



**HAL**  
open science

# Test dynamique des convertisseurs analogiques-numériques rapides : caractérisation par analyse temporelle, test industriel et considérations instrumentales

Djamel Haddadi

► **To cite this version:**

Djamel Haddadi. Test dynamique des convertisseurs analogiques-numériques rapides : caractérisation par analyse temporelle, test industriel et considérations instrumentales. Micro et nanotechnologies/Microélectronique. Université Bordeaux 1, 2002. Français. NNT: . tel-01219969

**HAL Id: tel-01219969**

**<https://hal.science/tel-01219969>**

Submitted on 23 Oct 2015

**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.

N° d'ordre : 2622

# THÈSE

présentée à

## L'UNIVERSITÉ BORDEAUX I

ÉCOLE DOCTORALE DES SCIENCES PHYSIQUES ET DE L'INGENIEUR

par **Djamel HADDADI**

POUR OBTENIR LE GRADE DE

### DOCTEUR

SPÉCIALITÉ : **ELECTRONIQUE**

\*\*\*\*\*

**TEST DYNAMIQUE DES CONVERTISSEURS ANALOGIQUES-NUMERIQUES  
RAPIDES: CARACTERISATION PAR ANALYSE TEMPORELLE, TEST  
INDUSTRIEL ET CONSIDERATIONS INSTRUMENTALES**

\*\*\*\*\*

Soutenue le : 19 Décembre 2002

Après avis de :

**M. C. MORANDI**  
**M. Y. BERTRAND**

Professeur, Université de Parme (Italie)  
Professeur, Université Montpellier II

**Rapporteur**  
**Rapporteur**

Devant la commission d'examen formée de :

**M. Y. DANTO**  
**M. D. DALLET**  
**M. Y. BERTRAND**  
**M. N. HAKIM**  
**M. Ph. MARCHEGAY**  
**M. C. MORANDI**

Professeur, Université de Bordeaux I  
Maître de Conférences, ENSEIRB  
Professeur, Université Montpellier II  
Chef de Projet, STMicroelectronics  
Professeur, ENSEIRB  
Professeur, Université de Parme (Italie)

**Président**  
**Rapporteur**  
**Examineurs**



N° d'ordre : 2622

# THÈSE

présentée à

## L'UNIVERSITÉ BORDEAUX I

ÉCOLE DOCTORALE DES SCIENCES PHYSIQUES ET DE L'INGENIEUR

par **Djamel HADDADI**

POUR OBTENIR LE GRADE DE

### DOCTEUR

SPÉCIALITÉ : **ELECTRONIQUE**

\*\*\*\*\*

**TEST DYNAMIQUE DES CONVERTISSEURS ANALOGIQUES-NUMERIQUES  
RAPIDES: CARACTERISATION PAR ANALYSE TEMPORELLE, TEST  
INDUSTRIEL ET CONSIDERATIONS INSTRUMENTALES**

\*\*\*\*\*

Soutenue le : 19 Décembre 2002

Après avis de :

**M. C. MORANDI**  
**M. Y. BERTRAND**

Professeur, Université de Parme (Italie)  
Professeur, Université Montpellier II

**Rapporteur**  
**Rapporteur**

Devant la commission d'examen formée de :

**M. Y. DANTO**  
**M. D. DALLET**  
**M. Y. BERTRAND**  
**M. N. HAKIM**  
**M. Ph. MARCHEGAY**  
**M. C. MORANDI**

Professeur, Université de Bordeaux I  
Maître de Conférences, ENSEIRB  
Professeur, Université Montpellier II  
Chef de Projet, STMicroelectronics  
Professeur, ENSEIRB  
Professeur, Université de Parme (Italie)

**Président**  
**Rapporteur**  
**Examineurs**



***A mes Parents***

***A Hassiba***

***A Youcef***

*Les travaux présentés dans ce mémoire ont été effectués en grande partie au sein du Laboratoire de Microélectronique IXL de l'université de Bordeaux I dirigé par Monsieur le professeur A. TOUBOUL. Je tiens à lui exprimer ma profonde reconnaissance pour la confiance qu'il m'a témoignée en m'accueillant au sein de son laboratoire.*

*Que Monsieur le Professeur Philippe MARCHEGAY, directeur de cette thèse, sache que sa confiance et son soutien ont été précieux pour la réalisation de ce travail et qu'il soit assuré de ma profonde gratitude.*

*Je tiens à témoigner ma reconnaissance à Monsieur Dominique DALLET, Maître de Conférences à l'ENSEIRB, pour ses conseils, sa grande disponibilité, son suivi ainsi que sa profonde amitié.*

*Je tiens à exprimer mes sincères remerciements à Monsieur Y. DANTO, Professeur à l'Université de Bordeaux I, pour l'honneur qu'il me fait en présidant ce jury.*

*Je remercie vivement Monsieur Carlo MORANDI, Professeur à l'Université de Parme, et Monsieur Yves BERTRAND, Professeur à l'Université Montpellier II, pour l'intérêt qu'ils ont porté à mes travaux en acceptant d'examiner ce mémoire et d'en être les rapporteurs.*

*Que Monsieur Nazir HAKIM, Chef de Projet R&D chez STMicroelectronics Crolles, soit assuré de ma reconnaissance pour avoir accepté de corriger ce mémoire et pour sa participation à ce jury, apportant ainsi l'avis du milieu industriel.*

*Que Monsieur A. CHASSAGNEUX, Responsable du département Product Engineering de la division Standard Linear IC's de STMicroelectronics Grenoble, soit vivement remercié pour son soutien et pour l'intérêt qu'il a porté à mes travaux.*

*Je remercie également Monsieur Guy JESTIN du Laboratoire CELAR de la DGA, pour l'intérêt qu'il a porté à mes travaux et pour les nombreuses données qu'il m'a fournies.*

*Enfin, j'exprime ma sympathie à tous les membres du laboratoire que je n'ai pas nommés, aux collaborateurs du Projet Européen DYNAD, et aux membres de l'équipe "High Speed Data Converters" de chez STMicroelectronics Grenoble.*





**SOMMAIRE**

---

## SOMMAIRE

<b>INTRODUCTION.....</b>	<b>1</b>
<b>PARTIE I: CARACTERISATION PAR ANALYSE TEMPORELLE.....</b>	<b>7</b>
<b>CHAPITRE 1: PARAMETRES DYNAMIQUES DES CONVERTISSEURS ANALOGIQUES-NUMÉRIQUES.....</b>	<b>8</b>
1. Introduction.....	11
2. Les sources d'imperfections des CANs.....	12
3. Les paramètres clés des CANs.....	13
4. Définition temporelle des paramètres dynamiques.....	16
4.1. Mode single-tone.....	16
4.2. Mode dual-tone.....	20
5. Conclusion.....	22
<b>CHAPITRE 2: LA MÉTHODE DES MOINDRES CARRÉS NON LINÉAIRES.....</b>	<b>23</b>
1. Introduction.....	25
2. L'estimateur des moindres carrés.....	25
2.1. Les Bornes de Cramer-Rao.....	27
2. Algorithmes pour des problèmes de moindres carrés non linéaires.....	29
3. Méthode de séparation des variables.....	31
4. Critère d'arrêt et valeurs initiales.....	32
5. Conclusion.....	33
<b>CHAPITRE 3: MODE SINGLE-TONE: METHODE SANS SÉPARATION DES VARIABLES.....</b>	<b>35</b>
1. Introduction.....	37
2. Estimation des paramètres du signal de test.....	38
2.1. Estimation des valeurs initiales.....	41

---

2.1.1. Résultats de simulation.....	43
2.2. Une méthode globalement convergente.....	47
2.2.1. Résultats de simulation.....	49
3. Estimation des Produits Harmoniques.....	49
4. Simulations de Monte-Carlo et Résultats réels.....	51
5. Conclusion.....	54
<b>CHAPITRE 4: MODE SINGLE-TONE: METHODE</b>	
<b>DE SÉPARATION DES VARIABLES.....</b>	<b>55</b>
1. Introduction.....	57
2. Estimation des paramètres du modèle.....	58
2.1. Paramètres linéaires.....	58
2.2. Paramètre non linéaire.....	59
2.2.1. Négliger la contribution de la distorsion harmonique...60	
2.2.2. L'incertitude tolérée sur la valeur initiale.....	62
2.2.3. Un algorithme robuste.....	64
4. Simulations de Monte-Carlo et résultats réels.....	66
5. Conclusion.....	72
<b>CHAPITRE 5: ANALYSE TEMPORELLE EN MODE DUAL-TONE.....</b>	<b>73</b>
1. Introduction.....	75
2. Estimation des paramètres du signal de test.....	76
2.1. Estimation des paramètres non linéaires.....	77
2.1.1. Problème des valeurs initiales.....	77
3. Estimation des produits d'intermodulation.....	78
4. Simulations de Monte-Carlo et résultats réels.....	79
5. Conclusion.....	82
<b>PARTIE II: TEST INDUSTRIEL ET CONSIDERATIONS INSTRUMENTALES...85</b>	
<b>CHAPITRE 6: TEST INDUSTRIEL.....</b>	<b>87</b>
1. Introduction.....	89
2. L'algorithme.....	89
3. Temps de test.....	92

---

4. Répétabilité.....	95
5. Conclusion.....	98
<b>CHAPITRE 7: MINIMISER LES PERTURBATIONS DE LA CARTE DE TEST.....</b>	<b>99</b>
1. Introduction.....	101
2. Le bruit des sources d'alimentation.....	102
2. La pureté spectrale du signal test.....	104
3. Bruit de commutation et diaphonie.....	106
4. Le Signal d'horloge.....	111
5. Empilement des couches et choix du diélectrique.....	113
6. Conclusion.....	114
<b>CHAPITRE 8: COMPENSATION DU BRUIT DE PHASE DU GENERATEUR DE TEST.....</b>	<b>115</b>
1. Introduction.....	117
2. Le bruit de phase.....	118
4. Effet du bruit de phase sur les performances d'un CAN.....	120
2. Compensation du bruit de phase.....	121
2.1. Spectre en puissance moyenné.....	122
2.2. Le filtrage médian.....	122
2.3. Compensation.....	124
3. Résultats.....	125
4. Conclusion.....	127
<b>CONCLUSION.....</b>	<b>129</b>
<b>ANNEXE A.....</b>	<b>135</b>
A.1. Dérivation des bornes de Cramer-Rao.....	137
A.1.1. Mode Single-tone.....	137
A.1.2. Mode Dual-tone.....	141
A.2. Détermination de la densité de probabilité.....	142
A.3. La méthode "Line-search".....	143
<b>REFERENCES BIBLIOGRAPHIQUES.....</b>	<b>145</b>

**INTRODUCTION**



## INTRODUCTION

Les convertisseurs analogiques-numériques (CANs) représentent, d'un point de vue économique, un grand marché qui se chiffre en milliard de dollars. C'est aussi un marché compétitif et à forte croissance. Les CANs sont des produits à forte valeur ajoutée, mais d'une grande difficulté technique sur le plan de la conception comme sur celui du test. Cette activité nécessite des investissements lourds en termes de matériels et de ressources humaines. En effet, dépasser la taille critique est indispensable pour se faire une place sur ce marché. Dans les applications à fort volume (marché de masse), telles que les disques durs ou encore les Setup Box, les CANs sont purement et simplement intégrés avec les autres composants pour donner le système sur puce (system on chip). Le marché des CANs discrets est limité à des applications industrielles ou à moyen, voire faible volume.

L'intérêt des CANs réside dans le fait qu'ils remplacent avantageusement les vieilles solutions analogiques rigides, encombrantes et coûteuses. L'exemple typique est celui des stations de base, où les CANs permettent de réduire le nombre de fréquences intermédiaires et de gérer plusieurs standards à la fois. La tendance actuelle pour les CANs est l'intégration du VGA (variable gain amplifier) d'entrée et du traitement digital, tel que le filtrage numérique. Les performances des CANs rapides ne cessent de croître, tirées vers le haut par deux types d'applications. D'une part, les applications orientées Télécom, telles que les stations de base, exigent des CANs rapides avec une bande passante effective large et des performances dynamiques élevées. Les CANs destinés à ces applications sont généralement conçus en technologie BiCMOS et sont donc gourmands en consommation d'énergie. D'autre part, les applications portables, telles que l'imagerie ou l'instrumentation, qui sont moins contraignantes en termes de vitesse, mais exigeantes au niveau de la consommation et de l'encombrement. Les CANs dédiés à ces applications sont généralement réalisés en technologie



CMOS qui permet de réduire la consommation et d'augmenter le nombre de canaux sur la même puce.

Réussir l'introduction d'un nouveau produit sur le marché passe par trois étapes essentielles:

1. Le recensement des besoins clients. C'est le travail des commerciaux techniques.
2. La conception de produits qui répondent parfaitement aux besoins des clients. Ce travail est celui des concepteurs (designers).
3. Enfin, l'industrialisation des produits. C'est le travail des ingénieurs de test et produit.

La réduction des durées de ces différentes étapes, permet de mettre le produit rapidement sur le marché et d'arriver avant les concurrents. Avec l'augmentation des performances des produits, la phase de test devient de plus en plus lente et coûteuse. La maîtrise des méthodes de test permet justement de pallier à ce problème.

La difficulté que présente le test des CANs se situe à deux niveaux complémentaires. D'une part, il faut des méthodes rapides et précises pour extraire les paramètres de performance du produit à tester. D'autre part, il faut un matériel de test (carte et instrumentation de test) performant afin de ne pas dégrader les performances du produit à tester. La maîtrise de ses deux aspects permet de réduire considérablement le temps nécessaire à l'industrialisation du produit.

Sur l'aspect algorithmique, il existent trois méthodes pour calculer les paramètres statiques et dynamiques des CANs. Les paramètres statiques sont calculés par l'analyse statistique [Blair, 1994], [Morandi et al., 1994], [Dallet, 1995], [dynad, 2001], [ieeestd1241, 2001], alors que les paramètres dynamiques sont estimés par les analyses spectrale et temporelle. L'analyse spectrale, basée sur la transformée de Fourier rapide, est sans doute la méthode la plus utilisée pour calculer les paramètres dynamiques [Benkais, 1993], [Dallet, 1995], [dynad, 2001], [ieeestd1241, 2001]. Cependant, cette approche exige que la condition dite de cohérence soit satisfaite. Cette dernière nécessite des générateurs très

précis et une synchronisation entre le générateur d'horloge et celui du signal de test. Sinon, un phénomène d'étalement spectral (leakage) apparaît, empêchant l'application directe de l'analyse spectrale. En pratique, il n'est pas toujours possible de satisfaire la condition de cohérence, pour des raisons soit de coût soit d'application. En effet, d'une part les générateurs précis sont coûteux, et d'autre part, dans l'application, l'horloge est en général générée par un oscillateur local, qui est forcément non synchronisé par rapport au signal de test. Dans ces situations il y a deux alternatives possibles. La première consiste à utiliser l'analyse spectrale avec des fenêtres de pondération [**dynad, 2001**]. Cette solution pose cependant le problème délicat du choix de la bonne fenêtre de pondération [**Nuttall, 1981**], [**Harris, 1978**]. La deuxième alternative consiste à utiliser l'analyse temporelle qui, de part son principe, est insensible au phénomène d'étalement spectral (leakage).

La première partie de cette étude, réalisée dans le cadre du projet européen DYAND [**Morandi et al., 2000**], [**dynad, 2001**], est consacrée au test dynamique des CANs par l'analyse temporelle. En terme de précision, cette méthode est parfaitement adaptée au test des CANs, dans la mesure où nous avons à résoudre un problème avec un grand nombre d'échantillons et un rapport signal à bruit élevé. Cependant, la vitesse de la méthode décroît lorsque le nombre d'échantillons augmente. La vitesse diminue également avec l'augmentation du nombre de paramètres du modèle à estimer, autrement dit la couverture de test. L'analyse temporelle est basée sur des méthodes itératives, ce qui soulève le problème épineux de la convergence.

A cause des problèmes qui viennent d'être évoqués, l'analyse temporelle est souvent utilisée pour estimer un seul paramètre dynamique, à savoir le nombre effectif de bits [**Malewski et al., 1987**], [**Leroux, 1989**], [**Hejn et al., 1992**], [**ieeestd1241, 2001**]. En effet, dans ce cas le problème à résoudre est relativement simple. L'approche présentée dans [**Pintelon et al., 1996**] permet d'estimer pratiquement tous les paramètres du CANs en mode single-tone. Cependant, l'implantation de cette méthode est plutôt lourde et elle est relativement lente, notamment lorsque le nombre d'échantillons augmente. Cette méthode est plutôt adaptée aux problèmes ayant une forte distorsion harmonique,

ce qui n'est pas notre cas. Dans tous les travaux que nous venons de citer, le problème de convergence n'a pas été étudié. On se contente généralement de souligner l'importance de choisir des valeurs initiales précises afin d'assurer la convergence des algorithmes, ce qui est une évidence. Par ailleurs, la caractérisation des CANs par l'analyse temporelle en mode dual-tone n'a, à notre connaissance, jamais été traitée.

Dans la première partie de cette étude, on se propose d'étudier avec plus de détails les problèmes de l'analyse temporelle qui viennent d'être évoqués, à savoir: précision, convergence, valeurs initiales, temps d'exécution et simplicité d'implantation. Cette analyse sera réalisée en mode single-tone comme en mode dual-tone.

La deuxième partie de cette étude, réalisée au sein de la société STMicroelectronics, traite de l'implantation industrielle des analyses spectrale et temporelle et des considérations matérielles, à savoir la carte et l'instrumentation de test. L'implantation industrielle abordera les problèmes de répétabilité et de temps de test. La réalisation d'une carte mixte performante qui ne dégrade pas les performances du produit à tester, constitue un défi de plus en plus difficile pour les ingénieurs de test. La mise en œuvre des différents tests (single-tone et dual-tone) nécessite un certain nombre de précautions qui seront également abordées. Enfin, nous aborderons les problèmes posés par les limitations de l'instrumentation de test, et nous proposerons une méthode pour compenser le bruit de phase du générateur de test.

**PARTIE I**

**CARACTERISATION PAR ANALYSE TEMPORELLE**



**CHAPITRE 1**

**PARAMETRES DYNAMIQUES DES  
CONVERTISSEURS ANALOGIQUES-NUMERIQUES**



# CHAPITRE 1

## PARAMETRES DYNAMIQUES DES CONVERTISSEURS ANALOGIQUES-NUMERIQUES

### 1. Introduction

Les convertisseurs analogiques-numériques (CANs) sont, avec les convertisseurs numériques-analogiques (CNAs), les composants clés de l'électronique d'aujourd'hui. En effet, les traitements effectués sur les signaux sont de plus en plus réalisés avec des outils numériques. Par contre, le monde qui nous entoure est complètement analogique et les capteurs récupérant une information à travers un signal de type tension ou courant, le sont aussi.

