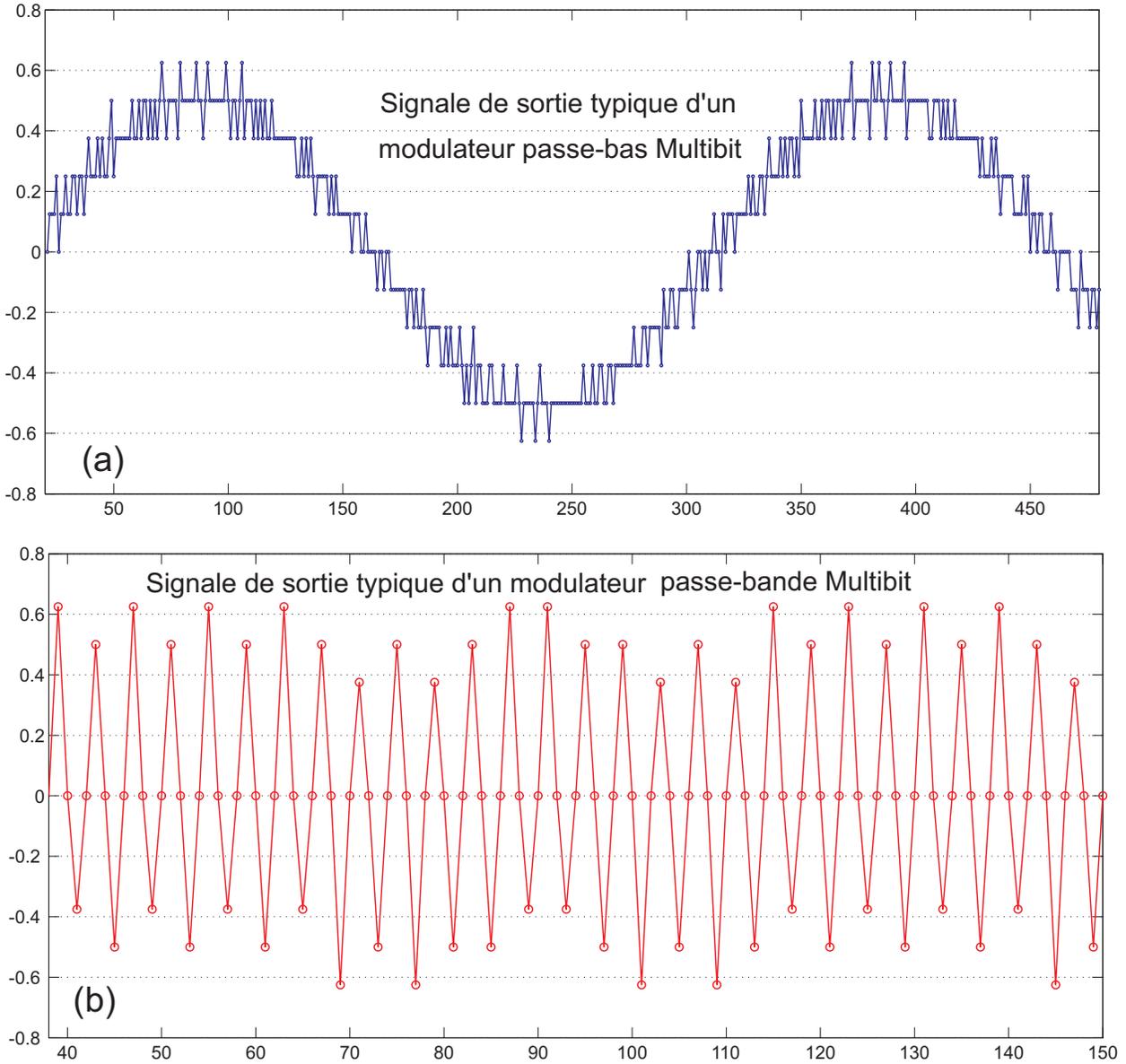


FIG. 6.21 – Signal de sortie du modulateur a) passe-bas b) passe-bande.

Cela se simplifie si la fréquence de l'entrée vaut le quart de la fréquence d'échantillonnage ($\omega_{in} = 2\pi \frac{f_e}{4}$) avec l'approximation suivante: $\Delta t \simeq T_e = \frac{1}{f_e}$. On obtient alors :

$$\Delta_{v,rms} \simeq \frac{A\pi}{2\sqrt{2}} \quad \text{pour une entrée grande } A > \text{LSB} \quad (6.7)$$

Notons que la valeur maximale A est en général limitée à $\frac{M}{2}$, donc la variation de sortie du modulateur passe-bande sera limitée en valeur inférieure à $\frac{M}{2}$ ce qui est confirmé par la simulation. Cette estimation n'est valable que pour le CNA NRZ ainsi que semi-RZ [3]. Cette estimation approximative montre que l'effet de gigue d'horloge d'un modulateur multibit passe-bande CT dépend de l'amplitude de l'entrée, ce qui est une caractéristique qui le distingue tout spécialement du modulateur passe-bas similaire.

Cependant, dans le cas d'une entrée faible ($x(t) < \frac{1}{M}$), le comportement d'un modulateur passe-bande est similaire à celui d'un passe-bas donc le signal de sortie est de l'ordre du LSB, ce qui met bien en évidence l'avantage d'utiliser un modulateur multibit afin de limiter l'effet de gigue d'horloge. Ce comportement semblable concernant l'effet de gigue d'horloge pour une entrée de faible amplitude dans deux cas passe-bas et passe-bande, est particulièrement intéressant car il détermine la plage dynamique du modulateur (DR). De toute façon, le fonctionnement dynamique et l'intérêt d'un CNA semi-RZ vis à vis de la gigue d'horloge est toujours très proche à celle d'un CNA NRZ, que ce soit dans le cas passe-bas que dans le cas passe-bande, alors qu'il ne présente pas de problème lié aux impulsions transitoires notamment du "glitch".

La seule question dans un CNA semi-RZ reste la détermination exacte de la durée de commande de remise à zéro (t_d). En général, imposer un retard en utilisant quelques portes numériques placées en série offre moins de précision que d'utiliser un front d'horloge. Le retard obtenu peut avoir une valeur approximative, par exemple, $\pm 5\%$ autour de sa valeur prévue. Avec l'outil de conception et de simulation actuel et la base de données fourni par les fabricant de circuits intégrées, l'éventuelle variation (erreur) de ce type de retard peut être assez limitée, par exemple inférieur à 1%. Un tel générateur de retard est souvent utilisé dans la réalisation des circuits à capacités commutés afin de préparer les signaux d'horloge multi-phase pour qu'ils ne se recouvrent pas. Pour les raisons suivantes, la précision finie de ce retard (Δt_d) ne peut avoir qu'une influence négligeable sur la performance d'un CNA semi-RZ :

1. Δt_d a une valeur petite par rapport à la valeur absolue de t_d , et t_d est aussi petit devant la période d'échantillonnage, donc, l'erreur finale sur la charge appliquée $I_u \times (T_e - t_d)$ sera très petit, par exemple de l'ordre de $0.01\% = 10^{-4}$,
2. Si une erreur de retard est appliquée à toutes les cellules du CNA en utilisant un circuit commun, son effet sera linéaire, donc, ne limitera pas du tout la linéarité de système et apparaîtra comme une erreur de gain linéaire $\bar{\alpha}$,
3. Si t_d n'est pas identique pour toutes les cellules, mais qu'il ne présente qu'une erreur modérée, l'influence de l'erreur Δt_d pourrait se traduire par une erreur statique de cellule laquelle sera mise en forme par la méthode de DEM.

6.4 Compensation des erreurs dynamiques du CAN multibit

Les erreurs d'un CAN sont analysées au chapitre 3. L'effet le plus important de non idéalité d'un CAN employé dans une boucle de modulateur $\Sigma\Delta$ est le temps d'établissement qui limite la vitesse de conversion analogique numérique. En effet, tous les temps de propagations générés par les différents blocs du modulateur $\Sigma\Delta$ peuvent être compensés au niveau système en ajoutant une boucle de rétroaction rapide, à l'exception du CAN. Ce dernier constitue une limite principale à l'augmentation de la fréquence d'échantillonnage. D'après plusieurs rapports analytiques et expérimentales, la vitesse du quantificateur est le premier facteur limitant à la fréquence maximale d'échantillonnage dans le modulateur $\Sigma\Delta$ à temps continu [24,33,34,48,51,109,157]. Pour une fréquence d'échantillonnage don-

née, deux effets de l'hystérésis et de la métastabilité de la boucle causées par le quantificateur doivent être correctement vérifiés. La compensation d'une éventuelle limite de vitesse par la quantification peut être réalisée tout d'abord par l'augmentation de la consommation afin d'obtenir un gain élevé du comparateur. Pour un CAN composé de comparateurs rapides, une telle méthode de compensation ne donne pas de résultats satisfaisants et elle exige une augmentation de consommation non acceptable en particulier pour des petites variations. Afin d'obtenir un CAN plus rapide, un deuxième moyen consiste à utiliser une architecture entrelacée. Dans le cas simple, cette méthode est appelée "Ping-Pong" si deux CAN sont utilisés. Cependant, une telle architecture souffre en pratique par défaut d'appariement entre les deux parties, notamment en terme d'erreur d'offset. Heureusement, toutes les erreurs du CAN qui affectent le modulateur en superposant au signal principal, comme l'erreur d'offset, sont mises en forme par la NTF au sein du modulateur $\Sigma\Delta$ donc posent moins de problèmes. Ainsi, nous proposons une nouvelle architecture du CAN de type RZ qui nous permet d'augmenter la fréquence d'échantillonnage du modulateur sans diminuer le temps d'établissement du CAN. À partir d'une certaine fréquence, la deuxième technique peut devenir plus économique que la première (en terme de consommation).

Le principe de l'architecture proposée est illustré à la figure 6.24. Il s'agit d'utiliser deux CAN-flash en alternance. Chaque CAN convertit le signal continu appliqué à son entrée sur le front montant de l'horloge. Une remise à zéro est alors nécessaire quand l'horloge passe à l'état bas. En plaçant à la sortie des deux CAN des portes *OR* de type drain ouvert (*Open-Drain*), la cadence du signal de sortie est alors deux fois plus grande sans réduction du temps d'établissement du CAN. Les CAN utilisent des comparateurs RZ identiques de type dégénératif avec une combinaison de rebouclage positif, comme sur la figure 6.22.

Les erreurs additionnelles entre des CAN entrelacés désappariés en terme d'offset, de retard ou autre, doivent être inférieures à l'erreur de quantification. Une prévision raisonnable consiste à limiter l'erreur de défaut pour qu'elle soit 3 fois plus petite que l'erreur de quantification soit d'atteindre environ 1.2dB de réduction en terme de SNDR final. Par exemple la puissance d'erreur d'appariement entre deux CAN doit être inférieure à $(\frac{1}{3\sqrt{12M}})^2$. Dans le cas RZ, le défaut d'appariement est en générale imposé par l'offset. Si les deux CNA entrelacés utilisent un diviseur de tension commun, l'offset entre deux CAN dépendra seulement de l'offset des comparateurs. Comme on l'a analysé au chapitre 3, dans un CAN, l'offset qui intervient dépend toujours de l'offset d'une seule paire différentielle et ne dépend pas de la somme des offsets des tous les comparateurs. Ainsi, l'offset du comparateur ne constituera pas une limite importante dans l'architecture ping-pong, si l'amplitude de l'offset ne varie pas avec le nombre des comparateurs. Parce que, dans le cas CNA l'offset de chaque bloc est la somme des offset de toutes les cellules, mais, dans le cas CAN dans chaque période de quantification il y a un seul niveau qui peut être influencé par l'erreur y compris l'offset, si bien que l'erreur reste inférieur à un LSB. La figure 6.23 montre la sortie du modulateur obtenu en utilisant un CNA d'architecture ping-pong avec une taux d'erreur supposée 10%, en comparaison au cas idéal. En appliquant le taux de défaut d'appariement entre deux CAN à l'ordre de 1%, la différence est invisible.