Il est donc nécessaire d'utiliser une interface pour passer du monde analogique au monde numérique et vice-versa. La numérisation s'effectue au moyen de composants électroniques non parfaits par principe. Cette opération va donc engendrer des déformations sur le signal à traiter. Celles-ci seront systématiques par rapport au processus de quantification et aléatoires par rapport à la non idéalité des composants.

Il est important de pouvoir qualifier une chaîne d'acquisition qui se fait à travers l'évaluation des performances du CAN. La définition de tous les paramètres caractérisant les CANs dépasse de loin le cadre de notre étude et il est conseillé pour cela de consulter l'abondante littérature dans ce domaine, notamment les deux standards [dynad, 2001], [ieeestd1241, 2001]. Ici, nous allons nous limiter à



la définition des paramètres clés et nous détaillerons surtout les paramètres dynamiques qui peuvent être estimés par l'analyse temporelle.

## 2. Les sources d'imperfections des CANs

Un CAN réel, quelle que soit son architecture, possède un certain nombre de sources de bruit et de distorsion, comme le montre la figure (1-1). Le buffer analogique d'entrée possède un bruit large bande, une non-linéarité, et une bande passante limitée. L'échantillonneur-bloqueur introduit une non-linéarité supplémentaire, une limitation de la bande passante, et de la gigue (jitter). Le quantificateur, quant à lui génère le bruit de quantification et les non-linéarités différentielle et intégrale.

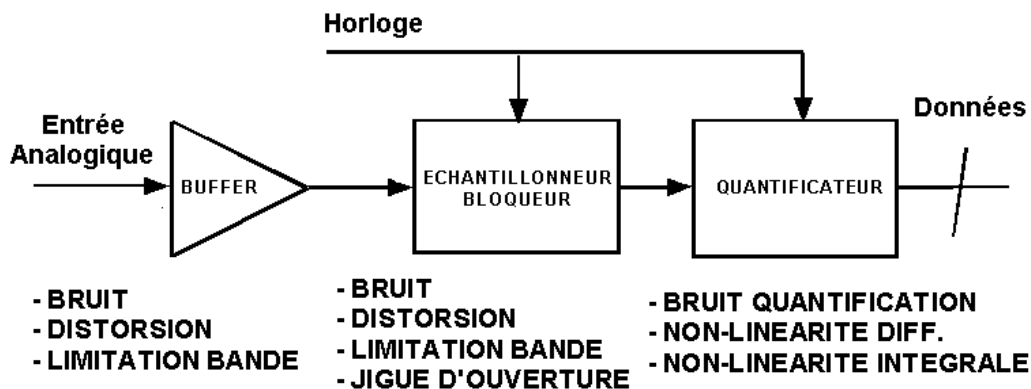


Figure 1-1: Modèle d'un CAN montrant les sources de bruit et de distorsion.

Les circuits large bande internes du CAN produisent un bruit large bande à cause des effets thermiques  $kT/C$ , C étant la capacité d'échantillonnage. Ce bruit est présent même pour un signal d'entrée DC pour lequel la distribution des codes de sortie est une gaussienne centrée autour de la valeur nominale du signal DC, comme le montre la figure (1-2). Pour mettre en évidence ce phénomène, il suffit de mettre à la masse l'entrée du CAN et de tracer l'histogramme des données ainsi obtenues.

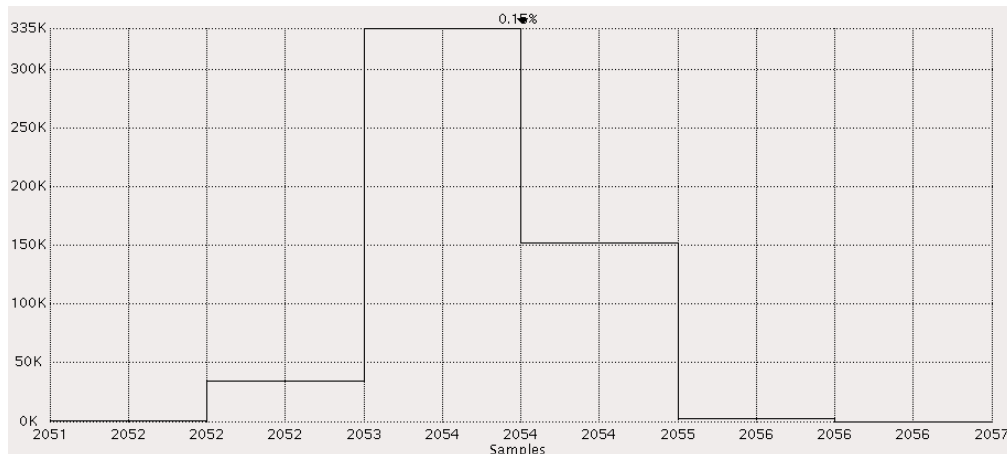


Figure 1-2: Histogramme des données d'un CAN réel 12 Bits avec un signal de test DC.

### 3. Les paramètres clés des CANs

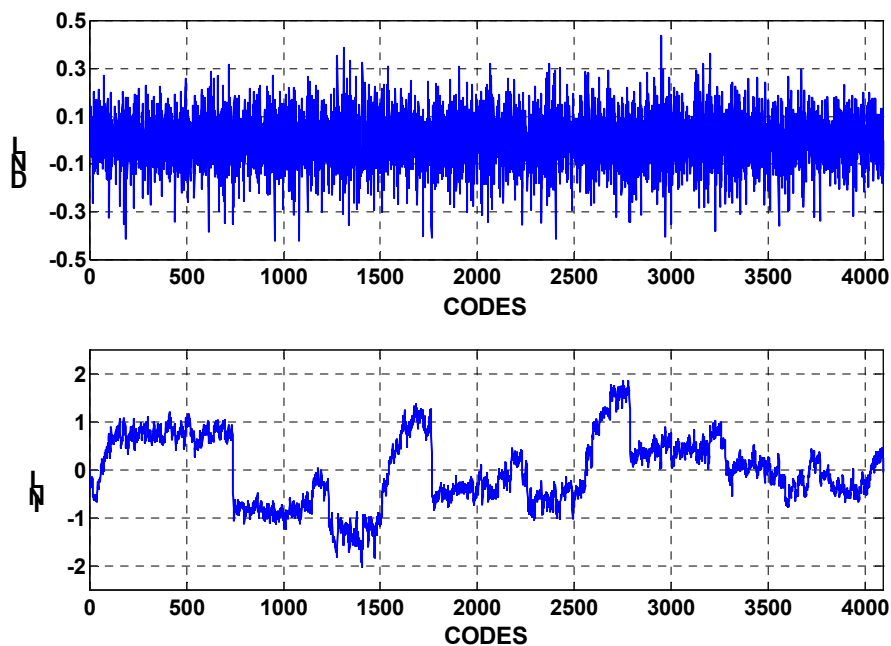
Les imperfections (bruit, distorsion) d'un CAN sont quantifiées par des paramètres statiques et dynamiques. Les principales erreurs statiques sont au nombre de quatre:

- l'erreur d'offset (OE) et l'erreur de gain (GE)
- la non-linéarité différentielle (DNL) et la non-linéarité intégrale (INL)

Les paramètres statiques sont calculés par une technique utilisant l'analyse statistique. La difficulté de cette méthode provient du fait qu'elle nécessite un grand nombre d'échantillons, notamment pour les hautes résolutions, et ce à cause de la précision requise sur l'INL et la DNL. La figure (1-3) montre un exemple donnant les courbes de linéarité d'un CAN réel obtenues avec une fréquence d'horloge de 60MHz, un signal de test de 10MHz, et 500K échantillons.

En pratique, ce test nécessite un système d'acquisition (analyseur logique, carte d'acquisition) ayant une grande taille mémoire, ce qui du coup augmente le coût du banc de test. Pour le test industriel, le problème n'est pas tellement la taille de la mémoire mais l'augmentation du temps de test due à l'acquisition et au traitement d'un grand nombre de points. En effet, plus la résolution d'un CAN, à

savoir le nombre de bits sur lequel est codé le signal, augmente, plus la fréquence d'échantillonnage sera faible. Il est alors évident que le test statistique d'un tel convertisseur sera très coûteux en terme de temps de test. Pour ce type de produits, des outils de réduction de temps de test basés sur la modélisation des linéarités du CAN ont été développées par la société ARDEXT Technologies [Cherubal et al.]. Ils permettent un gain d'au moins un facteur 2 sur le temps de test moyennant un coût software relativement élevé.



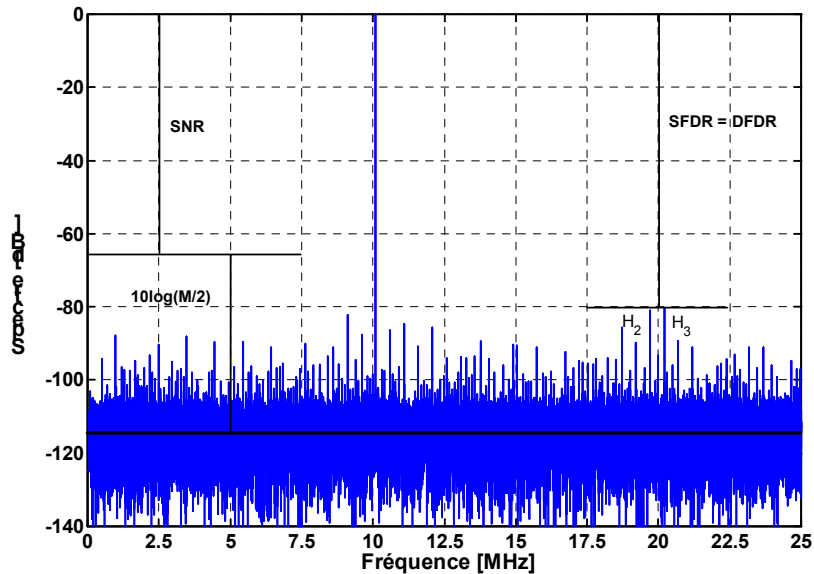
**Figure 1-3:** DNL et INL d'un CAN réel 12 bits.  $f_e = 10\text{MHz}$ ,  $f_h = 60\text{MHz}$ ,  $M = 500\text{K}$  échantillons.

En fait, on parle de paramètre statique lorsque les erreurs sont directement liées à la caractéristique de transfert d'un CAN. L'analyse statistique, par son principe, minimise l'influence des erreurs dynamiques telles que la gigue, le bruit thermique ... Il est important de quantifier ces phénomènes dynamiques à travers les principaux paramètres:

- la distorsion harmonique (THD),
- le rapport signal-à-bruit (SNR),
- le rapport signal-à-(bruit et distorsion) (SINAD),
- le nombre de bits effectifs (ENOB),
- la dynamique des parasites (SFDR),

- la dynamique de distorsion (DFDR)
- la distorsion d'intermodulation (IMD).

La figure (1-4) montre le spectre d'un CAN réel avec la définition spectrale d'un certain nombre de paramètres. En effet, la méthode la plus utilisée, mais aussi la



**Figure 1-4:** Spectre d'un CAN réel 12 bits avec la définition des principaux paramètres.  $f_h = 50\text{MHz}$ ,  $f_e = 10\text{MHz}$ .

plus étudiée, pour calculer ces paramètres est sans doute l'analyse spectrale, méthode basée sur la transformée de Fourier rapide (FFT). Cependant, comme nous le verrons par la suite, cette technique impose certaines contraintes à l'instrumentation de test, qui ne sont pas toujours faciles à satisfaire en pratique. Les paramètres dynamiques, excepté le SFDR, peuvent aussi être estimés par l'analyse temporelle, une méthode à laquelle on reproche souvent sa lenteur et des problèmes de convergence. Nous verrons par la suite que cela n'est pas tout à fait exact, et que les réticences à l'égard de cette méthode s'expliquent plutôt par sa plus grande complexité, comparée à l'analyse spectrale, du moment où elle fait appel à des techniques d'optimisation non-linéaires.

Il faut noter aussi que pour les CANs rapides actuels, la dynamique est généralement limitée par la distorsion et donc la DFDR est égale à la SFDR et ce à la fois en mode dual-tone et single-tone. Par conséquent, les deux méthodes

sont équivalentes en terme de couverture de test. Nous reviendrons ultérieurement sur les avantages qu'offre l'analyse temporelle.

## 4. Définition temporelle des paramètres dynamiques

Après avoir décrit qualitativement les différents paramètres d'erreurs, qu'ils soient statiques ou dynamiques, il est important de donner les relations mathématiques permettant de les quantifier ou du moins de les estimer. Nous distinguerons deux types de formulation selon que le test s'effectue avec une sinusoïde à l'entrée du CAN, mode 'single-tone', ou avec deux sinusoïdes, mode 'dual-tone'.

### 4.1. Mode single-tone

Les paramètres THD, SNR, SINAD, ENOB sont calculés en mode single-tone où le signal d'entrée est composé d'une seule sinusoïde:

$$e(t) = A \cos(2\pi f_e t + \varphi) + C \quad (1-1)$$

où  $A$ ,  $f_e$ ,  $\varphi$  et  $C$  sont respectivement l'amplitude, la fréquence, la phase et l'offset. En supposant que ce signal soit échantillonné à la fréquence  $f_h$ , la  $n^{\text{ème}}$  donnée à la sortie du CAN sera donnée par:

$$y_n = e_n + b_n + h_n \quad 0 \leq n \leq M-1$$

$$e_n = e\left(\frac{n}{f_h}\right) = A \cos(n\omega + \varphi) + C \quad (1-2)$$

où  $b_n$  et  $h_n$  représentent respectivement la contribution du bruit et de la distorsion,  $M$  est la taille de l'acquisition, et  $\omega = \frac{2\pi f_e}{f_h}$  est la pulsation normalisée. La distorsion harmonique peut s'écrire comme étant la somme de  $P$  raies harmoniques:

$$h_n = \sum_{i=2}^P A_i \cos(i \times n \times \omega + \varphi_i) \quad 0 \leq n \leq M-1 \quad (1-3)$$

En pratique seule les dix premières harmoniques sont significatives ( $P = 10$ ). Après ces notations, nous pouvons maintenant formuler les définitions mathématiques des paramètres dynamiques par rapport à l'analyse temporelle.

Le SINAD est défini comme étant le rapport entre la puissance du signal de test et la puissance du bruit plus la distorsion, le tout exprimé en dB:

$$\text{SINAD [dB]} = 10 \log \left( \frac{MA^2}{2 \sum_{n=0}^{M-1} (y_n - e_n)^2} \right) \quad (1-4)$$

La puissance du bruit plus distorsion est calculée à partir de l'équation (1-2). Le SINAD est une très bonne indication de la performance globale du CAN en fonction de la fréquence de test  $f_e$  dans la mesure où il tient compte, et du bruit, et de la distorsion. Le SINAD est aussi utilisé pour définir la bande passante effective d'un CAN, qui correspond à la fréquence de test où le SINAD se dégrade de 3dB. La courbe donnant le SINAD en fonction de la fréquence de test permet aussi d'évaluer les performances du composant pour des applications de sous-échantillonnage (la fréquence d'entrée est supérieure à la fréquence de Nyquist  $f_h/2$ ).

L'ENOB est un paramètre équivalent au SINAD et il est donné par:

$$\text{ENOB [Bits]} = N - 0.5 \times \log_2 \left( \frac{12 \sum_{n=0}^{M-1} (y_n - e_n)^2}{M} \right) \quad (1-5)$$

où  $N$  est la résolution du CAN. La variable de la fonction  $\log_2$  correspondant au rapport de la puissance du bruit plus distorsion à la puissance du bruit de quantification qui est égale à  $1/12$  LSB (Least Significant Bit). Le LSB est défini par le rapport entre la dynamique du CAN en Volts et son nombre de codes qui est égal à  $2^N$ . Par exemple, pour un convertisseur 12 bits avec une dynamique de 2 Volts, le LSB est égal à  $488.28 \mu\text{V}$ .

Le SNR est donné par le rapport entre la puissance du signal de test et la puissance du bruit (cf. éq. (1-2)):

$$\text{SNR [dB]} = 10 \log \left( \frac{MA^2}{2 \sum_{n=0}^{M-1} (y_n - e_n - h_n)^2} \right) \quad (1-6)$$

La courbe donnant le SNR en fonction de la fréquence de test se dégradera en hautes fréquences mais pas aussi rapidement que le SINAD à cause de l'exclusion des harmoniques. Le SNR, comme d'ailleurs tous les autres paramètres, dépend non seulement du CAN, mais aussi des performances du banc de test, la carte de test incluse. Le SNR peut aussi s'exprimer en fonction des différentes composantes du bruit (jitter, DNL, bruit thermique) par [Brannon]:

$$\text{SNR [dB]} = -10 \log \left( (2\pi f_e \sigma_j)^2 + \left( \frac{1 + \text{DNL}_m}{2^N} \right)^2 + \left( \frac{V_T}{2^N} \right)^2 \right) \quad (1-7)$$

avec:

$\sigma_j$  : l'écart type du jitter

$\text{DNL}_m$  : la valeur moyenne de la DNL

$V_T$  : la racine carrée de la valeur quadratique moyenne du bruit thermique

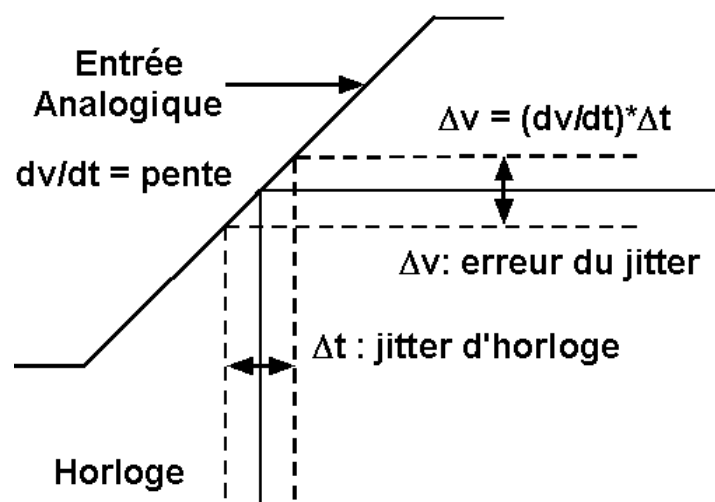


Figure 1-5: L'effet du jitter.