De l'expression 3.35, la variation minimale de l'amplitude d'un signal analogique appliqué à l'entrée

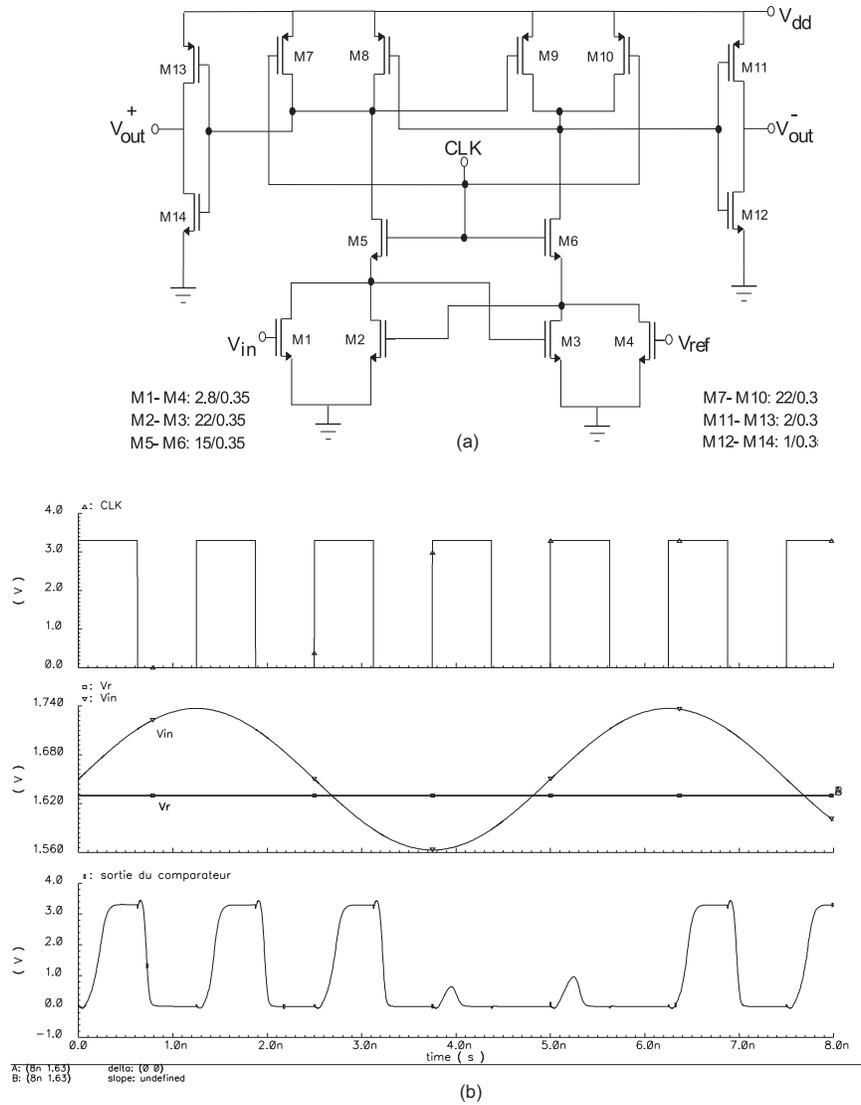


FIG. 6.22 – Comparateur RZ dégénératif avec un rebouclage positive : a) circuit, b) réponse transitoire

du comparateur, par l'utilisation de l'architecture entrelacée des deux CNA peut être exprimée par la relation suivante :

$$v_{min} = \Delta V_{logique} \cdot e^{-\frac{T_{com}}{\tau_l}} \quad (6.8)$$

où, $\Delta V_{logique}$ est environ deux tiers de la tension d'alimentation des portes numériques. T_{com} est le temps d'établissement d'une bonne comparaison, qui, avec une architecture ping-pong, peut durer une période complète d'échantillonnage, alors qu'elle ne dure que la moitié de ce temps si un seul CAN est employé. τ_l est la constante de temps du comparateur au moment de la comparaison. Cette tension minimale correspond à la résolution du comparateur ou à la valeur équivalant d'un LSB. L'erreur de métastabilité du comparateur peut être définie comme :

$$P_e = \frac{v_{min}}{V_{LSB}} = \frac{\Delta V_{logique}}{V_{LSB}} \cdot e^{-\frac{T_{com}}{\tau_l}} = 2^B \cdot e^{-\frac{T_{com}}{\tau_l}} \quad (6.9)$$

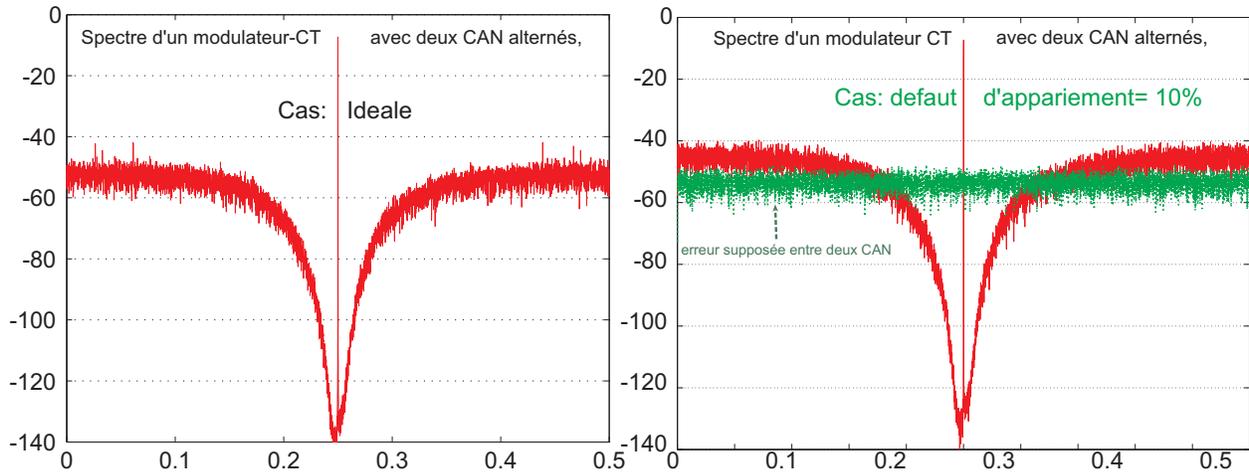


FIG. 6.23 – Spectre de sortie du modulateur passe-bande en utilisant un quantificateur d’une architecture ping-pong : a) cas idéal, b) avec 10 pourcent d’erreur

Par conséquent, la densité d’erreur de métastabilité pour un CNA B bits de structure ping-pong peut s’exprimer comme suit :

$$P_e = 2^B \cdot e^{-\frac{2\pi \times BW}{f_e}} \tag{6.10}$$

où, $f_e = \frac{1}{T_e}$ est la fréquence d’échantillonnage $BW = f_{-3dB,comparateur} = \frac{g_m}{2 \cdot \pi \cdot C_L}$ est la largeur de bande du comparateur pendant la phase de comparaison. De plus, si le comparateur suit un étage pré-amplificateur de gain A_p , la valeur de v_{min} (équation 6.8) ainsi que la probabilité d’erreur (équation 6.10) sont divisées par le gain A_p .

Pour conclure, l’erreur de métastabilité diminue d’une façon linéaire en augmentant le gain du comparateur, alors qu’elle est exponentiellement atténuée quand on augmente le temps de comparaison. Ainsi, l’utilisation de deux séries de comparateurs entrelacés peut repousser la limite de la fréquence d’échantillonnage jusqu’à la faire doubler. Compte tenu des limites de la technologie uti-

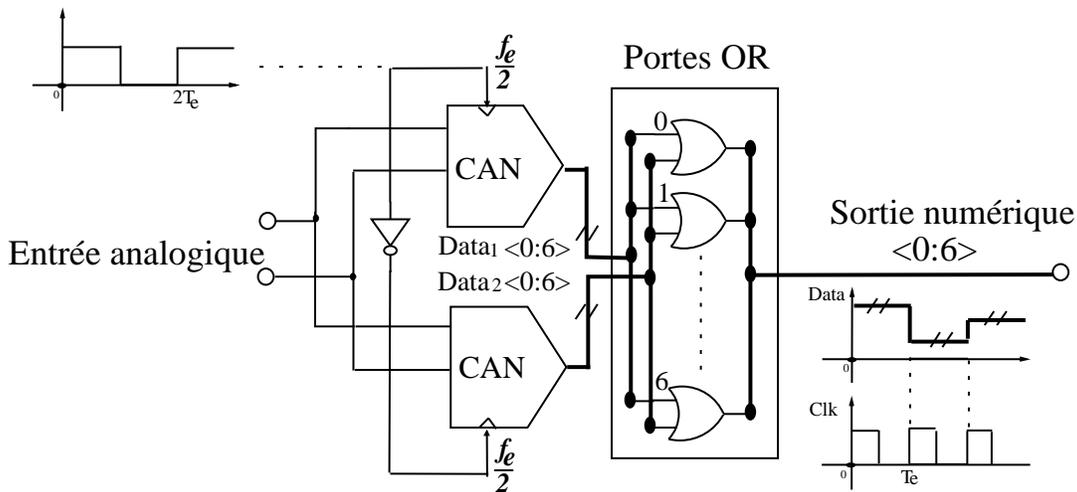


FIG. 6.24 – Deux CAN-flash en alternance.

lisée, cette proposition permet de concevoir un modulateur passe-bande à temps continu avec une fréquence d'échantillonnage égale à $1.2GHz$ ce qui permet de numériser un signal à une fréquence de $300MHz$ [24].

6.5 Conclusion

Nous avons présenté dans ce chapitre une conception à l'échelle du transistor d'un CNA multibit et une implantation des différentes méthodes de compensation des erreurs statiques et des erreurs dynamiques en technologie standard CMOS $0.35\mu m$.

Dans un premier temps, nous avons implanté les deux nouvelles méthodes d'appariement dynamique des composants (STDEM et MDEM) proposées au chapitre 5, ainsi qu'une méthode conventionnelle (SDEM), toutes avec des filtres identiques d'ordre 4 passe-bande. L'implantation des DEM est suivie du routage qui constitue l'étape finale; on vérifie la fréquence maximale que l'on peut appliquer ainsi que la surface nécessaire pour une technologie donnée. Leur comparaison met en évidence une très nette amélioration de la performance des méthodes proposées par rapport à la méthode conventionnelle. Les résultats des simulations confirment les analyses théoriques. La méthode STDEM est la plus recommandable puisqu'elle consomme moins de surface et peut être commutée à une fréquence allant au delà de $330MHz$. Il est clair que ces améliorations permettent de concevoir aujourd'hui d'un modulateur multibit avec une précision de plus de 16 bits sans être limité par la précision du CNA interne. Par ailleurs, obtenir une fréquence maximale de commutation aussi élevée grâce à une méthode DEM, peut être intéressant si on utilise un modulateur aux mêmes fréquence d'échantillonnage que les modulateurs discrets actuels et que certains modulateurs à temps continu.

Afin de compenser les erreurs dynamiques du CNA, nous avons présenté une conception au niveau transistor des trois différentes techniques, en particulier une nouvelle conception d'un bloc SRD. De plus, afin de limiter l'effet de gigue d'horloge en éliminant les impulsions transitoires du CNA surtout de glitch, nous avons présenté un CNA semi-RZ. Le fonctionnement dynamique du CNA semi-RZ est comparable à celui d'un CNA-RZ, alors que son comportement vis-à-vis la gigue d'horloge est aussi peu sensible que celui d'un CNA-NRZ. De plus, nous avons fait quelques remarques importantes sur le comportement de la gigue d'horloge dans le cas d'un modulateur passe-bande à temps continu qui diffère sensiblement de celui d'un modulateur passe-bas CT.

Puis, nous avons utilisé une nouvelle architecture de CAN qui a permis d'augmenter la fréquence d'échantillonnage (par exemple de $800MHz$ à $1.6GHz$). Notons que cette proposition se montrera utile si la fréquence d'échantillonnage est limitée par la vitesse maximale du quantificateur.

Chapitre 7

Conclusion générale et perspectives

L'objectif de ce travail était d'étudier les techniques d'amélioration de la linéarité des modulateurs $\Sigma\Delta$ multibit en technologie standard CMOS pour la numérisation des signaux rapides. Nous nous sommes en particulier intéressés à réduire l'effet de la non idéalité du CNA utilisé sur le chemin du retour du modulateur et nous avons mis en évidence les principales contraintes et les éventuelles sources d'erreurs liées aux circuits réalisant les cellules du CNA. L'étude analytique de ces sources a permis de mettre en évidence les différentes erreurs statiques et dynamiques présentes dans un modulateur multibit.

À l'exception du bruit thermique ($\frac{KT}{C}$) dans les circuits d'échantillonnage, les erreurs les plus importantes sont les suivantes :

- défaut d'appariement entre les cellules d'un CNA multibit,
- transitoires du signal de sortie du CNA (telles que le temps de montée et de descente, *glitch*, *CFT*) critiques essentiellement dans le cas des modulateurs à temps continu,
- effet de la gigue d'horloge (modulateurs CT),
- retard consacré au temps d'établissement du CAN.

En raison des limites de précision des technologies d'intégration actuelles, certaines erreurs dans les circuits analogiques sont inévitables; on cherche alors à minimiser leur effet grâce à des techniques de correction additionnelles.