On voit bien l'influence du jitter sur le SNR, notamment en haute fréquence de test ou d'application (cf. figure (1-5)). La pureté de l'horloge, dans ce cas, est d'une

extrême importance. Bien que des générateurs d'horloge ayant un très faible jitter (< 1ps) existent sur le marché, leur coût est relativement élevé. Une technique permettant de surmonter la limitation de l'instrumentation en terme de jitter a été développée moyennant une plus grande complexité de mise en œuvre et une légère augmentation du temps de test [Cauvet et al., 2001].

Il faut noter aussi que le jitter est un bruit large bande de distribution normale, c'est pourquoi dans le domaine spectral il se manifeste par un rehaussement du plancher de bruit (cf. figure (1-6)).

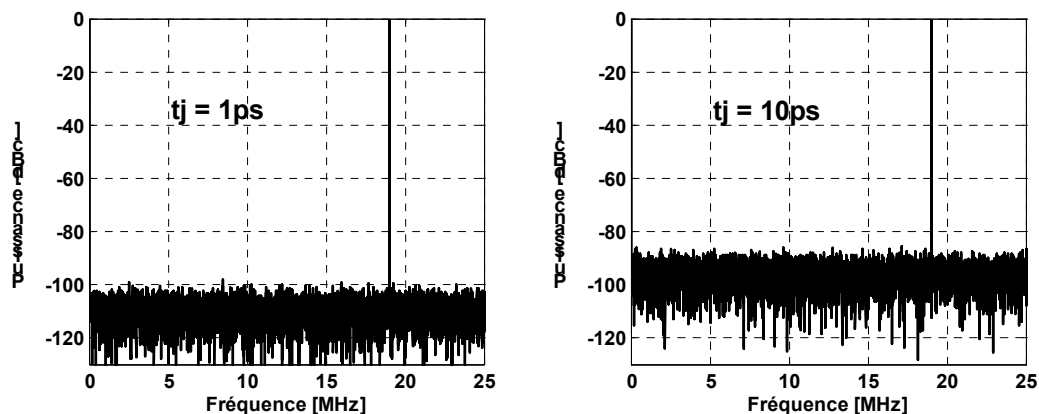


Figure 1-6: Influence du jitter dans le domaine spectral pour un CAN idéal.

La distorsion harmonique est égale au rapport entre la puissance des raies harmoniques et la puissance du signal de test :

$$\text{THD [dB]} = 10 \log \left( \frac{\sum_{i=2}^P A_i^2}{A^2} \right) \quad (1-8)$$

La THD, comme d'ailleurs les autres paramètres, est généralement spécifiée pour un signal d'entrée proche de la pleine échelle, généralement 0.5 à 1dB par rapport à la pleine échelle du CAN (par la suite nous utiliserons la notation dBFS pour référencer une grandeur par rapport à la pleine échelle). La majorité des CANs actuels sont conçus de sorte que la DNL soit répartie sur toute la dynamique. Ainsi, pour les signaux qui sont à quelques dB de la pleine échelle, la distorsion est due essentiellement à l'INL de la fonction de transfert. En revanche, pour les petits signaux, c'est la distorsion parasite (non-harmonique) générée par la DNL



qui domine. Cette dernière ne décroît pas forcément proportionnellement à l'amplitude du signal de test. Cela explique pourquoi, pour les signaux de niveaux élevés, le DFDR est égal au SFDR, alors que, pour les petits signaux, le SFDR est généralement limité par des raies parasites dues à la DNL.

Le DFDR est donné par le rapport entre la puissance de la plus grande raie harmonique et la puissance du signal de test:

$$\text{DFDR [dBc]} = 10 \log \left( \frac{\max \{A_i^2\}_{2 \leq i \leq P}}{A^2} \right) \quad (1-9)$$

L'unité dBc signifie que la référence est la porteuse. Le DFDR peut aussi être exprimé relativement à la pleine échelle (PE =  $2^N$ ) en remplaçant dans l'équation (1-9) l'unité par dBFS et A par PE/2. Un DFDR élevé augmente la capacité d'un récepteur à capter des petits signaux en présence de grands signaux et éviter qu'ils soient masqués par les produits de distorsion.

Finalement, l'amplitude du signal de test est généralement exprimée par rapport à la pleine échelle en utilisant le paramètre suivant:

$$\text{SFSR [dBFS]} = 20 \log \left( \frac{2A}{\text{PE}} \right) \quad (1-10)$$

#### 4.2. Mode dual-tone

En mode dual-tone le signal de test consiste en la combinaison de deux sinusoïdes de fréquences différentes:

$$e(t) = A_1 \cos(2\pi f_{e1} t + \varphi_1) + A_2 \cos(2\pi f_{e2} t + \varphi_2) + C \quad (1-11)$$

La réponse du CAN à ce stimulus peut s'écrire:

$$\begin{aligned} y_n &= e_n + b_n + h_n + d_n \quad 0 \leq n \leq M-1 \\ e_n &= e \left( \frac{n}{f_h} \right) = A_1 \cos(n\omega_1 + \varphi_1) + A_2 \cos(n\omega_2 + \varphi_2) + C \\ \omega_1 &= \frac{2\pi f_{e1}}{f_h}, \quad \omega_2 = \frac{2\pi f_{e2}}{f_h} \end{aligned} \quad (1-12)$$

où  $b_n$  et  $h_n$  représentent le bruit et la distorsion harmonique et  $d_n$  la contribution de la distorsion d'intermodulation. Les produits d'intermodulation sont donnés par l'expression suivante:

$$d_n = \sum_{\substack{2 \leq i+j \leq q \\ i>0, j>0}} \{A_{ij}^+ \cos(n\omega_{ij}^+ + \varphi_{ij}^+) + A_{ij}^- \cos(n\omega_{ij}^- + \varphi_{ij}^-)\} \quad (1-13)$$

avec  $\omega_{ij}^\pm = |i\omega_1 \pm j\omega_2|$

Les paramètres  $\omega_{ij}^\pm$  représentent les fréquences des raies d'intermodulations. Le nombre entier  $q$  est l'ordre le plus élevé des produits d'intermodulation qui en pratique est limité à 5.

En mode dual-tone deux paramètres dynamiques peuvent être calculés par l'analyse temporelle: l'IMD et le DFDR. L'IMD est définie comme étant le rapport entre la puissance des raies d'intermodulation et la puissance du signal de test:

$$\text{IMD [dB]} = 10 \log \left( \frac{\sum_{\substack{2 \leq i+j \leq q \\ i>0, j>0}} (A_{ij}^{+2} + A_{ij}^{-2})}{A_1^2 + A_2^2} \right) \quad (1-14)$$

Quant au DFDR, il est défini comme étant la puissance de la plus grande raie d'intermodulation exprimée relativement à la pleine échelle:

$$\text{DFDR [dBFS]} = 10 \log \left( \frac{\max_{2 \leq i+j \leq q} \{a_{ij}^{+2}, a_{ij}^{-2}\}}{\left(\frac{PE}{2}\right)^2} \right) \quad (1-15)$$

Le DFDR peut aussi s'exprimer en dBc en prenant comme référence la plus grande raie du signal de test.

En pratique, les amplitudes des deux sinusoïdes sont environ à  $-7\text{dBFS}$  pour éviter la saturation dans le cas où elles sont en phases. La différence entre les fréquences des deux sinusoïdes dépend de l'application et de l'information recherchée. Pour les CANs rapides, l'écart de fréquence se situe autour de 1MHz.

## 5. Conclusion

Après avoir décrit les différentes sources d'erreurs et leurs influences sur les paramètres, ce chapitre présente les paramètres clés des CANs en modes single-tone et dual-tone. Les paramètres pouvant être estimés à l'aide de l'analyse temporelle sont décrits avec plus de détails. A partir des équations donnant ces paramètres, on voit que leur calcul exige la connaissance des paramètres du signal de test et des produits de distorsion harmonique et d'intermodulation (fréquences, amplitudes, phases, et offset). Ces derniers peuvent être estimés par des méthodes d'optimisation non linéaires qui feront l'objet des chapitres suivants.

## **CHAPITRE 2**

### **LA METHODE DES MOINDRES CARRES NON LINEAIRES**



## CHAPITRE 2

### LA METHODE DES MOINDRES CARRES NON LINEAIRES

#### 1. Introduction

Nous avons vu dans le premier chapitre que les paramètres dynamiques d'un CAN dépendent uniquement des paramètres (fréquences, amplitudes, phases, offset) du signal de test et des produits de distorsion harmonique et d'intermodulation. Dans le domaine temporel, ces paramètres sont estimés par la méthode des moindres carrés. Dans ce chapitre, il s'agit d'abord de rappeler le principe de cette méthode ainsi que les critères permettant de déterminer ses performances, à savoir le biais et les bornes de Cramer-Rao. Ensuite, nous décrivons un certain nombre d'algorithmes permettant de résoudre ce problème d'estimation non linéaire. La complexité de ce problème est grandement simplifiée par la méthode de séparation des variables, qui consiste à séparer l'estimation des paramètres linéaires des paramètres non linéaires du modèle. Enfin, les algorithmes étant itératifs, le problème des valeurs initiales ainsi que le critère d'arrêt seront également abordés.

#### 2. L'estimateur des moindres carrés

Les observations  $y_n$  contiennent le signal utile (le modèle)  $s_n$ , à savoir le signal de test et les produits de distorsion, et le bruit dû au composant sous test et à

l'instrumentation de test. Le résiduel, défini comme étant l'écart entre le modèle et l'observation, sera noté:

$$r_n(\mathbf{x}) = y_n - s_n(\mathbf{x}) \quad 0 \leq n \leq M-1 \quad (2-1)$$

où  $\mathbf{x}$  est un vecteur contenant les paramètres du modèle. On suppose que les valeurs  $r_n$  sont indépendantes et identiquement distribuées avec une variance  $\sigma^2$  et une densité de probabilité normale. Sous ces hypothèses, la vraisemblance d'un ensemble d'observations  $\mathbf{y} = [y_0, y_1, \dots, y_{M-1}]^T$ , connaissant les paramètres  $\mathbf{x}$  du modèle, est donnée par la fonction [Nocedal et al., 1999, pp. 255]:

$$p(\mathbf{y}; \mathbf{x}, \sigma) = (2\pi\sigma^2)^{-M/2} e^{-\frac{1}{2\sigma^2} \sum_{n=0}^{M-1} r_n^2} \quad (2-2)$$

Pour une variance fixe, il est évident que  $p(\mathbf{y}; \mathbf{x}, \sigma)$  est maximale quand l'écart quadratique défini par la relation suivante:

$$r(\mathbf{x}) = \sum_{n=0}^{M-1} r_n^2 \quad (2-3)$$

est minimale.

Ainsi l'estimateur du maximum de vraisemblance est identique à celui des moindres carrés. Cependant, lorsque les hypothèses ci-dessus ne sont pas remplies les performances des deux estimateurs sont différentes. L'estimateur des moindres carrés est utilisé lorsqu'on ne dispose d'aucune connaissance probabiliste sur les perturbations agissant sur les mesures.

Les hypothèses ci-dessus sont approximativement satisfaites pour les CANs. La densité spectrale de puissance du bruit est constante (cf. figure (1-4)), donc le bruit est blanc. Le bruit additif étant dû à des sources diverses, d'après le théorème central limite [Papoulis, 1991, pp. 214], sa distribution tend vers une distribution normale. Ce fait est illustré sur la figure (2-1), où nous avons représenté le résiduel d'un CAN 12 bits réel ainsi que sa densité de probabilité (PDF). Cette dernière est une gaussienne de moyenne nulle et d'écart type égal à la racine carrée de la valeur quadratique moyenne (RMS) du résiduel:

$$\sigma = \sqrt{\frac{1}{M} \sum_{n=0}^{M-1} r_n^2} = 0.5\text{LSB}. \text{ Les données ont été obtenues avec une fréquence}$$

d'échantillonnage de 20MHz, une fréquence de test de 10.5MHz et un SFSR de

-1dBFS. La méthode de calcul de la densité de probabilité est présentée en annexe (A.2). La propriété d'indépendance découle du fait qu'un bruit blanc et de distribution normale est forcément indépendant.

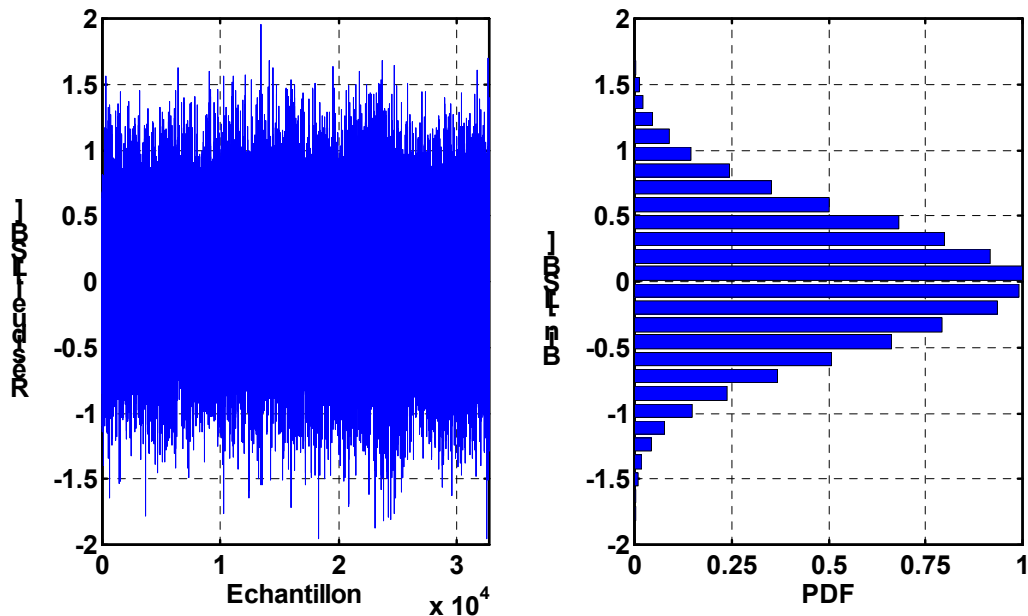


Figure 2-1: Résiduel d'un CAN 12 bits réel avec sa densité de probabilité.

### 2.1. Les Bornes de Cramer-Rao

Les performances d'un estimateur sont caractérisées par son biais et sa variance. Le biais mesure la différence entre la moyenne des réalisations de l'estimateur et la véritable valeur du paramètre à estimer. La variance mesure les fluctuations autour de la valeur moyenne. En ce qui concerne le biais, la borne minimale à atteindre est zéro. Quant à la variance, la borne minimale la plus connue est celle dite de Cramer-Rao. L'estimateur qui atteint cette borne est dit efficace. On dira aussi qu'il est asymptotiquement efficace s'il est efficace lorsque le nombre de mesures tend vers l'infini.

Soit  $\mathbf{F}(\mathbf{x})$  la matrice dont le coefficient  $(i, j)$  est donné par la relation suivante:



$$F_{i,j}(\mathbf{x}) = -E \left[ \frac{\partial^2 \ln(p(\mathbf{y}; \mathbf{x}))}{\partial x_i \partial x_j} \right] \quad (2-4)$$

où  $E$  désigne l'opérateur espérance mathématique qui se fait par rapport au mesures  $\mathbf{y}$ . La matrice  $\mathbf{F}(\mathbf{x})$  s'appelle la matrice d'information de Fisher. Soit  $(\mathbf{F}(\mathbf{x}))^{-1}[i, i]$  l'élément  $(i, j)$  de l'inverse de la matrice de Fisher. Alors, la variance  $V(\hat{x}_i)$  de tout estimateur  $\hat{x}_i$  est telle que [Dufflos et al., 2000, pp. 27]:

$$V(\hat{x}_i) \geq (\mathbf{F}(\mathbf{x}))^{-1}[i, i] \quad (2-5)$$

où  $x_i$  est la  $i^{\text{ème}}$  composante du vecteur  $\mathbf{x}$ . Ainsi, les éléments diagonaux de l'inverse de la matrice de Fisher définissent les bornes de Cramer-Rao. Les éléments de la matrice d'information de Fisher sont donnés par la relation suivante [Rife et al., 1976]:

$$F_{i,j}(\mathbf{x}) = \frac{1}{\sigma^2} \sum_{n=0}^{M-1} \frac{\partial s_n}{\partial x_i} \frac{\partial s_n}{\partial x_j} \quad (2-6)$$

L'étude des bornes de Cramer-Rao pour un signal multi-tone général a été faite dans [Rife et al., 1976]. Dans [Nehorai et al., 1986] et [James et al., 1994], le cas particulier d'un signal multi-harmonique, respectivement réel et complexe, a été traité, et des formules analytiques asymptotiques simples ont été dérivées. Cependant, le modèle traité supposait un offset nul, ce qui n'est pas toujours le cas pour notre problème. Ce cas est traité dans [Pintelon et al., 1996] avec cependant une formulation matricielle lourde à implanter. Par ailleurs, le cas d'un modèle dual-tone n'a, à notre connaissance, jamais été traité. En suivant la démonstration donnée dans [Nehorai et al., 1986] et [James et al., 1994], les expressions asymptotiques des bornes de Cramer-Rao d'un signal multi-harmonique avec offset sont dérivées en annexe (A.1.1). L'annexe (A.1.2) dérive les bornes de Cramer-Rao en mode dual-tone.

## 2. Algorithmes pour des problèmes de moindres carrés non linéaires

Le modèle de notre problème consiste en la combinaison d'un certain nombre de sinusoides. La fonction de coût (2-3) est donc non linéaire par rapport aux fréquences des sinusoides, mais linéaire par rapport aux amplitudes et à l'offset. La minimisation de cette fonction fait appel à des algorithmes itératives qui souffrent en général de l'existence de minima locaux et posent le problème de la vitesse de convergence. Notre objectif fut donc d'étudier et d'apporter des solutions à ces deux problèmes.

Le principe de ces algorithmes est le suivant. Soit  $\mathbf{x}_k$  l'estimateur de la  $k^{\text{ème}}$  itération. L'estimateur de l'itération suivante est donnée par:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \quad (2-7)$$

où  $\mathbf{p}_k$  est la direction de recherche et  $\alpha_k$  un scalaire positif représentant la longueur du pas. Ce scalaire permet d'améliorer la convergence de l'algorithme. A chaque itération, on estime tout d'abord la direction en supposant une longueur de 1, puis on estime la longueur par la résolution d'un problème unidimensionnel.