Nous avons tout d'abord effectué une étude bibliographique afin de classer les méthodes existantes de correction ou d'élimination de ces erreurs, dans le cadre des modulateurs $\Sigma\Delta$. La méthode la plus efficace pour diminuer l'effet des défauts d'appariement du CNA, (sauf pour les systèmes à très faible *OSR*, $OSR < 8$), est la DEM. Nous avons examiné les différentes techniques de DEM existantes en s'appuyant sur leurs avantages et leurs inconvénients. Pour une DEM passe-bas d'ordre "1", la méthode DWA est la meilleure mais souffre des tons parasites; différentes solutions à ce problème ont déjà été présentées. Nous avons montré l'intérêt de chacune de ces solutions, et nous avons proposé deux nouvelles améliorations de la méthode DWA, l'une pour la DEM passe-bas d'ordre 1 et l'autre pour la DEM passe-bande d'ordre 2.

En revanche, seules les deux méthodes TDEM et SDEM permettent la mise en œuvre d'algorithmes

capables de fournir une mise en forme d'ordre plus élevé. La performance de TDEM est limitée par sa grande sensibilité aux caractéristiques stochastiques de l'entrée. Cette sensibilité entraîne très souvent l'instabilité de l'algorithme, de sorte qu'une TDEM d'ordre deux passe-bas (ou d'ordre quatre passe-bande) n'est en pratique pas possible. La SDEM est la méthode la plus robuste, mais devient très complexe du fait de l'augmentation exponentielle des ressources de calcul quand le nombre de bits du CNA augmente.

Au cours de ce travail, nous avons proposé deux nouveaux algorithmes de DEM qui sont très peu sensibles voir insensibles aux caractéristiques stochastiques de l'entrée. La première méthode, utilise partiellement l'algorithme de TDEM ce qui diminue la complexité des circuits et d'autre part, utilise localement l'algorithme SDEM afin de diminuer la sensibilité aux caractéristiques stochastiques de l'entrée. Cette proposition est un mélange des deux algorithmes existants et est notée *MDEM*. Nous avons fourni des équations mathématiques unifiées et validées par la simulation au niveau système et au niveau transistor. La deuxième proposition consiste en un algorithme efficace qui est basé sur une modification profonde de l'algorithme TDEM. Cette méthode, qui est notée *STDEM* et qui ressemble à une TDEM tronquée, est la méthode la plus économique en ressources, la plus rapide et elle peut s'appliquer à des applications DEM d'ordre deux et trois passe-bas ainsi que d'ordre quatre et six passe-bande. La performance de cet algorithme a été vérifiée de façon théorique ainsi que par des simulations au niveau système et au niveau transistor.

En ce qui concerne la compensation des erreurs dynamiques du CNA, nous avons proposé des circuits de *deglitcher* et *SRD* qui diminuent les impulsions transitoires d'un nouveau type de CNA semi-RZ, qui a été également étudié au cours de ce travail. Ce dernier développement marque un pas important dans la diminution de l'effet de la gigue d'horloge avec une quantification multibit, effet qui s'aggrave habituellement du fait de l'accroissement du nombre de transitions, dû aux méthodes de DEM.

Nous avons également mis en évidence la limite de la fréquence maximale d'échantillonnage d'un modulateur $\Sigma\Delta$ à temps continu, imposée par le temps d'établissement du comparateur (dans le *CAN flash*), alors que le retard des autres éléments de la boucle qui peut être compensé à l'étape de la conception au niveau système. Afin d'améliorer la vitesse d'échantillonnage d'un modulateur $\Sigma\Delta$ à temps continu limité par la vitesse du comparateur, nous avons proposé une nouvelle architecture de CAN. Cette architecture utilise une seule chaîne de résistances fournissant les références de comparaison, mais avec deux séries de comparateurs fonctionnant en alternance. Par conséquent, la fréquence d'échantillonnage peut être doublée. Nous avons montré que l'offset entre comparateurs n'a pas d'influence supplémentaire sur les performances de cette architecture.

En conclusion, ce travail constitue une contribution pour la réalisation de modulateurs $\Sigma\Delta$ à hautes performances, particulièrement pour les applications passe-bande à temps continu et en technologie standard *CMOS*. Il ouvre des perspectives dans le domaine de la conversion analogique-numérique à grande précision et très large bande, et s'intègre parfaitement dans les actions du Département des Signaux et Systèmes Électroniques de Supélec concernant les systèmes de numérisation du futur.

Dans la perspective de ce travail, nous pouvons proposer deux axes de recherches suivantes :

- adaptation des algorithmes proposés au cas de modulateurs large-bande, par le déplacement des zéros de la MTF (*Mismatch-Transfer-Fonction*),
- placement du bloc de DEM au niveau du CAN plutôt que sur le chemin du CNA afin d'éliminer son retard dans la boucle,
- étude à des fréquences autres que 0.25 pour les méthodes proposées au cas passe-bande.

```
*** **  
*** **  
***
```

Liste des publications

[51, 132, 149, 150, 152] :

[1] E.N. Aghdam , P. Benabes , "A new mixed stable DEM algorithm or bandpass multibit delta sigma ADC", Proceedings of the 10th IEEE International Conference on Electronics, Circuits and Systems, ICECS 2003. Volume 3, 14-17 Dec. 2003 Page(s): 962 - 965,
IEEE Digital Object Identifier: 10.1109/*ICECS*.2003.1301668

[2] S. Benabid, E. Najafi Aghdam, P. Benabes, S. Guessab, R. Kielbasa, "CMOS Design of a multibit bandpass continuous-time sigma delta modulator running at 1.2GHz," International Caracas Conference on Devices, Circuits & Systems (*ICCDCS*), Nov. 2004. Page(s): 51-55
IEEE Digital Object Identifier: 10.1109/*ICCDCS*.2004.1393352

[3] Esmaeil Najafi Aghdam, Philippe Benabes, "Higher order dynamic element matching by shortened tree-structure in Delta-Sigma modulators", ECCTD 2005 - European Conference on Circuit Theory and Design, Cork Ireland, 29 August - 2 September 2005,
IEEE Digital Object Identifier: 0 – 7803 – 9066 – 0/05/20.002005*IEEE*

[4] Esmaeil Najafi Aghdam, Philippe Benabes, "A hardware efficient 3-bit second-order dynamic element matching circuit clocked at 300MHz", IEEE International Symposium on Circuits and Systems (ISCAS 2006). Kos, Grace, 21-24 May 2006, Page: 2977-2980.
IEEE Digital Object Identifier: (not yet)

[5] E. Najafi Aghdam, P. Benabes, R.Kielbasa; "Stability and performance improvement in high order tree-structured dynamic element matching algorithms in delta-sigma modulators", Soumis à IEEE-Transaction on *C&S – I*, Control Number: 2126 (revision of 1462).

Bibliographie

- [1] R. Schreier and G. Temes, *Understanding Delta-sigma data converters*. IEEE Press, 2005.
- [2] S. Norsworthy, R. Schreier, and G. Temes, *Delta-sigma data converters, Theory, design and simulation*. NJ: IEEE Press, 97.
- [3] M. Gustavsson, J. Wikner, and N. Tan, *CMOS Data Converters for Communications*. Kluwer A.P., 2000.
- [4] J. Cherry and W. Snelgrove, *Continuous-Time Delta-Sigma Modulators for High-Speed A/d Conversion: Theory, Practice and Fundamental Performance Limits*. Kluwer A.P., 1999.
- [5] O. Bajdechi, J. Huijsing, and J. Huijsing, *Systematic Design Of Sigma-delta Analog-to-digital Converters*. Kluwer A.P., 2004.
- [6] R. V. D. Plassche, "Dynamic element matching for high- accuracy monolithic d/a converters," *IEEE C&S*, vol. 11, pp. 795–00, 76.
- [7] L. Carley, "A noise shaping coder topology for 15+ bit converters," *IEEE solid state circuit*, vol. SC-24, pp. 267–73, 04 89.
- [8] B. Leung and S. Sutarja, "Multi-bit a/d converter incorporating a novel class of dynamic element matching techniques," *IEEE circuit & systems II*, vol. CASII-39, pp. 35–51, 01 92.
- [9] F. Chen and B. Leung, "A high resolution multibit sigma-delta modulator with individual level averaging," *IEEE J.Solid-State Circuits*, vol. SC-30, pp. 453–60, 04 95.
- [10] R. Baird and T. Fiez, "Linearity enhancement of multi-bit a/d and d/a converters using data weighted averaging," *IEEE Circuits & Systems II*, vol. CASII-42, pp. 753–62, 12 95.
- [11] I. Galton, "Spectral shaping of circuit errors in digital-to-analog converters," *IEEE circuit and systems II*, vol. 44, pp. 807–8, 97.
- [12] R. Schreier and B. Zhang, "Noise-shaped multibit d/a convertor employing unit elements," *Electronic letters*, vol. 31/20, pp. 1712–3, 09 95.
- [13] L. Larson, T. Cataltepe, and G. Temes, "Multi-bit oversampled $\delta\sigma$ a/d converter with digital error correction," *Electronic Letters*, vol. 24, pp. 1051–52, Aug. 1988.
- [14] E. COLINET, J. JUILLARD, S. GUESSAB, and R. KIELBASA, "Actuation of resonant mems using short pulsed forces," *Proc. Sensors and Actuators A : Physical*, no. 115, pp. 118–125, 2004.
- [15] P. Benabes, *Etude de nouvelles structures de convertisseurs Sigma-Delta passe-bande*. PhD thesis, Dept.SSE, SUPELEC, Université Paris XI, 1994.
- [16] R. Adam, "Design aspects of high-order delta-sigma a/d converters," *Tutorial ISCAS'94*, pp. 235–260, 1994.
- [17] G. Ushaw and S. McLaughlin, "On the stability and configuration of sigma delta modulators," *Proceedings of ISCAS'94*, vol. 5, pp. 349–352, 1994.
- [18] S. Ardalan and J. Paulos, "An analysis of nonlinear behavior in delta-sigma modulators," *IEEE Trans. on Circuits and Systems*, vol. 34, pp. 593–603, June 1987.

- [19] R. Schreier and Y. Yang, "Stability tests for single-bit sigma-delta modulators with second-order fir noise transfer functions," *Proceedings of ISCAS'92*, pp. 1316–1319, May 1992.
- [20] R. Baird and T. Fiez, "Stability analysis of high-order delta-sigma modulation for adc's," *IEEE Trans. on Circuit and Systems*, vol. 41, pp. 59–62, Jan. 1994.
- [21] R. Schreier, "An empirical study of high order single-bit delta-sigma modulators," *IEEE Trans. on Circuits and Syst.-II*, vol. 40, pp. 429–39, Jul. 1993.
- [22] R. Schreier and M. Snelgrove, "Bandpass sigma-delta modulation," *Electronic letters*, vol. 25, pp. 1560–61, Nov. 89.
- [23] O. Shoaie and W. Snelgrove, "A multi-feedback design for *lc* bandpass delta-sigma modulators," *Proceedings Inter Symposium on Circuits and Systems*, vol. 1, pp. 171–174, May 1995.
- [24] S. Benabid, *Architecture et conception d'un modulateur sigma delta passe-bande à filtre LC adaptés à la numérisation des signaux rapides*. PhD thesis, Dept.SSE, SUPELEC, Université Paris XI, 2005.
- [25] L. Hernandez, P. Rombouts, E. Prefasi, M. Garcia, and C. Lopez, "A jitter insensitive continuous-time $\delta\sigma$ modulator using transmission lines," *Proc. of Conf. IEEE*, pp. 109–12, 04.
- [26] L. Hernandez, E. Prefasi, and P. Rombouts, "Sixth-order programmable bandwidth bandpass sigma-delta modulator implemented with transmission lines," *Elec. letters*, vol. 40, Dec. 04.
- [27] T. kaplan, J. jensen, C. fields, and M. chang, "A 2-gs/s 3-bit $\delta\sigma$ -modulated dac with tunable bandpass mismatch shaping," *IEEE J. of Solid-state Circuits*, vol. 40, pp. 603–10, 03 05.
- [28] A. Yahia, P. Benabes, and R. Kielbasa, "Bandpass delta-sigma modulators synthesis with high loop delay," *The 2001 IEEE International Symposium on Circuits and Systems*, vol. 1, pp. 344–7, May 01.
- [29] A. Yahya, *Architecture et conception d'un modulateur sigma delta passe-bande à filtre LC adaptés à la numérisation des signaux rapides*. PhD thesis, Dept.SSE, SUPELEC, Université Paris XI, Dec. 2002.
- [30] P. Benabes, M. Keramat, and R. Kielbasa, "A methodology for designing continuous-time sigma-delta modulators," *Proc. European Design and Test Conf.97*, pp. 46–50, 1997.
- [31] P. Benabes and J. Degouy, "article - fr," tech. rep., article - fr, France, 03 2002.
- [32] J. Cherry and W. Snelgrove, "Loop delay and jitter in continuous-time delta sigma modulators," *Proc. IEEE ISCAS98*, vol. 1, pp. 596–599, 1998.
- [33] J. Cherry and M. Snelgrove, "Excess loop delay in continuous-time delta-sigma modulator," *IEEE Trans. on Circuits and Syst-II*, vol. 46, pp. 376–89, April 1999.
- [34] W.Gao, O. Shoaie, and W. Snelgrove, "Excess loop delay effects in continuous-time delta-sigma modulators and the compensation solution," *Proc. IEEE ISCAS97*, vol. 1, pp. 65–68, 1997.
- [35] A. Yahia, P. Benabes, and R. Kielbasa, "Influence of the feedback dac delay on a continuous-time bandpass /spl delta//spl sigma/ converter," *IEEE International Symposium on Circuits and Systems*, vol. 2, pp. II-648–51, May 02.
- [36] A. Yahia, P. Benabes, and R. Kielbasa, "A new technique for compensating the influence of the feedback dac delay in continuous-time bandpass delta-sigma converters," *IMTC 2001*, vol. 2, pp. 716–9, May 01.
- [37] R. Adam, "Design and implementation of an audio 18-bit analog-to-digital converter using oversampling techniques," *J. Audio Eng. Society*, pp. 153–166, Mar. 1986.