Pour estimer la direction, la méthode de Newton modélise  $r(\mathbf{x}_k + \mathbf{p}_k)$  par une fonction de second ordre en utilisant le développement en série de Taylor [Dennis et al., 1996]:

$$r(\mathbf{x}_k + \mathbf{p}_k) \approx r(\mathbf{x}_k) + \mathbf{p}_k^T \nabla r(\mathbf{x}_k) + \frac{1}{2} \mathbf{p}_k^T \nabla^2 r(\mathbf{x}_k) \mathbf{p}_k \quad (2-8)$$

où  $\nabla r$  et  $\nabla^2 r$  sont le gradient et le hessien de la fonction de coût. La direction de Newton qui minimise (2-8) est donnée par:

$$\nabla^2 r(\mathbf{x}_k) \mathbf{p}_k = -\nabla r(\mathbf{x}_k) \quad (2-9)$$

L'inversion de ce système linéaire permet de calculer la direction  $\mathbf{p}_k$  à chaque itération.

Une fois la direction calculée, la longueur du pas est obtenue en minimisant la fonction à une seule variable  $r(\mathbf{x}_k + \alpha_k \mathbf{p}_k)$  par rapport à  $\alpha_k$ , via une méthode dite "Line-search" [Nocedal et al., 1999] décrite en annexe (A.3).

On peut montrer que le gradient et le hessien de la fonction de coût sont donnés par les expressions suivantes:

$$\begin{aligned}\nabla r(\mathbf{x}) &= \mathbf{J}(\mathbf{x})^T \mathbf{r}(\mathbf{x}) \\ \nabla^2 r(\mathbf{x}) &= \mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x}) + \sum_{n=0}^{M-1} r_n \nabla^2 r_n\end{aligned}\quad (2-10)$$

où  $\mathbf{r} = [r_0 \cdots r_{M-1}]^T$ ,  $\mathbf{J}$  est la matrice jacobienne des  $r_i$ :

$$J_{ij} = \begin{cases} \frac{\partial r_{i-1}}{\partial x_j} \end{cases} \quad 1 \leq i \leq M, \quad 1 \leq j \leq L \quad (2-11)$$

,  $\nabla^2 r_n$  est le Hessian de  $r_n$  et  $L$  est le nombre de variable de  $\mathbf{x}$ .

D'après (2-10), la méthode de Newton souffre de la difficulté que présente le calcul des  $M$  hessian  $\nabla^2 r_n$  à chaque itération. Pour remédier à ce problème, le hessian de la fonction de coût est approché par son premier terme, soit:

$$\nabla^2 r(\mathbf{x}) \approx \mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x}) \quad (2-12)$$

La méthode qui en résulte est celle de Gauss-Newton [Dennis et al., 1996]. La méthode de Levenberg-Marquardt [Levenberg, 1944] [Marquardt, 1963] est une alternative quand l'approximation de Gauss-Newton (2-12) est imprécise et/ou quand la matrice est singulière. Cette méthode remplace (2-12) par:

$$\nabla^2 r(\mathbf{x}) \approx \mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x}) + \lambda \mathbf{I} \quad (2-13)$$

où  $\mathbf{I}$  est la matrice identité et  $\lambda$  un paramètre scalaire positif choisi à chaque itération par une technique décrite dans [Dennis et al., 1996].

---

<sup>1</sup> L'élément  $(i, j)$  du hessian d'une fonction  $f(\mathbf{x})$  est donné par  $\frac{\partial^2 f}{\partial x_i \partial x_j}$ .

### 3. Méthode de séparation des variables

Notre problème tel qu'il a été présenté jusqu'à présent est plutôt difficile à résoudre en pratique, et ce à cause du grand nombre de paramètres à estimer. Cela est particulièrement vrai en mode dual-tone. Cependant, le problème peut se simplifier considérablement en remarquant qu'une sinusoïde peut s'écrire sous la forme:

$$\begin{aligned} A \cos(\omega t + \varphi) &= a \cos(\omega t) + b \sin(\omega t) \\ \text{avec } a &= A \cos(\varphi) \text{ et } b = -A \sin(\varphi) \end{aligned} \quad (2-14)$$

Par conséquent, le vecteur  $\mathbf{x}$  des paramètres à estimer peut se décomposer en deux parties:

$$\mathbf{x} = [\boldsymbol{\Omega}, \boldsymbol{\theta}] \quad (2-15)$$

Le vecteur  $\boldsymbol{\Omega}$  contient les paramètres non linéaires, à savoir les fréquences angulaires normalisées des signaux de test, et  $\boldsymbol{\theta}$  est constitué des paramètres linéaires du modèle à estimer ( $a$  et  $b$ ). Notons qu'en mode single-tone,  $\boldsymbol{\Omega}$  se réduit à un scalaire. En utilisant une formulation matricielle, on peut montrer que le modèle peut s'écrire de la façon suivante:

$$\mathbf{s}(\mathbf{x}) = \mathbf{H}(\boldsymbol{\Omega})\boldsymbol{\theta} \quad \text{avec } \mathbf{s} = [s_0 \cdots s_{M-1}]^T \quad (2-16)$$

où  $\mathbf{H}$  est une matrice dite matrice du signal. La fonction de coût (2-3) peut se réécrire:

$$r(\boldsymbol{\Omega}, \boldsymbol{\theta}) = (\mathbf{y} - \mathbf{H}(\boldsymbol{\Omega})\boldsymbol{\theta})^T (\mathbf{y} - \mathbf{H}(\boldsymbol{\Omega})\boldsymbol{\theta}) \quad (2-17)$$

Il est montré dans [Golub et al., 1973] que la minimisation de (2-17) peut se faire en deux étapes:

**1<sup>ère</sup> étape:** un estimé  $\hat{\boldsymbol{\Omega}}$  de  $\boldsymbol{\Omega}$  est obtenu en minimisant la fonction de coût modifiée:

$$r(\boldsymbol{\Omega}) = [(\mathbf{I} - \mathbf{H}(\boldsymbol{\Omega})\mathbf{H}^+(\boldsymbol{\Omega}))\mathbf{y}]^T [(\mathbf{I} - \mathbf{H}(\boldsymbol{\Omega})\mathbf{H}^+(\boldsymbol{\Omega}))\mathbf{y}] \quad (2-18)$$

où  $\mathbf{H}^+$  est le pseudo-inverse de  $\mathbf{H}$ . En général  $\mathbf{H}^+$  est calculé par une méthode numériquement stable telle que la décomposition en valeurs singulières (SVD) [Press et al., 1992]. Lorsque  $\mathbf{H}^T\mathbf{H}$  est inversible, son pseudo-inverse est donné par l'expression suivante:

$$\mathbf{H}^+ = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T \quad (2-19)$$

**2<sup>ème</sup> étape:** une fois  $\hat{\Omega}$  obtenue, l'estimé  $\hat{\theta}$  des paramètres linéaires est donné par:

$$\hat{\theta} = \mathbf{H}^+(\hat{\Omega})\mathbf{y} \quad (2-20)$$

Lorsque  $\mathbf{H}^T\mathbf{H}$  est inversible, le système linéaire (2-20) pourrait être résolu plus rapidement à l'aide de la décomposition de Cholesky [**Press et al., 1992**], notamment quand le nombre d'échantillons devient important.

Les algorithmes décrits dans le paragraphe précédent s'appliquent bien sûr à la minimisation de (2-18), et ce en remplaçant  $\mathbf{x}$  par  $\Omega$  et en notant que le résiduel est donné par:  $\mathbf{r}(\Omega) = (\mathbf{I} - \mathbf{H}(\Omega)\mathbf{H}^+(\Omega))\mathbf{y}$ .

#### 4. Critère d'arrêt et valeurs initiales

Le processus itératif (2-7) nécessite des valeurs initiales  $\mathbf{x}_0$  pour l'amorcer et un critère pour l'arrêter. D'après notre expérience, le critère le plus simple et le plus pratique consiste à arrêter l'algorithme lorsqu'une des conditions suivantes est remplie:

1. Le nombre d'itérations  $k$  excède une valeur maximale  $K_{\max}$ . Cela évite une boucle infinie quand l'algorithme diverge à cause par exemple de la non fonctionnalité du composant sous test. D'après la pratique, on peut prendre  $K_{\max} = 50$ . En effet, nous avons constaté que l'algorithme converge en un faible nombre d'itérations, notamment si l'on dispose, comme nous le verrons par la suite, de techniques pour générer des valeurs initiales précises.
2. L'écart entre deux itérations successives est suffisamment proche de zéro:

$$\max_{1 \leq i \leq L} \left| \frac{\mathbf{x}_{k+1}[i] - \mathbf{x}_k[i]}{\mathbf{x}_{k+1}[i]} \right| < (\text{précision machine})^{\frac{2}{3}}.$$

Ce critère donné dans [**Dennis et al., 1996**] traduit le fait qu'il faut tenir compte de la limitation due à la précision de calcul de la machine.

La méthode de séparation des variables décrite ci-dessus simplifie grandement le problème d'estimation des valeurs initiales dans la mesure où il suffit de fournir les valeurs initiales des paramètres non linéaires. Le cas le plus simple est lorsqu'on dispose de générateurs de signaux précis. Il suffit alors d'utiliser les valeurs des fréquences des générateurs pour générer les valeurs initiales. Dans le cas contraire, les méthodes les plus précises et les plus rapides sont sans doute celles basées sur la transformée de Fourier rapide (FFT) et les fenêtres de pondération. Elles seront discutées dans le chapitre suivant.

## 5. Conclusion

Après avoir dérivé le critère des moindres carrés à partir de celui du maximum de vraisemblance, ce chapitre présente un critère fondamental, à savoir les bornes de Cramer-Rao, qui nous permettra d'évaluer les performances des méthodes d'estimation des paramètres dynamique du CAN. Ces méthodes sont essentiellement au nombre de trois: Newton, Gauss-Newton et Levenberg-Marquardt. Leur principe est rappelé et une méthode permettant de réduire la complexité du problème d'estimation a été présentée. Cette méthode, dite méthode de séparation des variables, sépare l'estimation des paramètres linéaires des paramètres non linéaires du modèle. Elle permet également de simplifier le problème d'estimation des valeurs initiales nécessaires aux différentes méthodes itératives. Ces dernières exigent aussi un critère d'arrêt qui a été défini. La fonction de coût, comme nous le verrons, possède des minima locaux et donc des valeurs initiales imprécises conduisent à la divergence des algorithmes. Dans les chapitres suivants, un critère de convergence global sera déterminé et des méthodes permettant de satisfaire ce critère seront également proposées. Les notions fondamentales discutées dans ce chapitre seront exploitées tout au long des chapitres suivants.



**CHAPITRE 3**

**MODE SINGLE-TONE**

**METHODE SANS SEPARATION DES VARIABLES**





## CHAPITRE 3

### MODE SINGLE-TONE: METHODE SANS SEPARATION DES VARIABLES

#### 1. Introduction

Comme nous l'avons vu au premier chapitre, les paramètres dynamiques d'un CAN en mode single-tone nécessitent la connaissance des paramètres du signal de test et des produits de distorsion harmonique. Dans ce chapitre, ces paramètres seront estimés en deux étapes. Dans un premier temps, on estimera les paramètres du signal de test sans séparer l'estimation des paramètres linéaires et non linéaires. C'est la méthode la plus utilisée pour estimer les paramètres SINAD et ENOB [dynad, 2001], [ieeestd1241, 2001], [Leroux, 1989],[Malewski et al., 1987]. Dans un second temps, on estimera les paramètres des raies harmoniques en utilisant les résultats obtenus. Cela nous permettra de calculer les paramètres dynamiques restants (SNR, THD).

Trois méthodes itératives d'estimation des paramètres du signal de test seront étudiées et comparées: Newton, Gauss-Newton et Levenberg-Marquardt. L'incertitude tolérée par chaque méthode sur les valeurs initiales sera déterminée par simulation. Ce résultat servira comme critère pour choisir la méthode d'estimation des valeurs initiales garantissant la convergence globale des différents algorithmes. Ce résultat nous permettra également de développer une méthode robuste, dans le sens où elle converge quelle que soit l'erreur sur les valeurs initiales. L'efficacité de cette méthode sera validée par simulation.

L'estimation des produits harmoniques revient essentiellement à résoudre un système linéaire dont la matrice dépend du nombre d'échantillons et d'harmoniques. Lorsque le nombre de ces paramètres augmente, le temps nécessaire à la résolution de ce système augmente également. Des alternatives seront proposées pour pallier à ce problème.

Enfin, les performances de la méthode présentée seront validées par des simulations de Monte-Carlo et des données réelles. Nous verrons notamment qu'elle atteint les bornes de Cramèr-Rao et qu'elle est en parfait accord avec l'analyse spectrale.

## 2. Estimation des paramètres du signal de test

Afin de simplifier les formules, nous adoptons une notation matricielle. Le modèle (1.2) du signal de test peut s'écrire sous la forme vectorielle suivante:

$$\begin{aligned} \mathbf{s}(\mathbf{x}) &= x_2 \cos(x_1 \mathbf{t}) - x_3 \sin(x_1 \mathbf{t}) + x_4 \cdot \mathbf{1} \\ \mathbf{t} &= [0, \dots, M-1]^T; \quad \mathbf{1} = [1, \dots, 1]^T \\ \mathbf{x} &= [\omega, A \cos(\varphi), A \sin(\varphi), C]^T \end{aligned} \quad (3-1)$$

La méthode sans séparation des variables consiste à estimer simultanément tous les éléments du vecteur  $\mathbf{x}$  par la méthode des moindres carrés décrite dans le deuxième chapitre. La matrice jacobienne du résiduel  $\mathbf{r} = \mathbf{y} - \mathbf{s}$  est donnée dans ce cas par l'expression suivante:

$$\mathbf{J} = [\mathbf{t}(x_2 \sin(x_1 \mathbf{t}) + x_3 \cos(x_1 \mathbf{t})) \quad -\cos(x_1 \mathbf{t}) \quad \sin(x_1 \mathbf{t}) \quad -\mathbf{1}] \quad (3-2)$$

Le premier terme  $\mathbf{J}^T \mathbf{J}$  du hessien de la fonction de coût (cf éq. (2-10)) est donné par la matrice suivante:

$$\begin{bmatrix} \|\mathbf{c}_1\|^2 & -\langle \mathbf{c}_1, \cos(x_1 \mathbf{t}) \rangle & \langle \mathbf{c}_1, \sin(x_1 \mathbf{t}) \rangle & -\langle \mathbf{c}_1, \mathbf{1} \rangle \\ -\langle \mathbf{c}_1, \cos(x_1 \mathbf{t}) \rangle & \|\cos(x_1 \mathbf{t})\|^2 & -\langle \cos(x_1 \mathbf{t}), \sin(x_1 \mathbf{t}) \rangle & \langle \cos(x_1 \mathbf{t}), \mathbf{1} \rangle \\ \langle \mathbf{c}_1, \sin(x_1 \mathbf{t}) \rangle & -\langle \cos(x_1 \mathbf{t}), \sin(x_1 \mathbf{t}) \rangle & \|\sin(x_1 \mathbf{t})\|^2 & -\langle \sin(x_1 \mathbf{t}), \mathbf{1} \rangle \\ -\langle \mathbf{c}_1, \mathbf{1} \rangle & \langle \cos(x_1 \mathbf{t}), \mathbf{1} \rangle & -\langle \sin(x_1 \mathbf{t}), \mathbf{1} \rangle & M \end{bmatrix} \quad (3-3)$$

où  $\langle \cdot, \cdot \rangle$  désigne l'opérateur produit scalaire<sup>2</sup> et  $\mathbf{c}_1 = \mathbf{t}(x_2 \sin(x_1 \mathbf{t}) + x_3 \cos(x_1 \mathbf{t}))$  la première colonne de la matrice  $\mathbf{J}$ . Le deuxième terme  $\sum_{n=0}^{M-1} r_n \nabla^2 r_n$  du hessian est

donné par la matrice suivante:

$$\begin{bmatrix} \langle \mathbf{r}(\mathbf{x}), \mathbf{t}^2(x_2 \cos(x_1 \mathbf{t}) + x_3 \sin(x_1 \mathbf{t})) \rangle & \langle \mathbf{r}(\mathbf{x}), \mathbf{t} \sin(x_1 \mathbf{t}) \rangle & \langle \mathbf{r}(\mathbf{x}), \mathbf{t} \cos(x_1 \mathbf{t}) \rangle & 0 \\ \langle \mathbf{r}(\mathbf{x}), \mathbf{t} \sin(x_1 \mathbf{t}) \rangle & 0 & 0 & 0 \\ \langle \mathbf{r}(\mathbf{x}), \mathbf{t} \cos(x_1 \mathbf{t}) \rangle & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (3-4)$$

Cette matrice est symétrique et singulière car sa dernière colonne est nulle. On voit que la plupart de ses éléments sont nuls, ce qui justifie le fait de la négliger devant la matrice  $\mathbf{J}^T \mathbf{J}$  qui est symétrique et inversible. D'après (2-10), le gradient de la fonction de coût est donné par:

$$\mathbf{J}^T \mathbf{r} = [\langle \mathbf{c}_1, \mathbf{r} \rangle \quad -\langle \cos(x_1 \mathbf{t}), \mathbf{r} \rangle \quad \langle \sin(x_1 \mathbf{t}), \mathbf{r} \rangle \quad -\langle \mathbf{1}, \mathbf{r} \rangle]^T \quad (3-5)$$

Les équations (3-3), (3-4), (3-5) permettent d'implanter les différents algorithmes (Newton, Gauss-Newton, Levenberg-Marquardt) sans avoir à stocker la matrice de grande taille  $\mathbf{J}$ . Le calcul de la matrice (3-3) fait intervenir des sommes ayant les formes suivantes:

$$S_{c,m} = \sum_{n=0}^{M-1} n^m \cos(n\theta); \quad S_{s,m} = \sum_{n=0}^{M-1} n^m \sin(n\theta) \quad (3-6)$$

avec  $m = 0, 1, 2$  et  $\theta = \omega, 2\omega$

L'évaluation de ces sommes est grandement simplifiée par les formules analytiques suivantes:

$$S_{c,0}(\theta) = \frac{\cos\left(\frac{M-1}{2}\theta\right) \sin\left(\frac{M}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)}; \quad S_{s,0}(\theta) = \frac{\sin\left(\frac{M-1}{2}\theta\right) \sin\left(\frac{M}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)} \quad (3-7)$$

<sup>2</sup> Le produit scalaire de deux vecteurs  $\mathbf{x}$  et  $\mathbf{y}$  est donné par:  $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_i x_i y_i$ . La norme d'un vecteur est définie par:  $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ .

$$\left\{ \begin{array}{l} S_{c,1}(\theta) = \frac{M \sin((M-0.5)\theta) - S_{s,0}(\theta) \cos\left(\frac{\theta}{2}\right)}{2 \sin\left(\frac{\theta}{2}\right)} - \frac{S_{c,0}(\theta)}{2} \\ S_{s,1}(\theta) = \frac{-M \cos((M-0.5)\theta) S_{c,0}(\theta) \cos\left(\frac{\theta}{2}\right)}{2 \sin\left(\frac{\theta}{2}\right)} - \frac{S_{s,0}(\theta)}{2} \end{array} \right. \quad (3-8)$$

$$\left\{ \begin{array}{l} S_{c,2}(\theta) = \frac{M(M-1) \sin((M-0.5)\theta) - 2S_{s,1}(\theta) \cos\left(\frac{\theta}{2}\right)}{2 \sin\left(\frac{\theta}{2}\right)} \\ S_{s,2}(\theta) = \frac{-M(M-1) \cos((M-0.5)\theta) + 2S_{c,1}(\theta) \cos\left(\frac{\theta}{2}\right)}{2 \sin\left(\frac{\theta}{2}\right)} \end{array} \right. \quad (3-9)$$

Il faut noter que ces formules ne sont valables que pour  $\theta \neq 2\pi k, k = 0, 1, \dots$ . En pratique, cette condition est toujours satisfaite dans la mesure où la pulsation normalisée est choisie de façon à satisfaire la condition dite de cohérence tel que:

$$\frac{\omega}{2\pi} = \frac{f_e}{f_h} = \frac{J}{M} \quad (3-10)$$

avec  $J$  un entier premier avec  $M$ , qui représente le nombre de périodes de l'acquisition. Si  $M$  est une puissance de deux, alors il suffit de choisir  $J$  impair. Cette condition permet de d'exiter tous les codes d'un CAN.

Un cas intéressant correspond à la situation où la pulsation normalisée  $\omega$  est connue avec une bonne précision. Le vecteur des paramètres à estimer se réduit alors aux paramètres linéaires du modèle:

$$\boldsymbol{\theta} = [A \cos(\varphi), A \sin(\varphi), C]^T \quad (3-11)$$

D'après (2-20), la valeur de  $\boldsymbol{\theta}$  qui minimise la fonction de coût est obtenue en résolvant le système linéaire suivant:

$$(\mathbf{H}^T \mathbf{H}) \hat{\boldsymbol{\theta}} = \mathbf{H}^T \mathbf{y} \quad (3-12)$$

avec la matrice du signal  $\mathbf{H}$  donnée par:

$$\mathbf{H} = [\cos(\omega \mathbf{t}) \quad -\sin(\omega \mathbf{t}) \quad \mathbf{1}] \quad (3-13)$$

La matrice  $\mathbf{H}$  étant de rang plein, alors la matrice  $\mathbf{H}^T \mathbf{H}$  est inversible. A partir de (3-6), (3-12) et (3-13) on obtient:

$$\begin{bmatrix} 0.5(M + S_{c,0}(2\omega)) & -0.5S_{s,0}(2\omega) & S_{c,0}(\omega) \\ -0.5S_{s,0}(2\omega) & 0.5(M - S_{c,0}(2\omega)) & -S_{s,0}(\omega) \\ S_{c,0}(\omega) & -S_{s,0}(\omega) & M \end{bmatrix} \hat{\boldsymbol{\theta}} = \sum_{n=0}^{M-1} \begin{bmatrix} y_n \cos(n\omega) \\ -y_n \sin(n\omega) \\ y_n \end{bmatrix} \quad (3-14)$$

Dans le cas où la condition de cohérence (3-10) est remplie, on peut facilement vérifier à partir de (3-7) que la matrice de ce système linéaire est diagonale et que sa solution est donnée par:

$$\hat{\boldsymbol{\theta}} = \begin{bmatrix} \frac{2}{M} \sum_{n=0}^{M-1} y_n \cos(n\omega) \\ -\frac{2}{M} \sum_{n=0}^{M-1} y_n \sin(n\omega) \\ \frac{1}{M} \sum_{n=0}^{M-1} y_n \end{bmatrix} \quad (3-15)$$

A partir de (3-11) on peut facilement remonter à l'amplitude, la phase et l'offset du signal de test. L'information sur la phase est notamment requise pour calculer la phase différentielle pour les CANs utilisés dans les applications vidéo, et pour évaluer la déviation de phase ainsi le temps d'ouverture [Reeder, 2002] entre deux canaux sous-jacents dans le cas d'un produit intégrant plusieurs CANs sur la même puce, comme c'est de plus en plus le cas en imagerie médicale.

Une fois que les paramètres du signal de test sont estimés, on peut alors calculer le paramètre SINAD à partir de (1-4).

### 2.1. Estimation des valeurs initiales

Les méthodes itératives utilisées pour estimer les paramètres du signal de test dans le cas où les fréquences sont inconnues nécessitent un vecteur initial  $\mathbf{x}_0$  pour démarrer le processus. Celui-ci sera estimé en deux étapes:

1. Une valeur initiale  $\omega_0$  de la pulsation normalisée sera estimée à l'aide d'une méthode décrite ci-après.
2. Les valeurs initiales  $\theta_0$  des paramètres linéaires sont calculées en résolvant le système linéaire (3-14).

La méthode choisie pour estimer  $\omega_0$  doit être suffisamment précise pour assurer la convergence vers le minimum global de la fonction de coût des algorithmes itératifs. L'étude analytique de la précision requise afin d'assurer la convergence globale étant délicate, nous avons procédé par simulation en suivant la méthodologie suivante:

- Le CAN est simulé par le modèle réaliste décrit dans le paragraphe 4 du quatrième chapitre. Ce modèle tient compte de trois sources d'erreur: la distorsion harmonique, bruit de quantification, bruit additif et le jitter. Il est complètement contrôlé par les paramètres suivants:  $N$ ,  $M$ , SFSR, SNR, THD,  $f_e$ ,  $f_h$  et  $\sigma_j$ . Pour refléter la réalité, les valeurs de ces paramètres seront calculées à partir de composants réels.
- L'écart entre la pulsation exacte  $\omega_e$  et  $\omega_0$  peut s'écrire:

$$\Delta\omega_0 = |\omega_0 - \omega_e| = \frac{2\pi\delta}{M} \quad 0 \leq \delta \leq 1 \quad (3-16)$$

Les paramètres du modèle étant fixés, on fait alors varier  $\delta$  par un pas fixe donné. Pour chaque valeur de l'erreur, on estime par un algorithme donné les paramètres du signal de test et on calcule la fonction de coût correspondante. L'analyse de la courbe ainsi obtenue donnera des informations sur la précision  $\delta_c$  requise pour garantir la convergence globale de chaque algorithme. Enfin on étudiera l'effet des différents paramètres du modèle sur  $\delta_c$ .

Lorsqu'on dispose de générateurs précis tel que le HP8644B de chez Agilent, la pulsation normalisée qui en résulte est suffisamment précise pour garantir la convergence globale. Une alternative consiste à utiliser les techniques basées sur

la transformée de Fourier discrète (DFT) et les fenêtres de pondération [Grandke, 1983], [Andria et al., 1989], [Bertocco et al., 1999]. Ces méthodes ont l'avantage d'être rapides et précises. L'implantation suivante de la technique présentée dans [Grandke, 1983] ne nécessite pas le calcul du spectre total des données. Les données  $y_n$  sont pondérées par la fenêtre de Hanning:

$$w_n = 0.5 - 0.5 \cos\left(2\pi \frac{n}{M}\right); \quad 0 \leq n \leq M-1 \quad (3-17)$$

Ensuite le spectre en amplitude des données pondérées est calculé aux points  $J-1, J, J+1$  en utilisant l'expression:

$$Y_k = \sqrt{\left(\sum_{n=0}^{M-1} y_n w_n \cos\left(2\pi \frac{nk}{M}\right)\right)^2 + \left(\sum_{n=0}^{M-1} y_n w_n \sin\left(2\pi \frac{nk}{M}\right)\right)^2} \quad (3-18)$$

La pulsation normalisée ainsi estimée est donnée par:

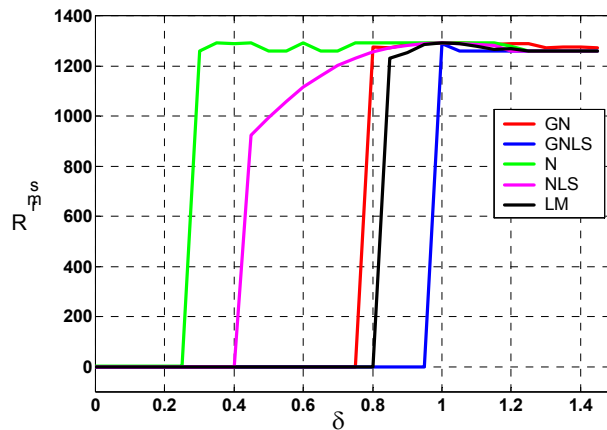
$$\omega_0 = \frac{2\pi}{M} \left( J + \frac{2 \max\{Y_{J+1}, Y_{J-1}\} - Y_J}{Y_J + \max\{Y_{J+1}, Y_{J-1}\}} \right) \quad (3-19)$$

### 2.1.1. Résultats de simulation

Pour la méthodologie d'estimation des valeurs initiales décrite ci-dessus, nous avons étudié l'effet de l'erreur (3-16) sur le résiduel RMS estimé à l'aide des quatre méthodes (cf. chapitre 2): Gauss-Newton (GN), Gauss-Newton avec recherche de  $\alpha_k$  (GNLS), Newton (N), Newton avec recherche de  $\alpha_k$  (NLS), et Levenberg-Marquardt (LM).

Les résultats obtenus sont donnés sur la figure (3-1). Pour chaque méthode, on distingue deux zones: une zone de convergence qui correspond à une erreur inférieure à une certaine valeur critique  $\delta_c$  et à un résiduel faible, et une zone de divergence qui correspond à un grand résiduel. La transition entre les deux zones étant brutale, on peut alors déterminer facilement le  $\delta_c$  de chaque méthode. On voit que la recherche de  $\alpha_k$  améliore significativement la largeur de la zone de convergence. La méthode de Newton est la plus exigeante en terme de précision, alors que celles de Gauss-Newton et Levenberg-Marquardt sont à peu près semblables.





**Figure 3-1:** Influence de l'erreur sur la pulsation initiale sur le résiduel RMS pour différents algorithmes. Paramètres: pas sur  $\delta = 0.05$ ,  $N = 12$ ,  $M = 16384$ ;  $f_e = 10.5\text{MHz}$ ,  $f_h = 50\text{MHz}$ ;  $\text{SFSR} = -1\text{dB}$ ,  $\text{SNR} = 65\text{dB}$ ,  $\text{THD} = -76\text{dB}$ .

Maintenant il faudrait s'assurer que  $\delta_c$  est constant quelque soit les conditions de test (THD, SNR, J). Pour cela nous avons étudié l'effet des paramètres SNR, THD, et le nombre de périodes J sur  $\delta_c$ , et ce pour chaque algorithme. L'effet est identique pour tous les algorithmes, c'est pourquoi nous ne donnons sur la figure (3-2) que les résultats de la méthode de Gauss-Newton. On voit que le SNR et la THD n'ont aucun effet sur  $\delta_c$  et on constate seulement une légère baisse pour les faibles nombres de périodes.

Le pire cas étant obtenu pour une valeur faible de J, nous avons alors tracé la variation du résiduel en fonction de l'erreur pour trois périodes avec une légère non-cohérence. Les résultats sont donnés sur la figure (3-3). A partir de cela, les valeurs données dans le tableau (3-1) représentent une bonne estimation de la précision requise pour chaque méthode sur la pulsation initiale et qui assure la convergence.

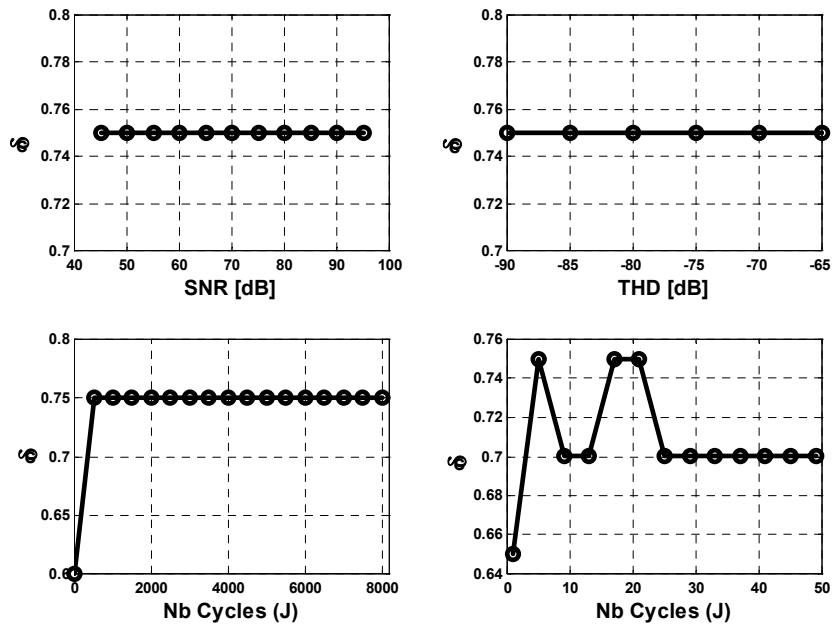


Figure 3-2: Effet du SNR, THD et J sur  $\delta_c$ . Paramètres:  $N = 12$ ;  $M = 16384$ ;  $f_e = 10.5\text{MHz}$ ;  $f_h = 50\text{MHz}$ ;  $SFSR = -1\text{dB}$ ;  $SNR = 65\text{dB}$ ;  $THD = -76\text{dB}$ .

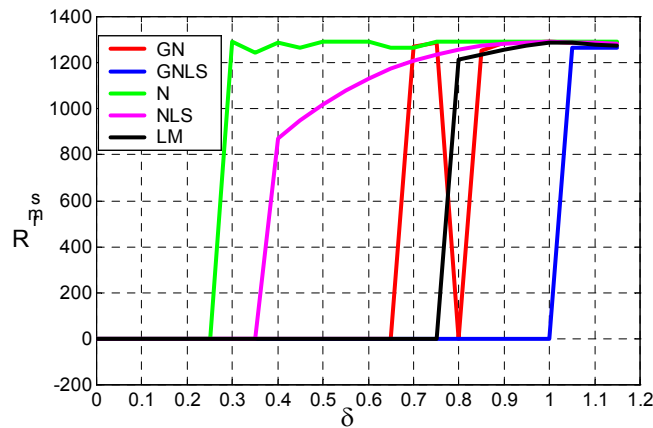


Figure 3-3: Influence de l'erreur sur la pulsation initiale sur le résiduel RMS pour différents algorithmes. Paramètres: pas sur  $\delta = 0.05$ ;  $N = 12$ ;  $M = 16384$ ;  $J = 3.01$ ;  $f_h = 50\text{MHz}$ ;  $SFSR = -1\text{dB}$ ;  $SNR = 65\text{dB}$ ;  $THD = -76\text{dB}$ .

Tableau 3-1: Incertitude tolérée sur la pulsation initiale afin d'assurer la convergence globale des différents algorithmes.

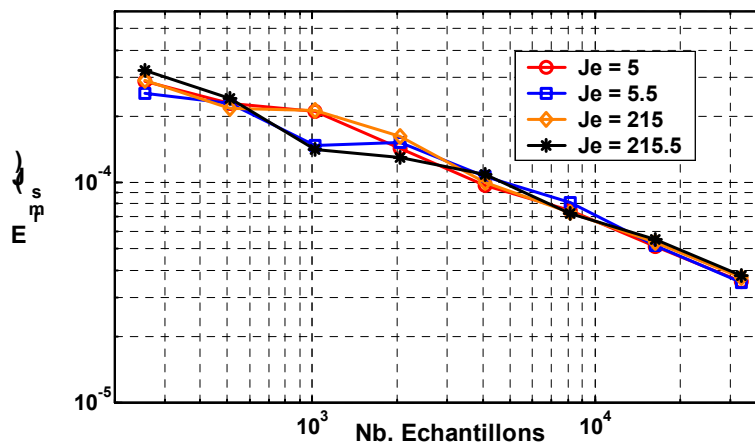
METHODE	$\delta_c$	$\Delta\omega_c$
Gauss-Newton (GN)	0.65	$2\pi \frac{0.60}{M}$

Gauss-Newton avec recherche de $\alpha$ (GNLS)	0.95	$2\pi \frac{0.90}{M}$
Newton (N)	0.25	$2\pi \frac{0.20}{M}$
Newton avec recherche de $\alpha$ (NLS)	0.35	$2\pi \frac{0.30}{M}$
Levenberg-Marquardt (LM)	0.75	$2\pi \frac{0.70}{M}$

Nous avons également étudié la précision de la méthode d'estimation de la pulsation initiale proposée ci-dessus (cf. eq. (3-19)) par des simulations de Monte-Carlo sur 100 réalisations indépendantes. On choisit le nombre de périodes  $J$  de façon à satisfaire ou non la condition de cohérence, ensuite on calcul la valeur RMS de l'erreur d'estimation par rapport à la valeur exacte  $J_e$  du nombre de périodes par l'expression suivante:

$$E_{\text{rms}}(J) = \sqrt{\frac{1}{100} \sum_{i=1}^{100} (\hat{J}_i - J_e)^2} \quad (3-20)$$

où  $\hat{J}_i$  est l'estimation obtenue à la  $i^{\text{ème}}$  réalisation.