- [38] M. Hauser, "Theory scaling and performance limitations in delta sigma analog to digital," *Proc. of ISCAS 90*, pp. 356–359, May 1990.
- [39] Y. Geerts and M. Steyaert, "Optimized topologies for multibit oversampling a/d converters," Tech. Rep. Deliverable D2.2.1, K. U. Leuven, ESAT - MICAS, K. U. Leuven, Belgique, 1999. ESPRIT project : SYSCONV Work Package 2.
- [40] F. Gerfers, M. Ortmanns, and Y. Manoli, "Design issues and performance limitations of a clock jitter insensitive multibit dac architecture for high-performance low-power $\Delta\Sigma$ modulators," *Proceedings of the ISCAS 2004*, vol. 1, pp. 1076–9, May 2004.
- [41] M. Ortmanns and F. Gerfers, *Continuous-Time Sigma-Delta A/D Conversion: Fundamentals, Performance Limits and Robust Implementations*. Springer, 2006.
- [42] L. Samid and N. Manoli, "The nonidealities of multibit continuous-time sigma delta modulators," *Proceedings of the ICECS 2003*, pp. 790–3, Dec. 03.
- [43] M. Ortmanns, F. Gerfers, and Y. Manoli, "Fundamental limits of jitter insensitivity in discrete and continuous-time dsm," *Proc. IEEE-CAS-Cof. ISCAS*, pp. 1037–40, 05 03.
- [44] M. Ortmanns, F. Gerfers, and Y. Manoli, "A continuous-time $\delta\sigma$ modulator with reduced sensitivity to clock jitter through scr feedback," *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS*, vol. 52, pp. 875–84, 05 05.
- [45] P. Benabes, P. Aldebert, and R. Kielbasa, "A matlab based tool for bandpass continuous-time modulators design," *Proc. IEEE Int. Symp. Circuits and Syst.*, pp. 274–277, June 1998.
- [46] Mathwork, "Matlab's user guide," <http://www.mathworks.com/matlabcentral/fileexchange>, 2005.
- [47] S. Paton, A. Gianomenico, L. Hernandez, A. Wiesbauer, T. Potscher, and M. Clara, "A 12 bit continuous-time $\delta\sigma$ modulator with 400mhz clock and low jitter sensitivity in 0.13um cmos," *Proc. of Symposium on VLSI, IEEE*, pp. 82–3, 04.
- [48] S. Paton, A. Gianomenico, L. Hernandez, A. Wiesbauer, T. Potscher, and M. Clara, "A 70-mw 300-mhz cmos continuous-time $\delta\sigma$ modulator with 15-mhz bandwidth and 11 bits of resolution," *IEEE J. of Solid-state Circuits*, vol. 39, no. 7, pp. 1056–63, 04.
- [49] C. Lelandais-Perrault, P. Benabes, J.-L. Gouy, and R. Kielbasa, "A parallel structure of a continuous-time filter for band-pass sigma-delta a/d converters," *10th IEEE Int. Conf. on Elec. Circuits and Systems*, vol. 1, pp. 954–57, Dec. 2003.
- [50] S. Benabid and P. Benabes, "High-linear integrated lc filter for a continuous-time band-pass sigma-delta adc," *Proc. 46th IEEE Midwest Symposium on Circuits and Systems*, Dec. 03.
- [51] S. Benabid, E. N. Aghdam, P. Benabes, S. Guessab, and R. Kielbasa, "Cmos design of a multibit bandpass continuous-time sigma delta modulator running at 1.2 ghz," *Proc. of ICCDCS.2004*, vol. 1, pp. 51–5, Nov. 04.
- [52] P. Benabes, A. Gautier, and D. Billet, "New wideband sigma-delta convertor," *Electronic letters*, vol. 29, pp. 1575–1577, Août 1993.
- [53] P. Benabes, A. Gauthier, and R. Kielbasa, "A multistage closed-loop sigma-delta modulator (mscl)," *Analog Integrated Circuits and Signal Processing*, pp. 195–204, Nov. 1996.
- [54] C. LELANDAIS-PERRAULT, *Systèmes de numérisation hautes performances - Étude des solutions à bancs de filtres hybrides - Extension des fonctionnalités*. PhD thesis, Dept.SSE, SUPELEC, Université Paris XI, 2005.
- [55] T. PETRESCU, *Systèmes de numérisation hautes performances - Étude des solutions à bancs de filtres hybrides - Caractérisation et optimisation*. PhD thesis, Dept.SSE, SUPELEC, Université Paris XI, 2005.

- [56] E. King and al., “A nyquist-rate delta sigma a/d converter,” *IEEE J. SOLID-STATE CIRCUITS*, vol. 33, Jan. 1998.
- [57] S. Kong and W. Ku, “Effects of non-ideal hadamard modulators on the performance of $\delta\sigma$ adc,” *Electronics Letters*, vol. 33, no. 2, pp. 109–110, 1997.
- [58] I. Mehr and D. Welland, “A cmos continuous-time gm-c filter for prml read channel applications at 150mv/s and beyond,” *IEEE Journal and Solid-State Circuits*, vol. 32, pp. 499–513, Apr. 1997.
- [59] T. kaplan, J. jensen, and C. fields, “Continuous-time $\delta\sigma$ modulators using distributed resonators,” *IEEE Trans. on Circuits and Syst.-I*, vol. 52, pp. 2397–2403, 11 05.
- [60] H. Inose, Y. Yasuda, and J. Murakami, “A telemetring system by code modulation- $\delta\sigma$ modulation,” *IRE Trans. Space Electron. Telemetry*, vol. 8, pp. 204–9, Sept. 1962.
- [61] G. Ritchie, J. Candy, and W. Ninke, “Interpolative digital to analog converters,” *IEEE Trans on Communications*, vol. 22, pp. 1797–1806, Nov. 1974.
- [62] J. Candy, “A use of double integration in sigma delta modulation,” *IEEE Trans. on Communication*, vol. 33, no. 3, pp. 249–58, 1985.
- [63] J. Candy and A. Huynh, “Double integration for digital-to-analog conversion,” *IEEE Trans. on Communication*, vol. 34, no. 1, pp. 77–81, 1986.
- [64] T. Hayashi, Y. Inabe, K. Uchimura, and A. Iwata, “A multistage delta-sigma modulator without double integration loop,” *ISSCC Digest of Tehnical Papers*, pp. 182–183, Feb. 1986.
- [65] R. Carley and J. Kenney, “A 16-bit 4th order noise-shaping d/a converter,” *IEEE Proceed of the Custom Integrated Circuits Conference*, pp. 2171–2174, 1988.
- [66] R. Adams and T. Kwan, “Data-directed scrambler for multibit noise-shapping d/a converters,” *U.S. patent number 5404142*, Apr. 95.
- [67] I. Galton, “Noise-shapping d/a converters for $\delta\sigma$ modulation,” *IEEE Trans. on Circuits and Syst-II*, vol. 1, pp. 441–44, May 1996.
- [68] T. shui, R. Schreier, and F. Hudson, “Modified mismatch-shaping for continuous-time delta-sigma modulators,” *IEEE 1998 Custom integrated circuits conf. p.*, vol. 1, pp. 225–28, 98.
- [69] M. Story, “Digital to analogue converter adapted to select input sources based on a preselected algorithm once per cycle of of a sampling signal,” *U.S. patent number 5138317*, Agu 92.
- [70] A. Thurston, T. Pearce, and M. Hawksford, “Bandpass implementation of the sigma-delta a-d conversion technique,” *Proc. IEE Inter. Conf. on Analogue-to-Digital and Digital-to-Analogue Conversion*, pp. 81–86, Sept. 1991.
- [71] G. Troster, H. Drebler, H. Golberg, W. Schardein, E. Zocher, A. Wedel, K. Schoppe, and J. Arndt, “An interpolative bandpass converter on a 1.2- μm bicmos analog/digital array,” *IEEE Journal Solid-State Circuits*, vol. 28, pp. 471–6, Apr. 1993.
- [72] R. S. et. al, “A 50 mw bandpass $\delta\sigma$ adc with 333 khz bw and 90db dr,” *IEEE Journal Solid-State Conference Digest of Technical Papers*, vol. 28, pp. 216–217, Feb. 2002.
- [73] S. Yan and E. Sánchez-Sinencio, “A continuous-time $\sigma\delta$ modulator with 88-db dynamic range and 1.1-mhz signal bandwidth,” *Solid-State Circuits, IEEE Journal of*, vol. 39, pp. 75–86, Jan. 2004.
- [74] M. Kappes, “A 2.2-mw cmos bandpass continuous-time multibit $\delta\text{-}\sigma$ adc with 68db of dynamic range and 1-mhz bandwidth for wireless applications,” *IEEE Journal and Solid-State Circuits*, vol. 38, pp. 1098–1104, July 2003.

- [75] Y. Geerts, M. Steyaert, and W. Sansen, “A high-performance multibit $\delta\sigma$ cmos adc,” *IEEE J. SOLID-STATE CIRCUITS*, vol. 35, pp. 1829–40, Dec. 2000.
- [76] K. Vleugels, S. Rabi, and B. Wooley, “A 2.5-v sigma-delta modulator for broadband communication applications,” *IEEE J. Solid-State Circuits*, pp. 1887–99, Dec. 01.
- [77] J. Silva, X. Wang, P. Kiss, U. Moon, and G. Temes, “Digital techniques for improved $\delta\sigma$ data conversion,” *IEEE custom integrated circuits conference 2002*, pp. 183–190, 02.
- [78] R. Katoh, S. Kobayashi, and T. Waho, “A dynamic element matching circuit for multi-bit delta-sigma modulators,” *proc. IEEE Int. conf.*, vol. 6c, pp. 24–5, 04.
- [79] A. Yasuda and et.al., “A third-order modulator using second-order noise-shaping dynamic element matching,” *IEEE Journal of Solid-State Circuits*, vol. 33, pp. 1879–86, Dec. 98.
- [80] X. Gong, E. Gaalaas, M. Alexander, D. Hester, E. Walburger, and J. Bian, “A 120 db multibit sc audio dac with second-order noise shaping,” *IEEE Solid-State Circuits Conf. ISSCC*, pp. 345–9, 02 00.
- [81] E. Fogleman and I. Galton, “A dynamic element matching technique for reduced-distortion multibit quantization in delta-sigma adcs,” *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 48, pp. 158–70, Feb. 01.
- [82] L. Haiqing and R. Schreier, “A bandpass mismatch-shaped multi-bit sigma-delta switched-capacitor dac using butterfly shuffler,” *ISSCC*.
- [83] T. Ueno and A. Yasuda, “A fourth-order bandpass-modulator using second-order bandpass noise-shaping dynamic element matching,” *IEEE Journal of Solid-State Circuits*, vol. 37, pp. 809–16, July 02.
- [84] I. Fujimori, A. Nogi, and T. Sugimoto, “A multibit delta-sigma audio dac with 120-db dynamic range,” *Solid-State Circuits, IEEE Journal of*, vol. 35, pp. 1066–73, Aug. 00.
- [85] A. Prasad, A. Chkhawala, K. Thompson, and J. Melanson, “A 120db 300mw stereo audio a/d converter with 110db thd+n,” *IEEE Conf. Proc.*, pp. 191–4, 04.
- [86] L. Breems, E. van der Zwan, and J. Huijsing, “A 1.8-mw cmos $\delta\sigma$ modulator with integrated mixer for a/d conversion of *if* signals,” *IEEE Journal and Solid-State Circuits*, vol. 35, pp. 468–75, Apr. 2000.
- [87] E. Dagher, P. Stubberud, W. Masenten, M. Conta, and T. Dinh, “A 2-ghz analog-to-digital delta-sigma modulator for cdma receivers with 79-db signal-to-noise ratio in 1.23-mhz bandwidth,” *IEEE J. of Solid-State Circuits*, vol. 39, pp. 1819–28, Nov. 04.
- [88] M. Vadipour, “Techniques for preventing tonal behavior of data weighted averaging algorithm in $\sigma - \delta$ modulator,” *Circuits and Systems II*, vol. 47, pp. 1137–44, Nov. 00.
- [89] O. Nys and R. Henderson, “An analysis of dynamic element matching techniques in sigma-delta modulation,” *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 231–4, 05 96.
- [90] R. Henderson and O. Nys, “Dynamic element matching techniques with arbitrary noise shaping function,” *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 239–6, 05 96.
- [91] R. J. Baker, *CMOS circuit design, layout, and simulation*. IEEE Press, 2005.
- [92] Y. Yang, A. Chokhawala, and et.al., “A 114-db 68-mw chopper-stabilized stereo multibit audio adc in 5.62 mm/sup 2/,” *Solid-State Circuits, IEEE Journal of*, vol. 38, pp. 2061–8, Dec. 2003.
- [93] R. Schreier, J. Silva, J. Steensgaard, and G. Temes, “Design-oriented estimation of thermal noise in switched-capacitor circuits,” *IEEE Trans. on Circuits and Syst.-I*, vol. 52, pp. 2358–68, Nov. 2005.

- [94] C. Shannon, "Communications in the presence of noise," *Proc. IRE.*, vol. 37, pp. 10–21, Jan. 1949.
- [95] E. Hegazi and N. Klemmer, "Accurate modeling of noise in switched-c /spl delta//spl sigma/ analog-to-digital converters," *IEEE Trans. on Circuits and Syst.-I*, vol. 52, pp. 2319–26, Nov. 2005.
- [96] V. NGUYEN, P. LOUMEAU, and J. NAVINER, "Modélisation et simulation d'un modulateur sigma delta passe haut avec vhdl-ams," *inproceedings-2003*, 2003.
- [97] H. Aboushady, *Conception en vue de la réutilisation de convertisseur analogique-numérique $\Delta\Sigma$ temps-continu mode courant*. PhD thesis, Dept. ELECTRONICS, COMMUNICATIONS AND COMPUTER SCIENCE, UNIVERSITY OF PARIS VI, 2002.
- [98] S. Kuboki, K. Kato, N. Miyakawa, and K. Matsubara, "Nonlinearity analysis of resistor string a/d converter," *IEEE Trans. on Circuits and Systems*, vol. 29, Juin 1982.
- [99] P. Benabes, P. Aldebert, and A. Yahia, "Influence of the feedback dac delay on continuous-time bandpass sigma-delta coverter," *Electronics Letters*, vol. 36, pp. 292–4, Feb. 00.
- [100] J. Fattaruso, S. Kiriaki, and et.al., "Self calibration techniques for a second order multibit sigma-delta modulator," *IEEE Solid-State Circuits*, vol. SC-28, pp. 1216–23, 12 93.
- [101] O. Shoaie and W. Snelgrove, "Design and implementation of a tunable 40mhz-70mhz gm-c bandpass $\delta\sigma$ modulator," *IEEE Trans. on Circuits and Syst-II*, vol. 44, pp. 521–530, 07 1997.
- [102] J. Wikner and N. Tan, "Modeling of cmos digital-to-analog converters for telecommunication," *Trans. on Circuits and Syst-II*, vol. 46, no. 5, pp. 489–99, 1999.
- [103] A. Bugeja, B. S. Song, P. Rakers, and S. Gillig, "A 14-b, 100-ms/s cmos dac designed for spectral performance," *IEEE J. SOLID-STATE CIRCUITS*, vol. 34, pp. 1719–32, Dec. 1999.
- [104] A. Bosch, M. Borremans, M. Steyaert, and W. Sansen, "A 10-bit 1-gsample/s nyquist current-steering cmos d/a converter," *IEEE J. SOLID-STATE CIRCUITS*, vol. 36, pp. 315–324, Mar. 2001.
- [105] A. Bosch, M. Steyaert, and W. Sansen, "An accurate statisticales yield model for cmos current-steering d/a converters," *Analog Integrated Circuits and Signal processing*, no. 29, pp. 173–180, 2001.
- [106] M. Vesterbacka, M. Rudberg, J. Wikner, and N. Andersson, "Dynamic element matching in d/a converters with restricted scrambling," *IEEE Proc. The 7th Int. Conf. CAS, ICECS*, pp. 36–9, 12 00.
- [107] P. Benabes, M. Keramat, and R. Kielbasa, "A methodology for designing continuous-time sigma-delta modulators," *European Design and Test Conference*, pp. 46–50, March 97.
- [108] H. SHAMSI, O. SHOAEI, and R. DOOST, "Analysis of the clock jitter effects in a time invariant model of continuous time delta sigma modulators," *IEICE TRANS. FUNDAMENTALS*, vol. E89-A, pp. 399–407, Feb. 2006.
- [109] J. Cherry and M. Snelgrove, "Clock jitter and quantizer metastability in continuous-time delta-sigma modulators," *IEEE Trans. on Circuits and Syst-II*, vol. 46, pp. 661–76, 07 1999.
- [110] O. Oliaei and H. Aboushady, "Jitter effects in continuous time $\delta\sigma$ modulators with delayed return-to-zero feedback," *Proc. Int. Conf Electronics Circuits and Systemes*, pp. 351–4, Sept. 1998.
- [111] O. Oliaei, "Clock jitter noise spectra in continuous-time delta-sigma modulators," *IEEE Conf. proc.*, vol. II, pp. II–192–5, 1999.

- [112] S. Luschas and H. Lee, "High-speed $\delta\sigma$ modulator with reduced timing jitter sensitivity," *IEEE Trans. on CaS-II*, vol. 49, pp. 712–20, Nov. 02.
- [113] A. Pulincherry, M. Hufford, E. Naviasky, and U. Moon, "A time-delay jitter-insensitive continuous-time banpass $\delta\sigma$ modulateur architecture," *IEEE Trans. Circuits Syst.II*, vol. 52, pp. 680–684, Oct. 2005.
- [114] L. Hernandez and A. G. A. Wiesbauer, S. Patón, "Modelling and optimization of low pass continuous-time sigma-delta modulators for clock jitter noise reduction," *Proc. of ISCAS2004*, vol. I, pp. I-1072–5, 04.
- [115] O. Oliaei, "Sigma-delta modulators with spectrally shaped feedback," *IEEE Trans. on Circuits and Syst.-II*, vol. 50, pp. 518–530, Sept. 2003.
- [116] Y. Sakina, "Multi-bit $\sigma\delta$ analog-to-digital converters with nonlinearity correction using dynamic barrel shifting," *University of California, Berkely CA, Repport No. UCB/ERL M93/63, 1993*.
- [117] K. Klaassen, "Digitally controlled absolute voltage division," *IEEE trans. Insrum. Measur.*, vol. 24, pp. 106–112, June 19975.
- [118] D. Maeyer, P. Rombouts, and L. Weyten, "Addressing static and dynamic errors in unit element multibit dacs," *Elect. lett.*, vol. 39, pp. 1038–9, july 03.
- [119] Z. Li and T. Fiez, "Dynamic element matching in low oversampling delta sigma adcs," *IEEE Proc.*, pp. IV-683–686, 02.
- [120] A. E. R.E. Radke and T. Fiez, "A 14-bit current-mode sigmadelta dac based upon rotated data weighted averaging," *IEEE Journal of Solid-State Circuits*, vol. 35, pp. 1074–84, 08 2000.
- [121] D. Cini, C. Samori, and A. Lacaita, "Double-index averaging: a novel technique for dynamic element matching in σ - δ a/d converters," *IEEE Trans. on Circuit and Systems-II*, vol. 46, pp. 353–58, 04 1999.
- [122] K. Chen and T. Kuo, "An improved technique for reducing baseband tones in sigmaÜdelta modulators employing data weighted averaging algorithm without adding dither," *IEEE Circuits & Systems II*, vol. 46, pp. 63–68, Jan. 99.
- [123] T. Kuo, K. Chen, and H. Yeng, "A wideband cmos sigmaÜdelta modulator with incremental data weighted averaging," *IEEE J.Solid-State Circuits*, vol. 37, pp. 11–17, Jan. 02.
- [124] I. Fujimori, L. Longo, A. Hairapetian, K. Seiyama, S. Kotic, J. Cao, and S. Chan, "A 90-db snr 2.5-mhz output-rate adc using cascaded multibit deltaÜsigma modulation at 8 oversampling ratio," *IEEE J.Solid-State Circuits*, vol. 35, pp. 1820–28, Dec. 00.
- [125] J. Welz and I. Galton, "Necessary and sufficient conditions for mismatch shaping in a general class of multibit dacs," *IEEE Trans. on Circuits and Systems II*, vol. 49, pp. 748–9, Dec. 02.
- [126] A. Hamoui and K. Martin, "Linearity enhancement of multibit /spl delta//spl sigma/ modulators using pseudo data-weighted averaging," *Circuits and Systems, 2002. ISCAS 2002. IEEE International Symposium on*, vol. 3, pp. III-285–8, May 02.
- [127] H. Lin, J. Silva, B. Zhang, and R. Schreier, "Multi-bit dac with noise-shaped element mismatch," *Circuits and Systems, 1996. ISCAS '96., 'Connecting the World'. 1996 IEEE International Symposium on*, vol. 1, pp. 235–8, May 96.
- [128] A. Yasuda, H. Tanimoto, and T. Lida, "A 100 khz 9.6 mw multi bits dac and adc using noise shaping dynamic elements matching with tree structure," *Solid-State Circuits Conference, 1998. Digest of Technical Papers. 45th ISSCC 1998 IEEE International*, pp. 64–5, 98.
- [129] T. Shui, R. Schreier, and F. Hudsonn, "Mismatch-shaping dac for lowpass and bandpass multi-bit delta-sigma modulators," *IEEE ISCAS-98*, 05 98.

- [130] T. shui, R. schreier, and F. hudson, "Mismatch shaping for a current-mode multibit delta-sigma dac," *IEEE J. of Solid-state Circuits*, vol. 34, pp. 331–8, 03 99.
- [131] T. Shui, R. Schreier, and F. Hudson, "Mismatch shaping for a current-mode multibit delta-sigma dac," *Solid-State Circuits, IEEE Journal of*, vol. 34, pp. 331–8, March 99.
- [132] E. N. Aghdam and P. Benabes, "A new mixed stable dem algorithm for bandpass multibit delta sigma adc," *Proc. ICECS2003*, vol. 3, pp. 962–5, Dec. 03.
- [133] E. Fogleman, J. Welz, and I. Galton, "An audio adc delta-sigma modulator with 100-db peak sinad and 102-db dr using a second-order mismatch-shaping dac," *IEEE J. Solid-State Circuits*, vol. 36, pp. 339–48, 03 01.
- [134] J. Welz, I. Galton, and E. Fogleman, "Simplified logic for first-order and second-order mismatch-shaping digital-to-analog converters," *IEEE Trans. on Circuit and Systems-II*, vol. 48, pp. 1014–27, Nov. 01.
- [135] J. Jensen and I. Galton, "A reduced-complexity mismatch-shaping dac for delta-sigma data converters," *Proc. of ISCAS'98*, pp. 504–7, 05 98.
- [136] J. Welz and I. Galton, "The mismatch-noise psd from a tree-structured dac in a second-order modulator with a midscale input," *IEEE-conf. ICASSP*, pp. 2625–8, 05 01.
- [137] J. Welz and I. Galton, "A tight signal-band power bound on mismatch noise in a mismatch-shaping digital-to-analog converter," *IEEE trans. on information theory*, vol. 50, Apr. 04.
- [138] A. Fishov, E. Siragusa, J. Welz, E. Fogleman, and I. Galton, "Segmented mismatch-shaping d/a conversion," *Circuits and Systems, 2002. ISCAS 2002. IEEE International Symposium on*, vol. 4, pp. 679–682, May 02.
- [139] E. Fogleman, I. Galton, W. Huff, and H. Jensen, "A 3.3 v single-poly cmos audio adc delta-sigma modulator with 98 db peak sinad," *Custom Integrated Circuits, 1999. Proceedings of the IEEE 1999*, pp. 121–4, May 99.
- [140] E. fogleman, I. galton, W. huff, and H. jensen, "A 3.3-v single-poly cmos audio adc delta-sigma modulator with 98-db peak sinad and 105-db peak sfdr," *IEEE J. of Solid-state Circuits*, vol. 35, pp. 297–307, 03 00.
- [141] P. Benabes and J. Degouy, "A third order wideband dynamic element matching algorithm for bandpass converters," tech. rep., Services des Mesures, SUPELEC, Gif sur Yvette, France, 03 2002.
- [142] T. Leslie and B. Singh, "An improved sigma-delta modulator architecture," *Proc. IEEE Int. Symp. Circuits and Systems*, pp. 372–75, May 90.
- [143] J. Steensgaard, U. Moon, and G. C. Temes, "Mismatch-shaping serial digital-to-analog converter," *Proc. IEEE Int. Symp. Circuits and Systems ISCAS99*, vol. II, pp. II–5–8, 1999.
- [144] M. Sarhang-Nejad and G. Temes, "A high-resolution multibit ea adc with digital correction and relaxed amplifier requirements," *IEEE JOURNAL OF SOLID-STATE CIRCUITS*, vol. 28, no. 6, pp. 648–60, 93.
- [145] A. Hamoui and K. Martin, "High-order multibit modulators and pseudo data-weighted-averaging in low-oversampling $\delta\sigma$ adcs for braod-band application," *IEEE Trans. on CaS-II*, vol. 51, pp. 72–85, Jan. 04.
- [146] H. Lin and R. Schreier, "A bandpass mismatch-shaped multi-bit /spl sigma//spl delta/ switched-capacitor dac using butterfly shuffler," *ISSCC. 1999*, pp. 58–9, Feb. 99.
- [147] M. Vadipour, "A bandpass mismatch noise-shaping technique for /spl sigma/-/spl delta/ modulators," *Circuits and Systems II*, vol. 3, pp. 130–5, March 04.