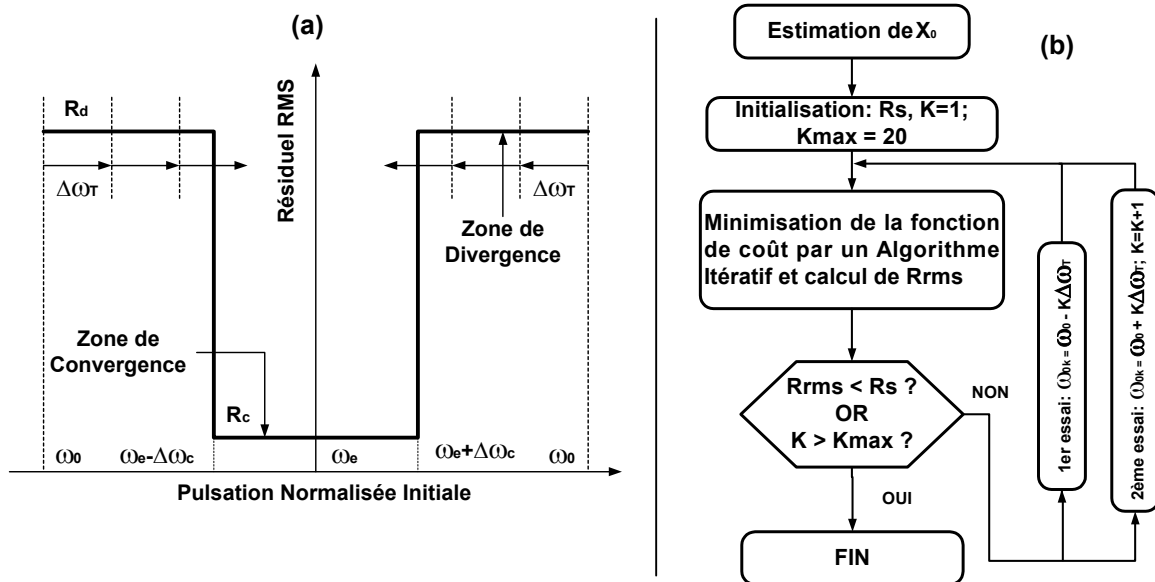
**Figure 3-4** :Précision de la méthode DFT interpolée et pondérée en fonction du nombre de points, de périodes et de la cohérence. Paramètres:  $N = 8$ ;  $f_h = 50\text{MHz}$ ;  $SFSR = -1\text{dB}$ ;  $SNR = 45.5\text{dB}$ ;  $THD = -66.5\text{dB}$ .

Le pire cas correspondant à un CAN 8 bits où le SNR est faible. Les résultats obtenus en fonction du nombre d'échantillons et du nombre de périodes avec et

sans cohérence sont représentés sur la figure (3-4). Notons que le paramètre  $\delta$  de l'équation (3-16) représente en fait l'erreur sur le nombre de périodes. En comparant les résultats du tableau (3-1) et ceux de la figure (3-4), on en déduit que la précision requise par les différentes méthodes est largement satisfaite.

## 2.2. Une méthode globalement convergente

D'après les résultats ainsi obtenues, le résiduel RMS a l'allure représentée à la figure (3-5.a) en fonction de la pulsation normalisée initiale.



**Figure 3-5:** Principe (a) et implantation (b) de la méthode des moindres carrés modifiée.

On distingue deux zones: une zone de convergence où le résiduel RMS est égal  $R_c$ , et une zone de divergence où le résiduel RMS vaut  $R_d$ . Le passage de  $R_c$  à  $R_d$  lorsque l'algorithme diverge est brutale. La largeur de la zone de convergence vaut  $2\Delta\omega_c$ . L'idée de base de l'algorithme modifié est que toute pulsation normalisée initiale située dans la zone de divergence peut être ramenée dans la zone de convergence par toute translation itérative de pas inférieure à  $2\Delta\omega_c$ . Un grand pas permet d'accélérer le processus mais risque de causer des phénomènes de bord. Un bon compromis consiste à choisir un pas égal à:

$$\Delta\omega_T = 1.5\Delta\omega_C \quad (3-21)$$

Il s'agit d'une translation croissante si  $\omega_0 < \omega_e$  et décroissante dans le cas contraire. En pratique, on dispose toujours d'une valeur minimale  $ENOB_s$  pour le nombre effectif de bits. En effet, on tolère rarement une dégradation d' $ENOB$  supérieure à 2 bits. C'est pourquoi pour notre problème on considérera le pire cas, à savoir:  $N - ENOB_s \approx 3$  bits. D'après (1-5), le résiduel RMS correspondant est donné par:

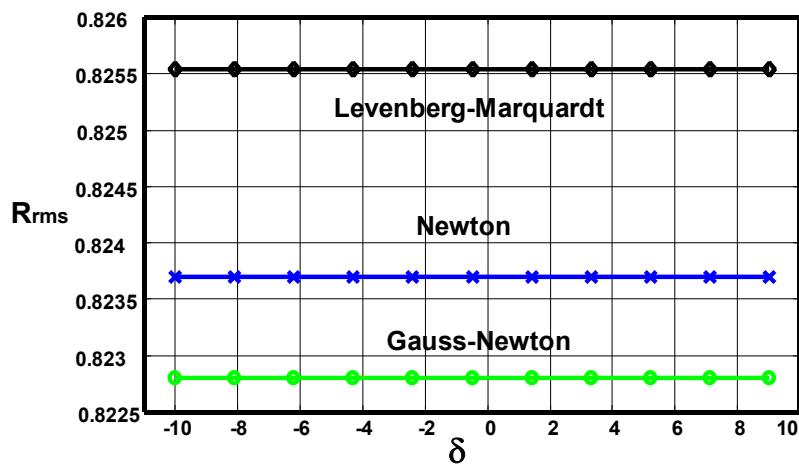
$$R_s = \frac{2^{(N-ENOB_s-1)}}{\sqrt{3}} \quad (3-22)$$

L'implantation de l'algorithme modifié est schématisée à travers la figure (3-5.b). Après l'estimation de la pulsation normalisée initiale, la fonction de coût est minimisée par un algorithme itératif donné (Newton, Gauss-Newton, Levenberg-Marquardt). Les paramètres ainsi estimés sont utilisés pour calculer le résiduel RMS. Tant que le processus diverge ( $R_{RMS} > R_s$ ), la valeur de la pulsation initiale est affinée. La position relative de  $\omega_0$  par rapport à  $\omega_e$  étant inconnue a priori, on applique alors deux translations, une croissante et l'autre décroissante de pas  $\pm k\Delta\omega_T$ , avant de passer à l'itération  $k + 1$ . Notons également que le processus est arrêté au bout de  $k_{max}$  itérations, et ce afin d'éviter une boucle infinie lorsque par exemple les données sont corrompues ou que le composant n'est pas fonctionnel. La valeur de  $k_{max}$  dépend de l'incertitude sur la valeur initiale et une valeur de 20 itérations est largement suffisante dans la plupart des situations.

Cette méthode permet de se passer complètement du problème d'estimation de la pulsation initiale. En effet, il suffit d'utiliser les valeurs des fréquences des générateurs pour générer la valeur initiale. L'algorithme se charge alors d'affiner cette valeur au cas où elle ne serait pas suffisamment précise pour assurer la convergence globale.

### 2.2.1. Résultats de simulation

La figure (3-6) montre l'effet de l'erreur sur la pulsation initiale sur la convergence de la méthode des moindres carrés modifiée appliquée à trois algorithmes. On voit bien que les différents algorithmes convergent vers le minimum global et ce malgré l'erreur importante sur la pulsation initiale. Par conséquent, la modification proposée ci-dessus est effectivement robuste vis-à-vis de l'erreur sur la pulsation initiale.



**Figure 3-6:** Résiduel RMS en fonction de l'erreur sur la pulsation initiale pour la méthode des moindres carrés modifiée. Paramètres:  $N = 12$ ;  $M = 16384$ ;  $SNR = 64.4\text{dB}$ ;  $THD = -75.5\text{dB}$ ;  $f_h = 50\text{MHz}$ ;  $f_e = 10.5\text{MHz}$ ;  $ENOBs = 9\text{ bits}$ .

## 3. Estimation des Produits Harmoniques

Le calcul des paramètres SNR, THD et DFDR nécessitent la connaissance des paramètres des produits de distorsion harmonique. Les paramètres du signal de test étant estimés, on peut alors calculer le résiduel  $\mathbf{r} = \mathbf{y} - \mathbf{s} = [r_0, \dots, r_{M-1}]^T$  qui est composé des harmoniques  $\mathbf{h} = [h_0, \dots, h_{M-1}]^T$  et du bruit additif  $\mathbf{b} = [b_0, \dots, b_{M-1}]^T$ . Le modèle (1-3) des composantes harmoniques peut s'écrire sous la forme matricielle suivante:

$$\mathbf{h}(\boldsymbol{\theta}) = \sum_{i=2}^P (\mathbf{a}_i \cos(i\omega \mathbf{t}) + \mathbf{b}_i \sin(i\omega \mathbf{t})) = \mathbf{H}(\omega) \boldsymbol{\theta}$$

avec  $\mathbf{a}_i = A_i \cos(\varphi_i)$ ;  $\mathbf{b}_i = -A_i \sin(\varphi_i)$  (3-23)

$$\boldsymbol{\theta} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_P \ \mathbf{b}_1 \ \cdots \ \mathbf{b}_P]^T$$

$$\mathbf{H} = [\cos(2\omega \mathbf{t}) \ \cdots \ \cos(P\omega \mathbf{t}) \ \sin(2\omega \mathbf{t}) \ \cdots \ \sin(P\omega \mathbf{t})]$$

Le modèle étant linéaire par rapport au paramètre  $\boldsymbol{\theta}$ , son estimé par la méthode des moindres carrés est donnée alors par le système linéaire suivant:

$$(\mathbf{H}^T \mathbf{H}) \hat{\boldsymbol{\theta}} = \mathbf{H}^T \mathbf{r} \quad (3-24)$$

La matrice  $\mathbf{H}$  est pratiquement de rang plein et le système (3-24) peut être résolu par la décomposition de Cholesky. Une alternative qui n'impose aucune condition sur le rang de  $\mathbf{H}$  est de résoudre le système surdéterminé suivant:

$$\mathbf{H}\boldsymbol{\theta} = \mathbf{r} \quad (3-25)$$

par la méthode de décomposition en valeurs singulières (SVD). L'inconvénient de cette dernière est le temps nécessaire au calcul du pseudo-inverse de  $\mathbf{H}$ , notamment quand le nombre d'échantillons et/ou d'harmoniques augmente.

Une deuxième approche pour estimer les produits harmoniques consiste à estimer une composante harmonique à la fois au lieu de les estimer toutes simultanément. L'avantage est sa rapidité et sa simplicité d'implantation. En effet, pour estimer l'harmonique d'ordre  $i$ , il suffit de résoudre le système 2x2 défini par la relation suivante:

$$\begin{bmatrix} 0.5(M + S_{c,0}(2i\omega)) & 0.5S_{s,0}(2i\omega) \\ 0.5S_{s,0}(2i\omega) & 0.5(M - S_{c,0}(2i\omega)) \end{bmatrix} \begin{bmatrix} \mathbf{a}_i \\ \mathbf{b}_i \end{bmatrix} = \sum_{n=0}^{M-1} \begin{bmatrix} r_{i,n} \cos(ni\omega) \\ r_{i,n} \sin(ni\omega) \end{bmatrix} \quad (3-26)$$

et déduit de (3-14) en considérant qu'il n'y a pas d'offset à estimer et en remplaçant  $\omega$  par  $i\omega$ . A chaque fois qu'une composante harmonique est estimée, elle est soustraite du résiduel courant pour donner les résiduels  $r_{i,n}$ . Il n'est pas nécessaire de respecter un ordre donné dans l'estimation des harmoniques, mais il faut essayer d'estimer en premier lieu les harmoniques ayant la contribution la plus importante, et ce afin de minimiser les erreurs qu'elles pourraient induire sur l'estimation des harmoniques faibles. Par exemple, pour les CANs rapides

modernes, il faut commencer par les harmoniques impaires car leur structure d'entrée différentielle atténue fortement les harmoniques paires.

#### 4. Simulations de Monte-Carlo et Résultats réels

Nous avons étudié les performances des algorithmes décrits ci-dessus à l'aide de simulations de Monte-Carlo sur 200 réalisations indépendantes. Les résultats donnés par les différentes méthodes (Newton, Gauss-Newton, Levenberg-Marquardt) sont semblables, c'est pourquoi par la suite nous ne donnerons que ceux de la méthode de Gauss-Newton. Le CAN a été simulé par le modèle décrit dans le chapitre suivant (paragraphe 4). Les expressions utilisées pour les bornes de Cramer-Rao sont données dans l'annexe (A.1.1).

Les résultats obtenus en fonction du nombre d'échantillons sont donnés sur la figure (3-7). On voit que la méthode de Gauss-Newton atteint les bornes de Cramèr-Rao pour l'estimation des paramètres du signal de test, à savoir la pulsation normalisée, la phase, l'amplitude et l'offset.

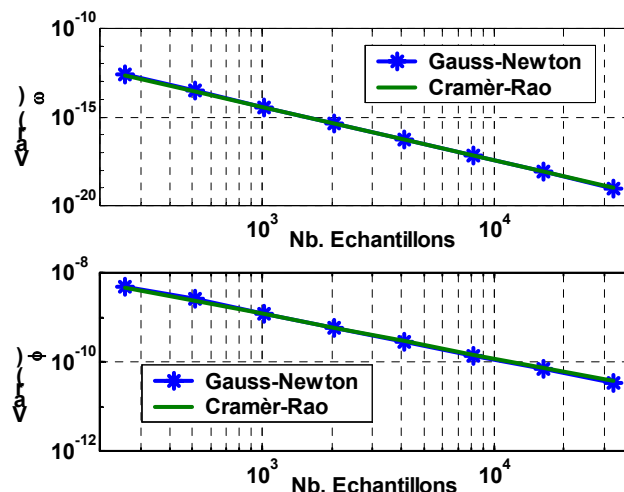
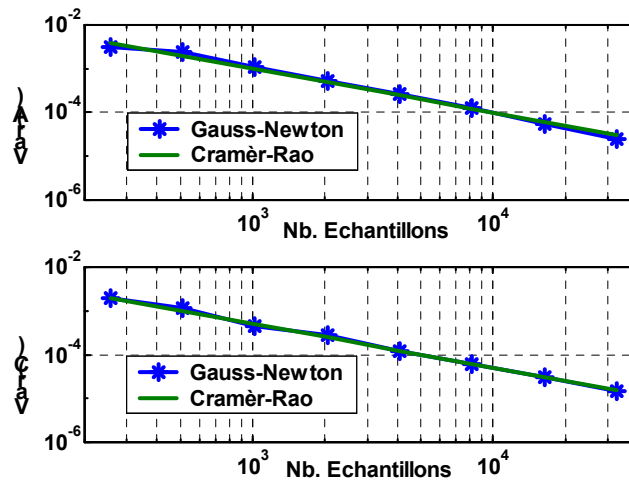


Figure 3-7.a: Comparaison de la méthode de Gauss-Newton avec les bornes de Cramèr-Rao pour l'estimation de la pulsation et de la phase.





**Figure 3-7.b:** Comparaison de la méthode de Gauss-Newton avec les bornes de Cramér-Rao pour l'estimation de l'amplitude et de l'offset.

Nous avons utilisé la méthode décrite dans ce chapitre pour caractériser deux CAN réels différents. Un CAN 12 bits CMOS de faible consommation destiné aux applications d'imagerie médicale, et un CAN 8 bits BiCMOS rapide pour des applications radar ou d'acquisition rapide. Les conditions de test sont données sur le tableau (3-2) et le spectre des données est représenté sur la figure (3-8).

**Tableau 3-2:** Conditions de test et paramètres des CANs.

CAN	fh [MHz]	fe [Hz]	M	N [Bits]
1	50	9898376.46484375	32768	12
2	1000	999862670.8984	65536	8

Pour utiliser l'analyse spectrale comme référence, il faut satisfaire la condition de cohérence avec une bonne précision. Pour cela, nous avons utilisé des générateurs de signaux très précis. Nous avons considéré 5 harmoniques lors du calcul des différents paramètres.

Les résultats obtenus par les différentes méthodes sont résumés dans le tableau (3-3). D'une part, on voit que les différents algorithmes de l'analyse temporelle donnent exactement les mêmes résultats. D'autre part, on constate un très bon

accord entre ces résultats et ceux obtenus par l'analyse spectrale. Par ailleurs, on voit un grand avantage pour l'analyse spectrale en terme de temps de test. Cependant, le temps de test des différents algorithmes de l'analyse temporelle reste parfaitement adapté au test en laboratoire et à l'engineering (phase de debug du test industriel). Ce temps sera fortement réduit en utilisant un matériel plus performant. En effet, le matériel que nous avons utilisé est relativement modeste en terme de performances (PC 400MHz, 32MO de RAM). De plus, l'implantation des algorithmes a été faite sous Matlab qui est un langage interprété pouvant augmenter le temps de calcul.

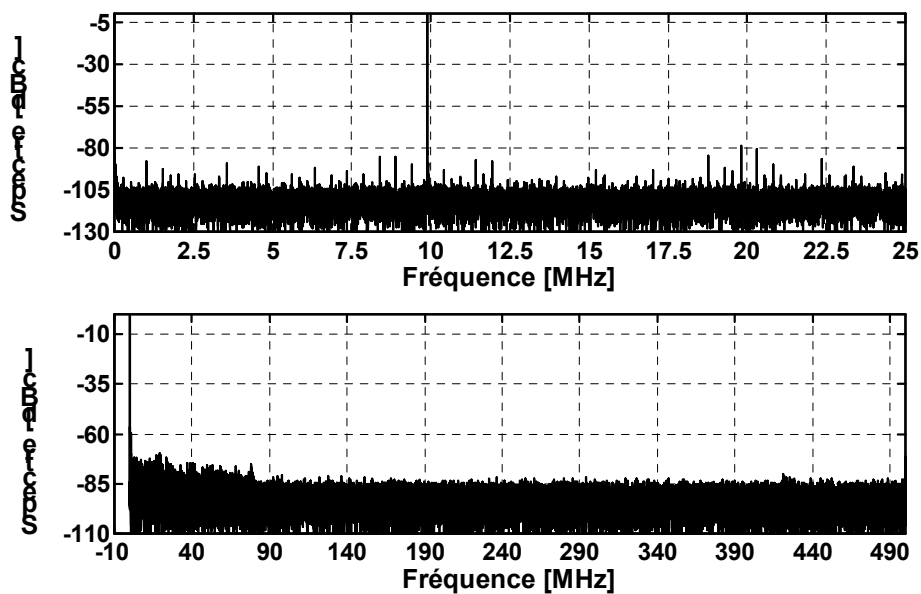


Figure 3-8: Spectre des données. CAN 1: spectre du haut; CAN 2: spectre du bas.

Tableau 3-3 : Résultats des analyses spectrale et temporelle.

CAN	Méthode	SFSR [dBFS]	SINAD [dB]	ENOB [Bits]	SNR [dB]	THD [dB]	DFDR/SFDR [dBFS]	Temps de Test [s]
1	Levenberg	-1.04	65.68		65.60	-76.40	-79.04/ *	2.60
	Newton	-1.04	65.68	10.79	66.06	-76.40	-79.04/ *	3.66
	Gauss	-1.04	65.68	10.79	66.06	-76.40	-79.04/ *	4.66
	FFT	-1.05	65.77	10.80	66.17	-76.40	-79.04/-79.04	0.32

<b>2</b>	<b>Levenberg</b>	-2.85	44.63	7.60	45.09	-54.52	-56.29	9.83
	<b>Newton</b>	-2.85	44.63	7.60	45.09	-54.52	-56.29/*	6.86
	<b>Gauss</b>	-2.85	44.63	7.60	45.09	-54.52	-56.29/ *	7.36
	<b>FFT</b>	-2.88	44.65	7.60	45.12	-54.52	-56.29/-56.29	0.5

## 5. Conclusion

Dans ce chapitre, nous avons étudié les propriétés de la méthode dite à quatre paramètres en terme de précision et de convergence. Nous l'avons aussi étendue à l'estimation des paramètres SNR, THD et DFDR. L'étude de l'effet des valeurs initiales sur la convergence des différents algorithmes nous a permis de développer une méthode globalement convergente dont les performances ont été validées par simulation.

**CHAPITRE 4**

**MODE SINGLE-TONE**

**METHODE DE SEPARATION DES VARIABLES**



## CHAPITRE 4

### MODE SINGLE-TONE: METHODE DE SEPARATION DES VARIABLES

#### 1. Introduction

Le modèle du signal à estimer consiste en la combinaison du signal de test et des harmoniques engendrés par la non linéarité du CAN. Il peut s'écrire sous la forme matricielle suivante:

$$\begin{aligned} \mathbf{s}(\omega, \boldsymbol{\theta}) &= \sum_{i=1}^P (a_i \cos(i\omega \mathbf{t}) + b_i \sin(i\omega \mathbf{t})) + C \cdot \mathbf{1} = \mathbf{H}(\omega) \boldsymbol{\theta} \\ \boldsymbol{\theta} &= [a_1, \dots, a_P, b_1, \dots, b_P, C]^T; a_i = A_i \cos(\varphi_i), b_i = -A_i \sin(\varphi_i) \\ \mathbf{H}(\omega) &= [\cos(\omega \mathbf{t}) \dots \cos(P\omega \mathbf{t}) \quad \sin(\omega \mathbf{t}) \dots \sin(P\omega \mathbf{t}) \quad \mathbf{1}] \\ \mathbf{t} &= [0, \dots, M-1]^T; \mathbf{1} = [1, \dots, 1]^T \end{aligned} \quad (4-1)$$

La fonction de coût à minimiser est alors donnée par:

$$r(\omega, \boldsymbol{\theta}) = (\mathbf{y} - \mathbf{H}(\omega) \boldsymbol{\theta})^T (\mathbf{y} - \mathbf{H}(\omega) \boldsymbol{\theta}) \quad (4-2)$$

La méthode de séparation de variables décrite dans le deuxième chapitre permet de séparer l'estimation des paramètres linéaires  $\boldsymbol{\theta}$  du paramètre non linéaire  $\omega$  du modèle (4-1).

L'objectif de ce chapitre est de montrer comment estimer les paramètres du modèle à l'aide de la méthode de séparation des variables. Le problème d'estimation de la pulsation normalisée sera considérablement simplifié en négligeant l'effet de la distorsion harmonique. L'incertitude tolérée sur la pulsation initiale sera également étudiée. Le résultat de cette étude sera utilisé pour développer une méthode robuste d'estimation de la pulsation normalisée. Enfin,

les performances de cette méthode seront validées à travers des simulations de Monte-Carlo et des données réelles.

## 2. Estimation des paramètres du modèle

### 2.1. Paramètres linéaires

La minimisation de la fonction de coût définie en (4-2) par rapport à  $\theta$  conduit à la minimisation du système linéaire suivant:

$$\hat{\theta} = \mathbf{H}^+(\omega)\mathbf{y} \quad (4-3)$$

Le pseudo-inverse  $\mathbf{H}^+$  peut être calculé à l'aide de la décomposition en valeurs singulières sans aucune information à priori sur le rang de  $\mathbf{H}$ . Cependant, le temps de calcul augmente beaucoup avec le nombre d'échantillons et/ou d'harmoniques. La matrice  $\mathbf{H}$  est pratiquement toujours de rang plein et son pseudo-inverse peut être calculé par:  $\mathbf{H}^+ = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$ . Ainsi, le système (4-3) peut être résolu de manière plus rapide à l'aide de la décomposition de Cholesky.

Pour démontrer ce résultat, supposant que les colonnes  $\mathbf{c}_k$  de  $\mathbf{H}$  soient liées. Autrement dit, ils existent  $2P + 1$  constantes  $\lambda_k$  tel que:

$$\sum_{k=1}^{2P+1} \lambda_k \mathbf{c}_k = \mathbf{0} \Rightarrow \sum_{k=1}^P [\lambda_k \cos(nk\omega) + \lambda_{k+P} \sin(n(k+P)\omega)] + \lambda_{2P+1} = 0; \forall 0 \leq n < M \quad (4-4)$$

Soit  $z_n$  le nombre complexe  $e^{jn\omega}$ . L'expression (4-4) peut se ré-écrire sous la forme suivante:

$$\sum_{k=0}^{2P} \gamma_k z_n^k = 0 \quad \forall 0 \leq n < M$$

$$\gamma_k = \begin{cases} \lambda_{P-k} + j\lambda_{2P-k} & 0 \leq k \leq P-1 \\ \lambda_{2P+1} & k = P \\ \lambda_{k-P} + j\lambda_k & P+1 \leq k \leq 2P \end{cases} \quad (4-5)$$

Cette équation signifie que le polynôme  $Q(z) = \sum_{k=0}^{2P} \gamma_k z^k$  possède  $M$  racines  $z_n$  alors que son degré n'est que de  $2P$ . Il suffit que  $2P + 1$  racines parmi tous les  $z_n$  soient distincts pour que cela soit une contradiction, et par conséquent (4-5) n'est possible que si tous les coefficients  $\gamma_k$  sont nuls. Ce qui veut dire que les colonnes de  $\mathbf{H}$  sont libres et cette matrice est de rang plein.

Maintenant il faut trouver les situations où nous avons  $2P + 1$  différents nombres complexes  $z_n$ . Si  $\omega < \pi/P$  alors  $n\omega < \pi \quad \forall 0 \leq n \leq P$  et donc les  $P + 1$  racines  $\{z_n\}_{0 \leq n \leq P}$  sont distinctes. Le polynôme  $Q(z)$  étant complexe, il admet aussi les racines suivantes:  $\{z_n^*\}_{0 \leq n \leq P} = \{z_{-n}\}_{0 \leq n \leq P}$  qui sont toutes différentes. Ainsi nous avons  $2P + 1$  racines distinctes. Supposons qu'il existe deux racines identiques  $z_{n_1}$  et  $z_{n_2}$  avec  $n_2 > n_1$ . Cela implique:

$$\omega(n_2 - n_1) = 2\pi k \quad k = 1, 2, \dots \quad (4-6)$$

La pulsation normalisée est toujours choisie de façon à satisfaire la condition de cohérence. L'imprécision de l'instrumentation et éventuellement le manque de synchronisation entre les générateurs d'horloge et du signal de test introduisent un écart  $\delta < 1$  par rapport à la condition de cohérence qui se formule par:

$$\omega = 2\pi \frac{J + \delta}{M} \quad (4-7)$$

La combinaison de (4-6) et (4-7) implique:

$$(n_2 - n_1)\delta = kM - J(n_2 - n_1) \quad (4-8)$$

Le terme de gauche de (4-8) est réel alors que celui de droite est entier, ce qui est contradictoire. Ainsi, il n'existe pas deux racines identiques si la condition (4-7) est satisfaite. Par conséquent, du moment où la pulsation normalisée satisfait (4-7) alors la matrice  $\mathbf{H}$  est de rang plein.

## 2.2. Paramètre non linéaire

En substituant (4-3) dans (4-2), la fonction de coût à minimiser devient:



$$\mathbf{r}(\omega) = \left[ (\mathbf{I} - \mathbf{H}(\omega)\mathbf{H}^+(\omega))\mathbf{y} \right]^T \left[ (\mathbf{I} - \mathbf{H}(\omega)\mathbf{H}^+(\omega))\mathbf{y} \right] \quad (4-9)$$

Ainsi le problème d'estimation de la pulsation normalisée revient à résoudre un problème de moindres carrés à une variable. D'après (2-9), la méthode de Gauss-Newton se traduit par le schéma itératif suivant:

$$\omega_{k+1} = \omega_k - \frac{\mathbf{j}^T(\omega_k)\mathbf{r}(\omega_k)}{\|\mathbf{j}(\omega_k)\|^2}; \quad \text{avec } \mathbf{r} = \mathbf{y} - \mathbf{H}\mathbf{H}^+\mathbf{y}$$

$$\text{et } \mathbf{j} = \frac{d\mathbf{r}}{d\omega} = \left[ (\mathbf{I} - \mathbf{H}\mathbf{H}^+) \frac{d\mathbf{H}}{d\omega} \mathbf{H}^+ + \left[ (\mathbf{I} - \mathbf{H}\mathbf{H}^+) \frac{d\mathbf{H}}{d\omega} \mathbf{H}^+ \right]^T \right] \mathbf{y} \quad (4-10)$$

$$\frac{d\mathbf{H}}{d\omega} = \left[ -\mathbf{t} \sin(\omega \mathbf{t}) \cdots \mathbf{P} \mathbf{t} \sin(\mathbf{P} \omega \mathbf{t}) \quad \mathbf{t} \cos(\omega \mathbf{t}) \cdots \mathbf{P} \mathbf{t} \cos(\omega \mathbf{t}) \quad \mathbf{0} \right]$$

L'outil fondamental qui a permis de dériver cette formule est le résultat lié à la dérivation du pseudo-inverse démontré dans [Golub et al., 1973]. La dérivée de la matrice du signal se calcule facilement à partir de (4-1).

La différence de cette approche par rapport à celle suivie dans [Pentelon et al., 1996] réside, d'une part dans le fait que nous avons préféré une formulation réelle à une formulation complexe et, d'autre part, le détail donné au niveau de la dérivation de (4-10) qui n'est pas du tout évident, à savoir le résultat lié à la dérivation du pseudo-inverse. La formulation réelle permet d'implanter facilement cette méthode dans des langages tel que le C ne possédant pas de type complexe.

Le temps de calcul et la mémoire nécessaire pour (4-10) augmente considérablement avec le nombre d'échantillons et/ou d'harmoniques. En pratique, le nombre d'harmoniques pris en compte ne dépasse pas 10, alors que le nombre d'échantillons augmente avec la résolution du CAN et peut atteindre 65536 pour un 14bits. Dans le paragraphe suivant nous réduirons considérablement cette complexité en utilisant une information à priori sur notre problème.

### 2.2.1. Négliger la contribution de la distorsion harmonique

La minimisation de (4-9) se simplifie encore considérablement en prenant en compte le fait que la contribution de la distorsion d'un CAN est largement

inférieure à celle du bruit. En négligeant la contribution de la distorsion harmonique, la matrice du signal, son pseudo-inverse, et la fonction de coût se réduisent à:

$$\begin{aligned} \mathbf{H}_a(\omega) &= [\cos(\omega t) \quad \sin(\omega t) \quad \mathbf{1}]; \quad \mathbf{H}_a^+ = (\mathbf{H}_a^T \mathbf{H}_a)^{-1} \mathbf{H}_a^T \\ r_a(\omega) &= [\mathbf{y} - \mathbf{H}_a (\mathbf{H}_a^T \mathbf{H}_a)^{-1} \mathbf{H}_a^T \mathbf{y}]^T [\mathbf{y} - \mathbf{H}_a (\mathbf{H}_a^T \mathbf{H}_a)^{-1} \mathbf{H}_a^T \mathbf{y}] \\ &= \|\mathbf{y}\|^2 - \mathbf{v}^T(\omega) \mathbf{E}^{-1}(\omega) \mathbf{v}(\omega); \quad \mathbf{v} = \mathbf{H}_a^T \mathbf{y}, \quad \mathbf{E} = \mathbf{H}_a^T \mathbf{H}_a \end{aligned} \quad (4-11)$$

La matrice  $\mathbf{H}_a$  étant de rang plein, donc la matrice  $\mathbf{E} = \mathbf{H}_a^T \mathbf{H}_a$  est inversible. Les éléments de  $\mathbf{E}$  et  $\mathbf{v}$  sont donnés par les formules suivantes:

$$\begin{aligned} \mathbf{E}(\omega) &= \begin{bmatrix} 0.5(M + S_{c,0}(2\omega)) & 0.5S_{s,0}(2\omega) & S_{c,0}(\omega) \\ 0.5S_{s,0}(2\omega) & 0.5(M - S_{c,0}(2\omega)) & S_{s,0}(\omega) \\ S_{c,0}(\omega) & S_{s,0}(\omega) & M \end{bmatrix} \\ \mathbf{v}(\omega) &= \left[ \sum_{n=0}^{M-1} y_n \cos(n\omega) \quad \sum_{n=0}^{M-1} y_n \sin(n\omega) \quad \sum_{n=0}^{M-1} y_n \right]^T \end{aligned} \quad (4-12)$$

L'application des méthodes de Newton et de Gauss-Newton à la minimisation de (4-11) conduit aux processus itératifs suivants:

$$\omega_{k+1} = \omega_k - \begin{cases} \frac{\mathbf{j}_a^T(\omega_k) \mathbf{r}_a(\omega_k)}{\|\mathbf{j}(\omega_k)\|^2} & \Rightarrow \text{Gauss - Newton} \\ \frac{\mathbf{r}'_a(\omega_k)}{\mathbf{r}''_a(\omega_k)} & \Rightarrow \text{Newton} \end{cases} \quad (4-13)$$

où  $\mathbf{r}_a(\omega) = \mathbf{y} - \mathbf{H}_a(\omega) \mathbf{E}^{-1}(\omega) \mathbf{v}(\omega)$  est le résiduel,  $\mathbf{j}_a(\omega) = d\mathbf{r}_a(\omega)/d\omega$ , et  $\mathbf{r}'_a$  et  $\mathbf{r}''_a$  sont les dérivées première et seconde de la fonction de coût. Les dérivées de  $\mathbf{v}$  se calculent facilement. Celles de  $\mathbf{H}_a$  se déduisent aisément en utilisant les règles suivantes:

$$\begin{aligned} \frac{dS_{c,m}(\lambda\omega)}{d\omega} &= -\lambda S_{s,m+1}(\lambda\omega); \quad \frac{dS_{s,m}(\lambda\omega)}{d\omega} = \lambda S_{c,m+1}(\lambda\omega) \\ \text{avec } m &= 0, 1 \quad \text{et } \lambda = 1, 2 \end{aligned} \quad (4-14)$$

Quant aux dérivées de  $\mathbf{E}^{-1}$ , elles sont données par les relations suivantes:

$$\begin{aligned} (\mathbf{E}^{-1})' &= -\mathbf{E}^{-1}\mathbf{E}'\mathbf{E}^{-1} \\ (\mathbf{E}^{-1})'' &= -\left[ (\mathbf{E}^{-1})'\mathbf{E}'\mathbf{E}^{-1} + \mathbf{E}^{-1}\mathbf{E}''\mathbf{E}^{-1} + \mathbf{E}^{-1}\mathbf{E}'(\mathbf{E}^{-1})' \right] \end{aligned} \quad (4-15)$$

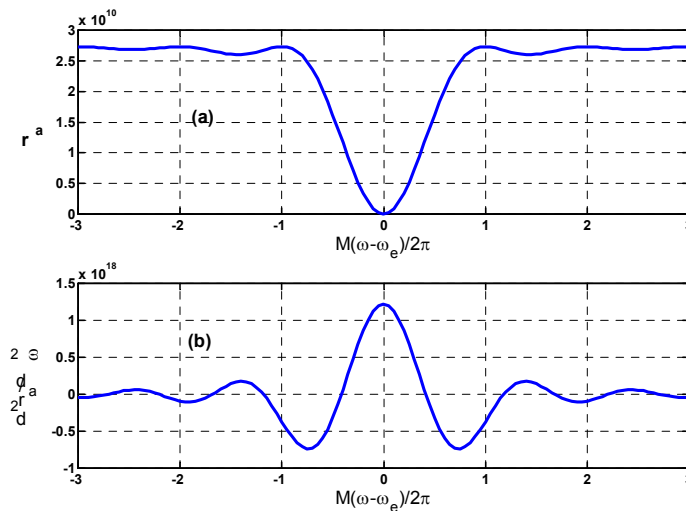
Il faut noter que les dérivées secondes ne sont requises que par la méthode de Newton.

### 2.2.2. L'incertitude tolérée sur la valeur initiale

La figure (4.1.a) montre un exemple de la fonction de coût (4-11) autour du minimum global. On voit qu'en dehors du lobe principal, il y a des minima locaux où l'algorithme d'estimation de la pulsation convergera si la valeur initiale fournie n'est pas suffisamment précise. On peut aussi noter que la largeur de ce lobe est égale à  $4\pi/M$ .

Une condition nécessaire pour que l'algorithme ne diverge pas est que la pulsation initiale se situe dans le lobe principal, ce qui se traduit par:

$$\Delta\omega_0 = |\omega_e - \omega_0| < \frac{2\pi}{M} \quad (4-16)$$



**Figure 4.1:** Variation de la fonction de coût (a) et de sa dérivée seconde (b) autour du minimum global. Paramètres:  $N = 12$ ;  $SNR = 65\text{dB}$ ;  $THD = -74.5\text{dB}$ ;  $M = 16384$ ;  $Jitter = 1\text{ps}$ ;  $\omega_e = 1.318\text{rad}$ .

Supposons que  $\omega_0$  satisfait (4-16). La condition suffisante pour que l'algorithme converge vers le minimum global est que le pas  $\omega_{k+1} - \omega_k$  soit une direction de descente, à savoir:

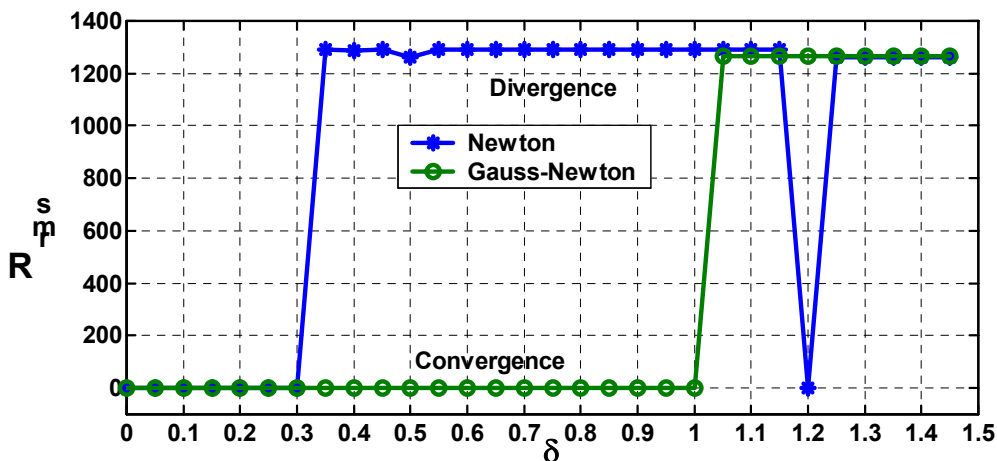
$$(\omega_{k+1} - \omega_k) \frac{dr_a(\omega_k)}{d\omega} < 0 \quad \forall k \quad (4-17)$$

En substituant dans (4-13) on obtient:

$$\begin{cases} \frac{d^2 r_a(\omega_k)}{d^2 \omega} > 0 & \Rightarrow \text{Newton} \\ \frac{(\mathbf{j}_a^T(\omega_k) \mathbf{r}_a(\omega_k))^2}{\|\mathbf{j}_a(\omega_k)\|^2} > 0 & \Rightarrow \text{Gauss - Newton} \end{cases} \quad (4-18)$$

On voit que la méthode de Gauss-Newton remplit la condition (4-18) qui devient une condition suffisante. En revanche, la méthode de Newton exige que la dérivé seconde de la fonction de coût soit positive.