- [148] J. Riches and N. Erdol, "Mismatch cancellation in quadrature bandpass ds modulators using an error shaping technique," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 49, pp. 73–85, Feb. 02.
- [149] E. N. Aghdam and P. Benabes, "Higher order dynamic element matching by shortened tree-structure in delta-sigma modulators," *P. ECCTD'05*, vol. I, pp. 201–4, Sept 2005.
- [150] E. N. Aghdam and P. Benabes, "A hardware efficient 3-bit second-order dynamic element matching circuit clocked at 300mhz," *P. ISCAS'06*, pp. 2977–2980, May 2006. lecture.
- [151] L. Zhimin and T. Fiez, "Dynamic element matching in low oversampling delta sigma adcs," *Circuits and Systems, 2002. ISCAS 2002. IEEE International Symposium on*, vol. 4, pp. IV–683–6, May 02.
- [152] E. N. Aghdam, P. Benabes, and R. Kielbasa, "Stability and performance improvement in high order tree-structured dynamic element matching algorithms in delta-sigma modulators," *Soumis à IEEE-Transaction on C&S – I*, pp. 1–10, 2006.
- [153] E. N. Aghdam and P. Benabes, "High performance $\delta\sigma$ modulateur for bandpass applications," technical report (in english), pages: 1-60, Dept. SSE, SUPELEC, Dec.2004.
- [154] L. Luh, J. Choma, and J. Draper, "A 50mhz continuous-time switched-current $\delta\sigma$ modulator," *Proc. of ISCAS 98*, 05 1998.
- [155] L. Luh, J. Choma, and J. Draper, "A high-speed fully differential current switch," *proc. icccs98*, 1998.
- [156] J. Pirkkalanemi, M. Waltari, M. Kosunen, L. Sumanen, and K. Halonen, "A 14-bit current-steering dac with current-mode deglitcher," *Analog Integr. Circuits Signal Process.*, vol. 35, no. 1, pp. 33–45, 2003.
- [157] W. Gao and W. Snelgrove, "A 950-mhz if second-order integrated lc bandpass delta-sigma modulator," *IEEE Journal and Solid-State Circuits*, vol. 33, pp. 723–732, May 1998.
- [158] F. Maloberti, *Analog Design for CMOS VLSI Systems*. Kluwer A.P., 2001.
- [159] T. Shui, R. Schreier, and F. Hudsonn, "Mismatch-shaping dac for lowpass and bandpass multi-bit delta-sigma modulators," *IEEE ISCAS-98*, 05 98.
- [160] X. Wang and et.al., "Digital estimation and correction of dac errors in multibit $\delta\sigma$ adc," *Electronics letters*, vol. 37, pp. 414–5, March 01.
- [161] C. Wang, "A 20-bit 25-khz delta-sigma a/d converter utilizing a frequency-shaped chopper stabilization scheme," *Solid-State Circuits, IEEE Journal of*, vol. 36, pp. 566–9, March 01.
- [162] R. Adams and K. Nguyen, "A 113-db snr oversampling dac with segmented noise-shaped scrambling," *Solid-State Circuits, IEEE Journal of*, vol. 33, pp. 1871–8, Dec. 98.
- [163] X. Gong, E. Gaalaas, M. Alexander, D. Hester, E. Walburger, and J. Bian, "A 120db multi-bit sc audio dac with second order noise shaping," *ISSCC 2000*, pp. 467–9, 00.
- [164] E. Tuijl, J. den homberg, D. reefman, C. Bastiaansen, and L. der dussen, "A 128f_s multi-bit $\delta\sigma$ cmos audio dac with real-time dem and 115db sfdr," *Proc. of ISSCC*, pp. 20.5–6, 04.
- [165] V. Colonna and A. Baschiroto, "A 10.7mhz selfcalibrated sc multibit 2nd-order bandpass ds modulator," *European Solid-State Circuits Conf.*, pp. 575–8, 02.
- [166] V. Colonna, G. gandolfi, F. stefani, and A. Baschiroto, "A 10.7mhz selfcalibrated switched-capacitor-based multibit bandpass $\delta\sigma$ modulator with on-chip switched buffer," *IEEE J. of Solid-state Circuits*, pp. 575–8, 08 04.
- [167] E. N. Aghdam and P. Benabes, "Increasing quantization frequency in a continuous-time delta-sigma modulator," *Soumis à Conference —*, pp. 1–2, 2006.

- [168] J. Cherry, M. Snelgrove, and W. Gao, "On the design of a forth-order continuous-time lc delta-sigma modulator for uhf a/d conversion," *IEEE Trans. on Circuits and Syst-II*, vol. 47, pp. 518–30, June 2000.
- [169] R. Schreier and B. Zhang, "Delta-sigma modulators employing continuous-time circuitry," *IEEE Trans. on Circuits and Syst.-I*, vol. 43, pp. 324–332, Apr. 96.
- [170] P. Coppejans, P. Vancorenland, W. Cock, and M. Steyaert, "Continuous time quadrature band-pass delta sigma modulator with input mixers," *IEE Proceedings G-Circuits, Devices and Systems*, pp. 331–6, Dec. 2002.
- [171] W. Bennett, "Spectra of quantized signals," *Bell Syst Tech. Journal*, vol. 27, pp. 446–472, July 1948.
- [172] V. Comino, M. Steyaert, and G. Temes, "A first-order current-steering sigma-delta modulator," *IEEE Journal and Solid-State Circuits*, vol. 26, pp. 176–182, March 1991.
- [173] V. Dias, G. Palmisano, and F. Maloberti, "Noise in mixed continuous-time switched-capacitors sigma-delta modulators," *IEE Proceeding-G*, vol. 139, pp. 680–84, Dec. 1985.
- [174] F. Henkel, U. Langmann, A. Hanke, S. Heinen, and E. Wagner, "A 1-mhz bandwidth second-order continuous-time quadrature bandpass sigma-delta modulator for low if radio receivers," *IEEE Journal and Solid-State Circuits*, vol. 37, pp. 1628–35, Dec. 2002.
- [175] I. Hsu and H. Luong, "A 70-mhz continuous-time cmos band-pass $\delta\sigma$ modulator for gsm receivers," *IEEE International Symposium on Circuits and Systems*, pp. 750–53, May 2000.
- [176] O. Shoaie and W. Snelgrove, "Optimal bandpass continuous time sigma-delta modulator," *Proc. IEEE ISCAS94*, pp. 489–492, 1994.
- [177] H. Tao and J. Khoury, "A 400-ms/s frequency translating bandpass sigma-delta modulator," *IEEE Journal and Solid-State Circuits*, vol. 34, pp. 1741–52, Dec. 1999.
- [178] K. Uchimura, T. Hayashi, T. Kimura, and A. Iwata, "Oversampling a/d and d/a converters with multi stage noise shaping modulators," *IEEE signal processing*, vol. 36, pp. 1899–1905, Dec. 1988.
- [179] L. Williams, "Third order sigma delta modulators," *IEEE Circuits and Syst*, vol. 38, pp. 489–498, May 1991.
- [180] E. va-der Zwan and E. Dijkmans, "A 0.2-mw cmos $\delta\sigma$ modulator for speech coding with 80db dynamic range," *IEEE Journal and Solid-State Circuits*, vol. 31, pp. 1873–1880, Dec. 1996.
- [181] J. Burghartz, D. Edelstein, M. Soyuer, H. Ainspan, and K. Jenkins, "Rf circuit design aspects of spiral inductors on silicon," *IEEE Journal and Solid-State Circuits*, vol. 33, Dec. 1998.
- [182] J. Yuan, W. Eisenstadt, and J. Liou, "A novel lossy and dispersive interconnect model for integrated circuit simulation," *IEEE Trans. Comp. Hybrids, Manufact. Technol*, vol. 13, pp. 275–80, June 1990.
- [183] C. Yue and S. Wong, "On-chip spiral inductors with patterned ground shields for si-based rf ic's," *IEEE Journal and Solid State Circuits*, vol. 33, pp. 101–109, May 1995.
- [184] P. Benabes, P. Aldebert, A. Yahia, and R. Kielbasa, "Influence of the feedback *dac* delay on continuous time band-pass sigma-delta coverter," *Electronics letters*, vol. 36, pp. 292–3, Feb. 2000.
- [185] J.-L. Degouy, P. Benabes, C. Flouzat, J.-M. Hode, and R. Kielbasa, "A new s-domain approach for designing continuous-time converters," *Proc. IEEE IMTC*, pp. 744–749, May 2001.

- [186] A. Yahia, P. Benabes, and R. Kielbasa, "A new technique to compensate the influence of the feedback dac delay in continuous-time bandpass ds converters," *Proc. IEEE Instrumentation and Measurement Technology Conference*, pp. 716–719, May 2001.
- [187] M. Albiol, J. Gonzalez, and E. Alarcon, "Mismatch and dynamic modeling of current sources in current-steering cmos d/a converters: an extended design procedure," *IEEE Trans. on Circuits and Syst-I*, vol. 51, pp. 159–169, Jan. 2004.
- [188] Y. Cong and R. Geiger, "Switching sequence optimization for gradient error compensation in thermometer-decoder dac arrays," *IEEE Trans. on Circuits and Syst-II*, vol. 47, pp. 585–595, July 2000.
- [189] M. Degrauwe, J. Rijmenants, E. Vittoz, and H. D. Man, "Adaptive biasing cmos amplifiers," *IEEE Journal and Solid-State Circuits*, vol. 17, pp. 522–28, June 1982.
- [190] H. Fiedler, "A 5-bit building-block for 20 mhz a/d converters," *IEEE Journal and Solid-State Circuits*, vol. 16, pp. 151–5, Sept. 1981.
- [191] S. Pipilos, Y. Tsvividis, J. Fenk, and Y. Papananos, "A si 1.8 ghz rlc filter with tunable center frequency and quality factor," *IEEE Journal and Solid-State Circuits*, pp. 1517–25, Oct 1996.
- [192] W. Song, H. Choi, S. Kwak, and B. Song, "A 10-b 20-msample/s low-power cmos adc," *IEEE Journal of Solid-State Circuits*, vol. 30, pp. 514–521, May 95.
- [193] Y. Tsvividis and M. Banu, "Continuous time mosfet-c filters in vlsi," *IEEE Journal and Solid-State Circuits*, vol. 21, pp. 15–30, Feb. 1986.
- [194] M. Banu and Y. Tsvividis, "Fully integrated active rc filters in mos technology," *IEEE Journal and Solid-State Circuits*, vol. 18, pp. 644–651, Dec. 1983.
- [195] Y. Tsvividis, "Integrated continuous-time filter design - an overview," *IEEE Journal and Solid State Circuits*, pp. 166–176, Mars 1994.
- [196] S. Tsukamoto, W. Schofield, and T. Endo, "A cmos 6-b 400-msamples/s adc with error correction," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 1939–47, 1998.
- [197] K. Yoon, S. Park, and W. Kim, "A 6b 500msample/s cmos flash adc with a background interpolated auto-zeroing technique," *IEEE Int. Solid-State Circuits Conference*, pp. 326–7, 1999.
- [198] S. Korotkov and D. Morozov, "A realisation of low-distortion cmos transconductance amplifier," *IEEE Trans. on Circuits and Sys.*, vol. 48, pp. 1138–40, Sept. 2001.
- [199] K. Kuo and A. Leuciuc, "A linear mos transistor using source degeneration and adaptive biasing," *IEEE Trans. on Circuits and Syst.-II*, vol. 48, pp. 937–943, Oct. 01.
- [200] E. Andre, "Etude d'une chaîne de conversion numérique-analogique pour applications multimédia," *ENSTBr*, Juin 1994.
- [201] J. Ho and H. Luong, "A 3-v, 1.47-mw, 120-mhz comparator for use in a pipeline adc," *Proc. of IEEE Asia Pacific Conference on Circuits and Syst.*, pp. 413–16, Nov. 96.
- [202] M. Ortmanns, F. Gerfers, and Y. Manoli, "Clock jitter insensitive continuous-time $\delta\sigma$ modulators," *IEEE Int. Conf. On Electronics, Circuits and Systems*, vol. 2, pp. 1049–52, 2001.
- [203] H. Tao, L. Toth, and J. Khoury, "Analysis of timing jitter in bandpass sigma-delta modulators," *IEEE Trans. Circuits and Systems II*, vol. 46, pp. 991–1001, 08 1999.
- [204] J. Arias, P. Kiss, V. Boccuzziand, L. Quintanilla, L. Enriquez, J. Vicente, D. Bisbal, J. Pablo, and J. Barbolla, "Nonlinearity correction for multibit $\delta\sigma$ dacs," *IEEE Trans. Circuits Syst.I*, vol. 52, pp. 1033–41, 06 2005.

- [205] A. Strak, A. Gothenberg, and H. Tenhunen, "Analysis of clock jitter effects in wideband sigma-delta modulators for rf-applications," *Analog Integrated Circuits and Signal Processing*, vol. 41, pp. 223–236, 2004.
- [206] S. Loeda, M. Reekie, and B. Mulgrew, "On the design of high performance wideband continuous-time sigma delta converters using numerical optimization," *IEEE Trans. Circuits Syst.I*, 2005.
- [207] M. Alboil, J. Gonzales, and E. Alarcon, "Improved design methodology for high-speed high-accuracy current steering d/a converters," *Design, Automation and Test in Europe Conference (DATE'03)*, pp. 1063–6, 09 2003.
- [208] S. Luschas, R. Schreier, and H. Lee, "Radio frequency digital-to-analog converter," *IEEE J. SOLID-STATE CIRCUITS*, vol. 39, pp. 1462–1467, 09 2004.
- [209] A. D. E. M. T. for Flash Analog-to Digital Converters, "E. säll and k. ola andersson and m. vesterbacka," *Proc. 6th Nordic Signal Processing Symposium-NORSIG 2004*, pp. 137–40, june 2004.
- [210] K. Lakshmikummar, R. Hadaway, and M. Copeland, "Characterization and modeling of mismatch in mos transistors for precision analog design," *IEEE J. Solid-State Circuits*, vol. 21, pp. 1057–66, 1986.
- [211] A. Cosand and A. Hills, "Multi-bit delta-sigma analog-to-digital converter with error shaping," *US.Patent*, Dec. 2002.

Liste des tableaux

2.1	Valeurs optimales des fréquences de résonance et amélioration de la résolution	29
2.2	Performances et caractéristiques des modulateurs $\Sigma\Delta$ multibits publiés ces dernières années.	48
2.3	Différents catégories de $\Sigma\Delta$ modulateur et leurs limites	49
3.1	Effets de la non idéalité des blocs du modulateur $\Sigma\Delta$ sur ses performance comme: " f_e " - " $ENOB$ ou $SNDR$ " - " $SFDR$ "	92
4.1	Comparaison des différentes solutions pour blanchir les tons de DWA dans un modulateur passe-bas	121
4.2	Différentes fonctions transferts pour SDEM d'ordre L	123
4.3	Propriétés principal des différentes méthodes de DEM	140
5.1	Comparaison des erreurs produites par une TDEM et celles d'une SDEM similaire, pour 4 éléments	155
5.2	Distribution de l'entrée aux cellules du CNA, TDEM normale et TDEM- tronquée . . .	164
5.3	Logique de décision employé dans l'ESB de l'algorithme STDEM	165
6.1	Taille nécessaire des registres des différents algorithmes d'ordre 4 passe-bande:	185
6.2	Nombres de blocs Tri-4 nécessaires aux méthodes MDEM et SDEM en fonction du nombre de cellules du CNA (M)	186
6.3	Comparaison des méthodes STDEM, MDEM et SDEM pour la correction des défauts d'appariement d'un CNA 3 bit.	187

Table des figures

2.1	Schéma général d'un convertisseur analogique-numérique (<i>CAN</i>) classique.	7
2.2	Quantification: (a) fonction de transfert, (b) bruit de quantification, (c) puissance de bruit de quantification.	8
2.3	Chaîne de conversion analogique-numérique.	9
2.4	Bruit de quantification dans le cas d'une conversion, (a) à la fréquence de <i>Nyquist</i> , (b) à suréchantillonnage, (c) sigma-delta.	10
2.5	Modulateur $\Sigma\Delta$	11
2.6	Modèle linéaire du modulateur $\Sigma\Delta$	13
2.7	schématique-simple d'un modulateur d'ordre un passe-bas	13
2.8	sortie numérique et spectre de la sortie d'un modulateur passe-bas monobit du premier ordre	14
2.9	Modèle quasi linéaire du modulateur en absence du signal d'entrée.	15
2.10	Modèle quasi linéaire du modulateur.	16
2.11	a)Modulateur passe-bas d'ordre 2, b)lieu des pôles du modulateur d'ordre 2	17
2.12	(a) Modulateur passe-bas d'ordre 3 instable, (b)Position des pôles en fonction du gain du comparateur d'un modulateur $\Sigma\Delta$ du 3 ^{ieme} ordre.	18
2.13	Modulateurs $\Sigma\Delta$: (a)passe-bas et (b) passe-bande	18
2.14	Chaîne de réception radio utilisant un passe-bande	19
2.15	Modulateur $\Sigma\Delta$ (a) à temps discret, (b) à temps continu.	20
2.16	Boucle ouverte du modulateur (a) à temps discret, (b) à temps continu.	21
2.17	Réponse temporelle du <i>CNA</i>	21
2.18	Réponse temporelle du <i>CNA</i> : (a) <i>NRZ</i> , (b) <i>RZ</i> , (c) <i>HZ</i>	22
2.19	L'effet du gigue d'horloge dans les modulateurs (a) à temps discret, (b) à temps continu.	23
2.20	SNR maximal du modulateur $\Sigma\Delta$ passe-bas; d'ordre $L_i = i^{th}$ et de quantification de $B_j = j - bits$ vis-à-vis du <i>OSR</i>	25
2.21	Architecture simple du modulateur $\Sigma\Delta$ passe-bas d'ordre L	26
2.22	Architecture simple du modulateur $\Sigma\Delta$ passe-bande d'ordre L	27
2.23	Performances du modulateur $\Sigma\Delta$ en fonction du l' <i>OSR</i> pour un $NTF = (1 - z^{-1})^L$ dans le cas d'un monobit d'ordre= {0...8} et idéal	28
2.24	<i>SNR/SNDR</i> et <i>DR</i>	31
2.25	Architecture d'un modulateur passe-bande d'ordre L avec multiple rétroaction retardée et des numérateurs Z^{-1}	34
2.26	Architecture général CIFB d'ordre pair du modulateur	34
2.27	Architecture général CIFB d'ordre impair du modulateur	35
2.28	Architecture d'un modulateur passe-bande d'ordre L avec une seul boucle et plusieurs alimentations-ver-l'avant <i>CIFF/CRFF</i>	36

2.29	Architecture d'un modulateur passe-bande d'ordre 6 avec une seule boucle et plusieurs branches parallèles <i>CIFF/CRFF</i> , développé à Dept. SSE-Supelec	36
2.30	Les modulateurs $\Sigma\Delta$ en cascade.	37
2.31	Modulateur $\Sigma\Delta$ <i>MASH 2</i>	38
2.32	Modulateur $\Sigma\Delta$ <i>DFOC</i>	39
2.33	Modulateur $\Sigma\Delta$ mixe 2-1.	40
2.34	Modulateur $\Sigma\Delta$ <i>MSCL</i> d'ordre 2.	40
2.35	Architecture du convertisseur $\Sigma\Delta$ passe-bande parallèle	41
2.36	a) Structure d'un modulateur avec dither, b) sortie du modulateur sans aucun dither, c) avec dither.	43
2.37	Le circuit d'intégrateur aux capacités commutées a) simple b) différentiel	44
2.38	Circuit d'intégrateur aux courants commutés simple	44
2.39	Différents types de réalisation d'un intégrateur à temps continu	45
2.40	Modulateur $\Sigma\Delta$ en utilisant une ligne de transmission <i>TL</i>	46
3.1	lieux présumés des erreurs du modulateur $\Sigma\Delta$	51
3.2	Exemple de la contribution des différents bruits dans un modulateur $\Sigma\Delta$, à la réalisation SC	53
3.3	Chaîne de conversion analogique-numérique (sans décimation)	54
3.4	a) échantillonnage, b) spectre du signal; suivant les différentes étapes d'échantillonnage avec un bloquer idéal	55
3.5	a) La structure d'un échantillonneur/bloqueur, b) le modèle électrique du SAH en phase d'échantillonnage	56
3.6	Effet de gigue d'horloge sur un signal sinusoïdal échantillonné	58
3.7	Structure d'un AOP simple avec ses sources de bruit thermique, b) circuit équivalent de bruit d'un AOP utilisé comme un amplificateur en combinaison rétroaction capacitif	59
3.8	Structure d'un modulateur $\Sigma\Delta$ passe-bas	61
3.9	Un intégrateur SC, qui est moins sensible aux imperfections des éléments	61
3.10	Modèle d'un intégrateur SC avec sources de bruit thermique a) quand $\phi_1 = 1$ b) quand $\phi_2 = 1$, c) l'équivalent	62
3.11	Quantificateur: a) le schéma bloc global, b) le schéma du comparateur c) la réponse d'un comparateur réel aux trois différents niveaux d'entrées	65
3.12	Un comparateur CMOS dynamique, il marque la sortie au front montant d'horloge	66
3.13	Une combinaison des deux amplificateurs d'une structure de rétroaction positive dans un comparateur	66
3.14	Modèle des erreurs du CNA interne dans un modulateur multibit	68
3.15	Exemple de fonction de transfert d'un convertisseur numérique-analogique multibit	69
3.16	Le schéma simple d'un CNA thermométrique sans DEM	70
3.17	Modèle statique d'une cellule du CNA réel située au chemin de retour	71
3.18	Le spectre de la sortie du modulateur passe-bande multibit a) idéal, b) avec un CNA réel ayant 1% d'erreur d'appariement sans aucune correction.	76
3.19	SNDR du modulateur multibit a) idéal, b) avec un CNA réel ayant 1%, 2% et 3% d'erreur d'appariement.	77
3.20	CNA de type source de courant pilotée a) schéma global de M cellules b) une source de courant et présentation d'erreur d'appariement	78

3.21	Modèle d'un CNA de type sources de courant piloté avec ses éléments parasites ainsi que sa réponse transitoire.	79
3.22	Les impulsions transitoires, a) <i>glitch</i> positif, b) <i>glitch</i> négatif.	81
3.23	Illustration de l'injection de charge d'horloge dans une cellule de CNA de type courant commuté.	82
3.24	Erreur de la sortie d'une cellule du CNA pour deux séries de donné équivalent : a)NRZ, b)RZ.	84
3.25	Lieu d'erreur de gigue d'horloge dans un modulateur : a) DT, b) CT général, c) exemple d'un CT de troisième ordre.	85
3.26	L'erreur de gigue d'horloge pour une cellule de CNA a) SC b) NRZ, c) RZ avec une durée de 50%.	85
3.27	La densité spectrale de gigue d'horloge réelle, ce qui n'est pas blanche	86
3.28	Modulateur à temps continu monobit a)avec un CNA de type SCR, b) avec un CNA sinusoïdal	88
3.29	Solution proposé afin de réduire l'effet de gigue d'hologe en faisant varier NTF a) modèle de gigue d'un CNA NRZ, b) l'optimisation de NTF dans la bande.	89
3.30	Solution proposé afin de réduire l'effet de gigue d'horloge par un filtrage Semi-FIR sur le signal de rebouclage a)position du filtre, b)un exemple defonctionnement de filtre. . .	90
3.31	Solutions proposées afin de réduire l'effet de retard dans la boucle d'un modulateur CT : a) rebouclage multiple [34], b) rebouclage rapide [30]	91
4.1	Schéma simple d'un CNA thermométrique avec DEM	95
4.2	Schéma d'algorithme RDEM	97
4.3	Spectre de la sortie du modulateur ayant 1% d'erreur de désappariement : a)idéal b) réel sans aucune correction c) correction par l'algorithme de RDEM d) cas similaire à (c) mais passe-bande	99
4.4	Méthode de commutation aléatoire quasi-RDEM en papillon ("Butterfly Shuffler") . .	100
4.5	Schéma interne d'une boîte en papillon pour réaliser RDEM	100
4.6	Schéma du brassage des sources par la méthode RDEM-segmentée, $M = 2^{\frac{B}{4}}$	101
4.7	Schéma de l'algorithme Clock Averaging, CLK-DEM, pour deux cellules.	103
4.8	Schéma général de l'algorithme Clock Averaging, dans le cas de 8-cellules.	103
4.9	Spectre de la sortie du modulateur multibit passe-bas avec un défaut d'appariement du CNA de 1% et correction par la méthode CLK-DEM, pour deux différents <i>OSR</i> . .	104
4.10	Tableau : l'approche additionnelle dans la méthode ILA-DEM, les éléments choisis pour une série des codes donnees, $x(n) \rightarrow x(n+9) \equiv \{2,5,4,7,5,2,2,4,5\}$, quant $M=8$	105
4.11	Spectre du modulateur passe-bas, a) idéal, b) méthode CLK-DEM, c) méthode ILA-DEM approche circulaire, d) méthode CLK-DEM, approche additionnelle, quand $M=8$	106
4.12	Les éléments choisis pour une série des codes donnés, $x(n) \rightarrow x(n+8) \equiv \{2,5,4,7,5,2,3,4\}$ par la méthode DWA, quand $M=8$	107
4.13	Spectre du modulateur passe-bas multibit ($M=16$) et $\sigma_\alpha = 1\%$: a) idéal, b) méthode DWA et l'entrée non périodique, c) méthode DWA avec une entrée périodique, d) méthode DWA avec une entrée DC.	110
4.14	Une réalisation simple de DWA conventionnelle pour un CNA de 16-éléments	111
4.15	Prévention de tons produit par la DWA dans un modulateur multibit à l'aide d'un dither, b) Le spectre du modulateur passe-bas multibit avec $M=16$, $ Dither _{max} = LSB$, $std(Dither) \simeq 0.01LSB$ et l'entrée périodique.	112

4.16	Quelques exemples de cycles utilisés pour l'algorithme de RDWA, b) le spectre d'un CNA multibit [120].	112
4.17	Cellules sélectionnées pour la séquence donnée $\{2,1,3,4,5,2,2,3,\dots\}$, quand $M=8$, a) algorithme de RnDWA, b) méthode de DWA classique.	113
4.18	Spectre de la sortie d'un modulateur multibit ayant algorithme RnDWA, b)l'amplitude maximal des tons dans la bande obtenue par la méthode RnDWA est comparée avec celle de DWA classique.	114
4.19	Structure du CNA entièrement différentiel à l'entrée d'un intégrateur SC, où la méthode de DIA peut être appliquée.	114
4.20	Résultats des simulations rapportés par [121]; SNDR maximum en fonction de l'OSR.	115
4.21	Spectre de la sortie du modulateur multibit [123] : a) DWA conventionnelle $M=16$ et $m = 0$, b)IDWA avec $M=16$ et $m = 1$, c)IDWA avec $M=16$ et $m = 3$	116
4.22	a)Une sélection possible par la méthode BIDWA pour $x=\{2,5,4,7,5,2,2,4,5,\dots\}$, $M=8$	117
4.23	a) Structure simple proposée pour la méthode P-DWA pour un CNA du 4-bits divisée en deux moitiés de 3-bit, chaque moitié utilise sa propre DWA conventionnelle	118
4.24	un exemple d'application passe-bas de la méthode P-DWA où le spectre d'erreur du CNA ($Do(n) - x(n)$) n'atteint pas zéro à la fréquence proche au DC ($f = 0$).	118
4.25	Méthode de DWA-partielle : a) le schéma bloc, b) un exemple de sélection des cellules pour un serai d'entrée	119
4.26	a)Une structure simple proposée pour la méthode Pseudo-DWA pour un CNA du 4-bits, b) spectre de la sortie avec Pseudo-DWA et $f_{Com_{inv}} = \frac{4*f_e}{OSR}$ puis $f_{Com_{inv}} = \frac{10*f_e}{OSR}$	120
4.27	Structure général de SDEM	122
4.28	Une boucle simplifiée de technique SDEM.	123
4.29	a)Spectre de la sortie de modulateur multibit avec 1% d'erreur corrigée par différents ordres de SDEM, b) la performance du CNA ayant un bloc SDEM qui est comparé par d'autre méthodes	125
4.30	Schéma bloc de la réalisation directe d'un algorithme de tri pour SDEM, avec $M=16$	125
4.31	Une stratégie simplifiée du tri pour SDEM, avec $M=8$	126
4.32	Structure générale de TDEM pour $M=16$	127
4.33	a) Structure générale de $SB_{k,r}$ pour l'algorithme TDEM, b) boucle de filtrage qui génère la séquence souhaitable pour $s_{k,r}(n)$	128
4.34	a) Schéma bloc simplifié d'une TDEM passe-bas : a)du premier ordre b) du deuxième ordre	128
4.35	a)Sortie du filtre dans une TDEM conventionnelle de second ordre, b) taille nécessaire des registres pour une TDEM de second ordre conventionnelle et une SDEM similaire.	130
4.36	SNDR du modulateur multibit utilisant une TDEM conventionnelle ou modifiée selon [133, 134].	132
4.37	a) Schéma de la fonction de transfert du quantificateur spécial qui produit $s_{k,r}$, b) schéma d'une TDEM passe-bas du troisième ordre utilisant un rebouclage non linéaire pour stabiliser sa boucle.	132
4.38	Spectre de la sortie de CNA lorsqu'on applique une entrée aléatoire : a)sans correction, b) corrigée par la TDEM d'ordre trois proposée en [141] et illustrée à la figure 4.37.	133
4.39	a) Schéma de TDEM segmenté, a)adapté aux éléments binaires, b) adapté aux CNA avec des éléments de poids différents.	133
4.40	Structure d'un modulateur multi-boucle "2-0", avec double quantificateur	135

4.41	Double-quantification, structure MASH qui n'a pas besoin de correction des erreur du CNA multibit car il est utilisé seulement au deuxième étage	135
4.42	Double-quantification, structure CIFB utilisent deux type de rebouclage; monobit pour les deux premiers intégrateurs et multibit pour les intégrateurs suivants.	136
4.43	Schéma d'un CNA série	136
4.44	Première stratégie l'étalonnage, a) le schéma global, b) le système en phase d'estimation des erreur du CNA multibit	138
4.45	Deuxième technique de l'étalonnage (background-calibration)	139
5.1	P-DWA modifiée, premier option qui emploie un diviseur aléatoire	142
5.2	Topologie générale proposée pour une B -bits MP-DWA	143
5.3	Bloc de commutation passe-bas qui fonctionne comme un diviseur spécial	143
5.4	Spectre du modulateur passe-bas multibit, $M=16$, $\sigma_\alpha = 1\%$: a) méthode P-DWA conventionnelle, b) P-DWA avec un diviseur aléatoire (premier option proposée), c) méthode MP-DWA proposé dans cette section	146
5.5	Performance d'un CNA 4-bits ayant 1% de défaut d'appariement qui est corrigée par une méthode de DWA passe-bande conventionnelle.	149
5.6	Sélection des éléments du CNA en passant de passe-bas à passe-bande: a) passe-bas du premier ordre, b) passe-haut du premier ordre, c) passe-bande du deuxième ordre.	150
5.7	Structure proposé pour le brassage des sources (DEM) de second ordre passe-bande, MP-DWA	151
5.8	Performance d'un CNA 4-bits ayant 1% de défaut d'appariement qui est corrigée par une méthode de MP-DWA passe-bande.	151
5.9	La structure proposé pour l'algorithme MDEM pratique, $M=16$	153
5.10	Comparaison de la fonctionnent des blocs de commutation: a,b) s_{41} et s_{21} dans une TDEM conventionnel, c) s_{41} dans l'MDEM-algorithme proposé, d) $sv_1(n) + sv_1(n) - sv_3(n) - sv_4(n)$ à SSB1 dans le MDEM proposé	156
5.11	Comparaison les différentes techniques du brassage des sources, TDEM, MDEM, SDEM, ayant 3% défaut d'appariement, $OSR=64$, $M=16$	160
5.12	Structure d'une SSB simple conçue pour une MDEM passe-bande d'ordre quatre	161
5.13	Comparaison de différentes techniques de DEM avec l'erreur de disparité de 1%.	161
5.14	a) L'algorithme Proposé en structure arborescence tronquée (STDEM), b) le schéma fonctionnel d'une ESB	163
5.15	Spectre de la sortie du modulateur sigma delta passe-bande d'ordre 8: a) avec CNA idéal, b) sans aucune correction mais $\sigma_\alpha = 3\%$, c) en utilisant une TDEM conventionnelle d'ordre 4 passe-bande, d) en utilisant une STDEM d'ordre 4 proposée dans cette section.	167
5.16	Comparaison de la performance des différentes techniques de DEM dans un modulateur passe-bande	168
5.17	$SB_{k,r}$ modifié pour STDEM, utilisé dans toutes les couches avant le dernier, $k \geq 3$	169
5.18	Schéma d'un bloc de ESB_r , utilisé dans la dernière couche de STDEM	169
5.19	Sortie du deuxième résonateur utilisé dans la STDEM qui dépasse rarement au dessus du 16, a) R11, b)R12 c)R21; expliquées au tableau 5.3.	171
5.20	Niveaux maximum de la sortie des registres, pour un ESB usées dans la STDEM et pour un SB utilisé dans la TDEM conventionnel	172

5.21	Simulation effectuée au niveau Transistor qui est implantée à l'environnement Cadence, pour un modulateur multibit qui emploie une algorithmes STDEM d'ordre quatre passe-bande	173
5.22	Rapporte de l'Ambit, pour une STDEM ayant 8-cellules	173
5.23	Comparaison le SNDR des différentes techniques de DEM du quatrième ordre passe-bande	174
6.1	Structure d'un modulateur passe-bas du troisième ordre cas: a) à filtre simple, b) à rebouclage multiple "CIFB", c) MSCL, d) à rebouclage multiple à temps continu . . .	178
6.2	Modulateur passe-bas idéal: a) signal de sortie $v(n)$, b) spectre de la sortie $PSD[v(n)]$, c) densité spectrale du bruit en sortie $PSD[v(n) - x(n)]$	179
6.3	SNR et SFDR du modulateur passe-bas idéal en fonction de l'amplitude de l'entrée, $OSR=64$, $M=16$	180
6.4	Structures passe-bande intéressantes: a) Cas discret, CRFB-retard-multiple, b) Cas continu, CRFB-retard-multiple avec des retards entiers, c) Cas continu, CRFB-retard-multiple avec des retard optimisés pour éliminer l'effet des imperfections des résonateurs d) Cas continu, rebouclage simple architecture parallèle	181
6.5	Modulateur passe-bande: a) sortie du modulateur dans le cas CRFB-retard-multiple discret, b) sortie dans le cas CRFB-retard-multiple à temps continu c) sortie dans le cas rebouclage simple architecture parallèle.	182
6.6	Spectre de la sortie d'un modulateur multibit passe-bande avec un CNA désapparié de 1%, a) corrigée par la MDEM, b) corrigée par la STDEM.	183
6.7	Schéma bloc de la MTF: a) SSB employé dans la SDEM ou dans la MDEM, b) ESB ou SB employé dans la STDEM ou dans les premières couches de la MDEM	184
6.8	Schéma bloc du tri d'un vecteur de 8 éléments n'utilisant que des blocs de tri de 4 éléments.	185
6.9	Tri d'un vecteur de 16 éléments en utilisant des blocs de tri de 8 éléments.	186
6.10	Schéma bloc du modulateur $\Sigma\Delta$ multibit passe-bande du 6 ^{ième} ordre à filtre LC.	187
6.11	Résultat de simulation au niveau Transistor dans l'environnement Cadence, pour un modulateur multibit en utilisant un algorithme passe-bande d'ordre quatre a) MDEM b) STDEM	188
6.12	Compte-rendu d'Ambit, donnant la surface minimale nécessaire à l'implanation d'un algorithme de DEM pour un CNA 3 bit: a) STDEM b)MDEM c)SDEM	188
6.13	Routage final pour les circuits conçus de la technique SDEM dans le cas $M=8$, ($area = 0.69mm^2$).	189
6.14	a) Surface final pour les circuits conçus de la technique MDEM dans le cas $M=8$, ($area = 0.53mm^2$), b) routage final pour les circuits conçus de la technique STDEM dans le cas $M=8$, ($area = 0.24mm^2$).	189
6.15	Compte-rendu d'Ambit pour les chemins les plus longs: a)STDEM b)SDEM	190
6.16	Circuit CNA 3-bit à code thermométrique.	190
6.17	Source de courant commutée.	191
6.18	a) Circuit d'un Déglitcher simple, b) circuit d'un SRD simple, c) les signaux d'entrants et les signaux sortants du deglitcher et du SRD, et de la sortie d'un cellule courant commuté	193
6.19	Réponse temporelle du CNA.	194
6.20	Cellule proposée pour un CNA semi-RZ a, b) circuit, c) les signaux	195

6.21	Signal de sortie du modulateur a) passe-bas b) passe-bande.	197
6.22	Comparateur RZ dégénératif avec un rebouclage positive : a)circuit, b)réponse transitoire	200
6.23	Spectre de sortie du modulateur passe-bande en utilisant un quantificateur d'une architecture ping-pong : a) cas idéal, b) avec 10 pourcent d'erreur	201
6.24	Deux <i>CAN</i> -flash en alternance.	201