La figure (4.1.b) montre un exemple de variation de la dérivé seconde de la fonction de coût autour du minimum global. On distingue bien la région où la condition (4-18) est satisfaite.



**Figure 4.2:** Variation du résiduel RMS obtenu par les méthodes de Newton et de Gauss-Newton en fonction de l'erreur sur la pulsation normalisée initiale. Paramètres: SNR = 65dB; THD = -74.5dB; M = 16384; Fin = 10.49MHz; Fh = 50MHz; Jitter = 1ps.

Pour confirmer ces résultats, nous avons étudié par simulation l'effet de l'erreur sur la pulsation initiale:

$$\delta = \frac{M\Delta\omega_0}{2\pi} = \frac{M|\omega_e - \omega_0|}{2\pi} \quad (4-19)$$

sur le résiduel RMS:

$$R_{\text{rms}} = \sqrt{\frac{r_a(\hat{\omega}_e)}{M}} \quad (4-20)$$

estimé par les méthodes de Newton et de Gauss-Newton. Le résultat est donné sur la figure (4-2). On fait bien apparaître la précision requise par les deux méthodes afin de converger vers le minimum global:

$$\begin{cases} \Delta\omega_c = \frac{2\pi}{M} & \Rightarrow \text{Gauss - Newton} \\ \Delta\omega_c = \frac{2\pi \times 0.3}{M} & \Rightarrow \text{Newton} \end{cases} \quad (4-21)$$

Ainsi la méthode de Gauss-Newton a l'avantage de tolérer une plus grande incertitude sur la valeur de la pulsation normalisée initiale comparée à la méthode de Newton. Cependant, on sait que le taux de convergence de la méthode de Newton est quadratique, alors que celui de la méthode de Gauss-newton n'est que super-linéaire [**Nocedal et al., 1999**].

La valeur de la pulsation initiale est générée à partir des fréquences des générateurs du signal de test et d'horloge. La méthode développée dans le paragraphe suivant permet d'affiner cette valeur au cas où elle n'est suffisamment précise pour assurer la convergence vers le minimum global.

### 2.2.3. Un algorithme robuste

Connaissant l'incertitude tolérée sur la pulsation initiale et en suivant la même démarche que celle décrite dans le chapitre précédent, une méthode robuste à l'égard de la valeur de la pulsation initiale est développée. Le principe d'implantation de cet algorithme est donné sur la figure (4.3).

Les paramètres de l'algorithme sont calculés à partir des mêmes considérations énoncées au chapitre précédent:

$$\Delta\omega_T = 1.5\Delta\omega_c; \quad R_s = \frac{2^{(N-ENOB_s-1)}}{\sqrt{3}}; \quad N-ENOB_s \approx 3\text{bits} \quad (3-22)$$

La figure (4-4) présente le résiduel RMS calculé par cette méthode en fonction de l'erreur sur la pulsation normalisée initiale. On voit bien que les résiduels RMS calculés par les méthodes de Newton et de Gauss-Newton sont en parfait accord avec le résiduel RMS exacte sur la plage de l'erreur. Cela montre bien la robustesse de cette méthode à l'égard de l'erreur sur la valeur initiale.

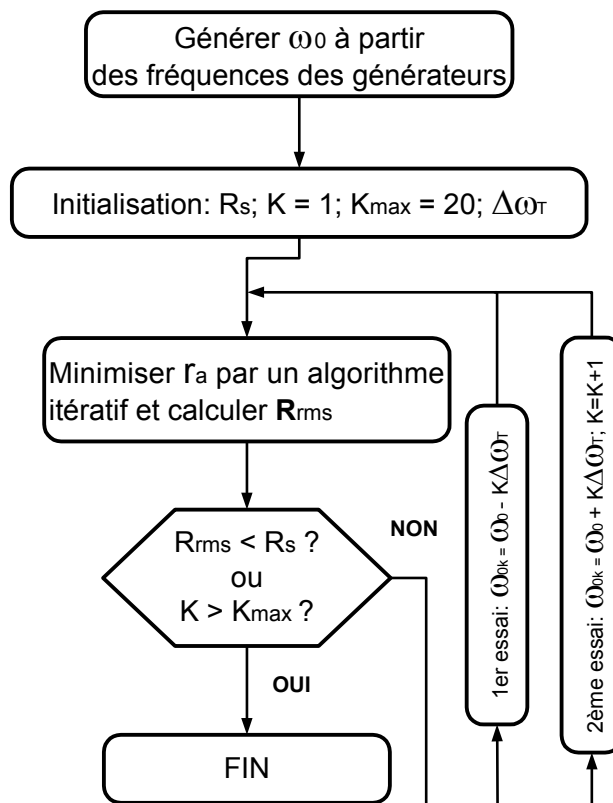
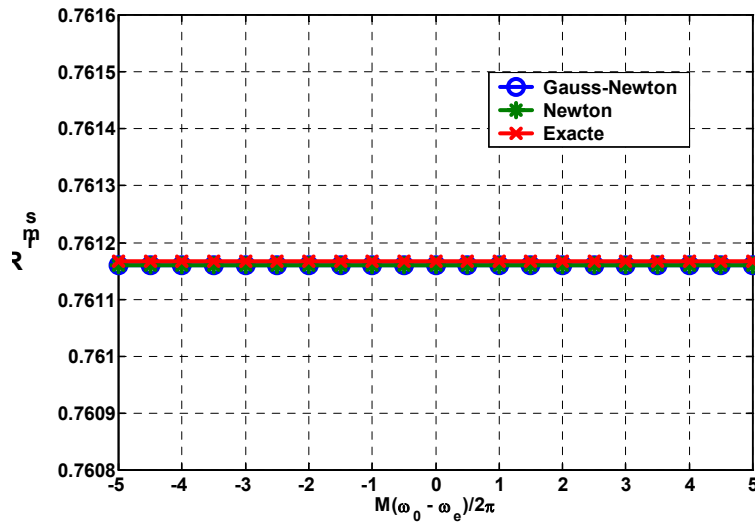


Figure 4-3: Schéma de l'algorithme globalement convergent.



**Figure 4.4:** Variation du résiduel RMS obtenu par les méthodes de Newton et de Gauss-Newton modifiées en fonction de l'erreur sur la pulsation normalisée initiale. Paramètres:  $N = 12$ ;  $SNR = 65.2\text{dB}$ ;  $THD = -75.5\text{dB}$ ;  $M = 8192$ ;  $Fin = 10.49\text{MHz}$ ;  $Fh = 50\text{MHz}$ ;  $Jitter = 1\text{ps}$ .

#### 4. Simulations de Monte-Carlo et résultats réels

Pour évaluer les performances statistiques de la méthode décrite dans ce chapitre, nous avons effectué des simulations de Monte-Carlo sur 200 réalisations. La variance de la pulsation estimée est calculée par l'expression suivante:

$$\text{Var}(\hat{\omega}_e) = \frac{1}{200} \sum_{n=1}^{200} (\hat{\omega}_{en} - \omega_e)^2 \quad (4-23)$$

où  $\hat{\omega}_{en}$  est l'estimation de la  $n^{\text{ème}}$  réalisation et  $\omega_e$  la valeur exacte. Notons que l'évaluation de la dispersion autour de la valeur exacte au lieu de la valeur moyenne permet de tenir compte du biais, à savoir l'écart entre la valeur moyenne

$\frac{1}{200} \sum_{n=1}^{200} \hat{\omega}_{en}$  et la valeur exacte.

La borne de Cramér-Rao de la pulsation normalisée est donnée par (cf. annexe (A.1.1)):

$$\text{Var}(\omega) = \frac{24\sigma^2}{M^3 \sum_{k=1}^P k^2 A_k^2} = \frac{12}{M^3 \text{SNR} \left( \sum_{k=1}^P k^2 \frac{A_k^2}{A_1^2} \right)}; \quad \text{SNR} = \frac{A_1^2}{2\sigma^2} \quad (4-24)$$

où  $\sigma$  est l'écart type du bruit additif,  $A_k$  l'amplitude de la  $K^{\text{ème}}$  harmonique et  $A_1$  l'amplitude du signal de test. On voit que cette variance est inversement proportionnelle au rapport signal à bruit et au cube du nombre d'échantillons. Par conséquent pour améliorer la précision d'estimation il suffit d'augmenter un de ces paramètres.

Pour être le plus réaliste possible, nous avons utilisé un modèle qui tient compte de la plupart des imperfections d'un CAN:

- Le jitter est modélisé par un bruit gaussien  $\sigma_j \boldsymbol{\epsilon}_j$  d'écart type  $\sigma_j$  [ps]. Les indices d'échantillonnage sont donnés par:

$$\mathbf{t} = [0, 1, \dots, M-1]^T + 10^{-12} F_h \sigma_j \boldsymbol{\epsilon}_j \quad (4-25)$$

- Le signal de test est donné par:

$$\mathbf{e}(\mathbf{t}) = A \cos(\omega_e \mathbf{t} + \varphi) + C$$

avec  $PE = 2^N$ ;  $A = \frac{PE}{2} 10^{\frac{\text{SFSR}}{20}}$ ;  $C = \frac{PE}{2}$ ;  $\omega_e = \frac{F_e}{F_h}$ ; (4-26)

- La distorsion harmonique est simulée par 5 raies d'amplitudes contrôlées par la THD du CAN:

$$\mathbf{h}(\mathbf{t}) = A_2 \cos(2\omega_e \mathbf{t}) + 1.3A_2 \cos(3\omega_e \mathbf{t}) + 0.7A_2 \cos(4\omega_e \mathbf{t}) + 1.2A_2 \cos(5\omega_e \mathbf{t}); \quad \text{avec} \quad A_2 = \frac{A}{2.11} 10^{\frac{\text{THD}}{20}} \quad (4-27)$$

La distorsion dominée par les harmoniques impaires permet de tenir compte de la structure d'entrée différentielle des CANs rapides qui atténue les harmoniques paires.

- Le bruit additif est simulé par un bruit gaussien  $\sigma_a \boldsymbol{\epsilon}_a$  d'écart type  $\sigma_a$  alors que celui de la quantification est obtenu grâce à la fonction arrondi. La valeur de

$$\sigma_a \text{ est calculée à partir du SNR du CAN par } \sigma_a = \frac{A}{\sqrt{2}} 10^{\frac{\text{SNR}}{20}}.$$



Le modèle final du CAN est donné par:

$$\mathbf{y} = \text{round}(\mathbf{e}(\mathbf{t}) + \mathbf{h}(\mathbf{t}) + \sigma_a \boldsymbol{\varepsilon}_a) \quad (4-28)$$

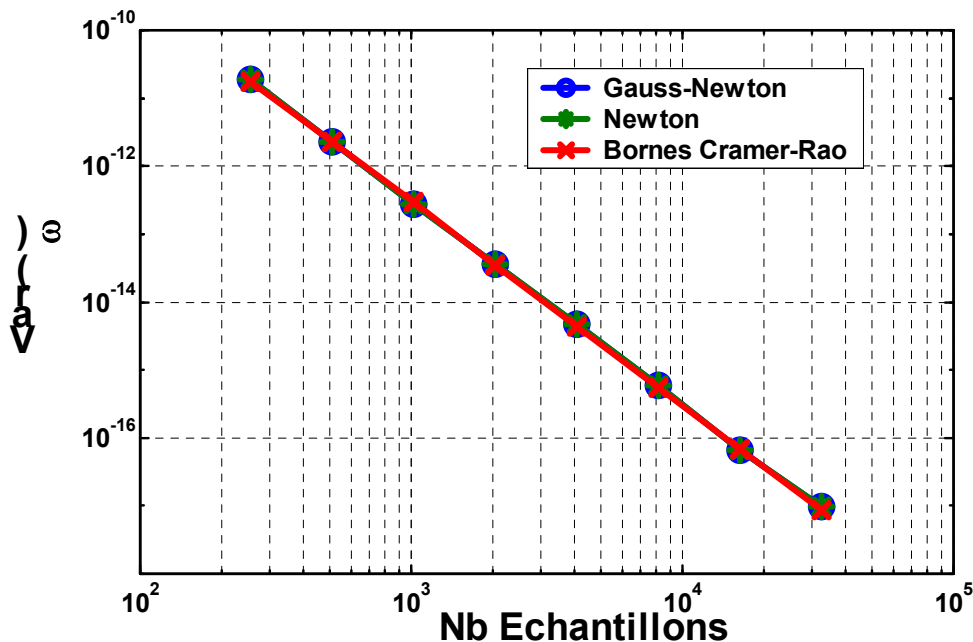
Paramètres : N; M; F<sub>e</sub>; F<sub>h</sub>; SFSR; SNR; THD;  $\sigma_j$

Ce modèle est complètement contrôlé par les 8 paramètres donnés dans (4-28). La valeur exacte du SNR est recalculé pour tenir compte du jitter et du bruit de quantification:

$$\mathbf{r} = \mathbf{y} - \mathbf{e}(\mathbf{t}) - \mathbf{h}(\mathbf{t}); \quad \sigma = \sqrt{\frac{\mathbf{r}^T \mathbf{r}}{M}}; \quad \text{SNR} = 10 \log_{10} \frac{A^2}{2\sigma^2} \text{ [dB]} \quad (4-29)$$

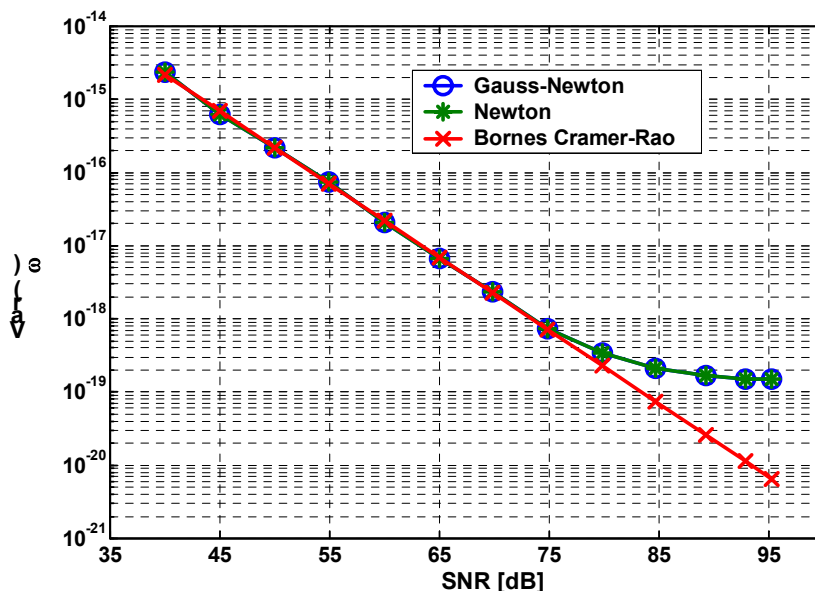
C'est cet écart type ( $\sigma$ ) qui est utilisé pour évaluer la borne de Cramer-Rao (4-24).

La figure (4-5) en échelle log-log montre la variation de la variance en fonction du nombre d'échantillons, et ce pour les méthodes de Newton et de Gauss-Newton. On voit bien que les deux méthodes atteignent la borne de Cramer-Rao sur toute la plage du nombre d'échantillons utilisés habituellement pour le test des CANs.



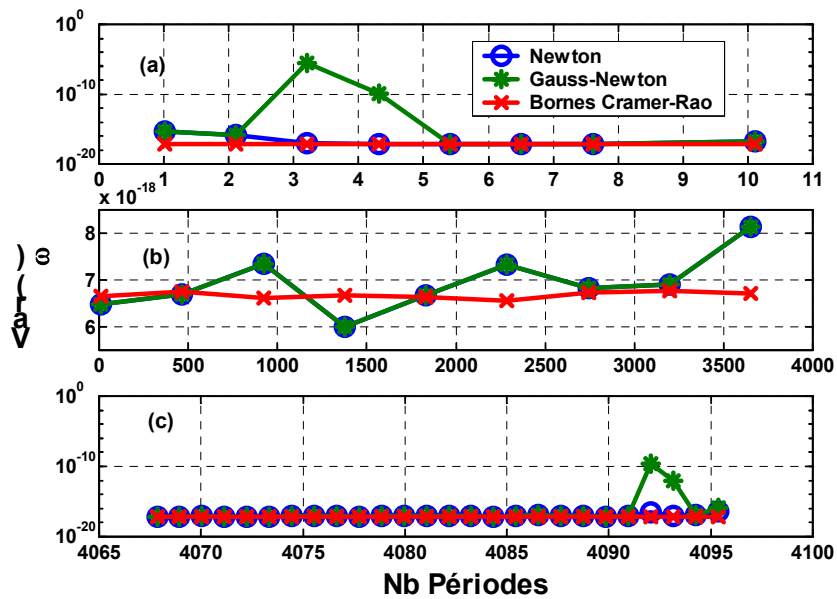
**Figure 4.5:** Variance de la pulsation normalisée calculée sur 200 réalisations et bornes de Cramer-Rao en fonction du nombre d'échantillons. Paramètres: N = 8; SNR = 46dB; THD = -71.5dB; fe = 10.49MHz; fh = 50MHz; Jitter = 1ps.

La figure (4-6) montre l'effet du SNR sur la variance. On distingue deux zones: la première zone correspond à celle où les méthodes de Newton et de Gauss-Newton atteignent la borne de Cramer-Rao. La deuxième zone correspond à celle où les deux méthodes s'écartent de cette borne. La première zone correspond à une contribution du bruit supérieure à celle de la distorsion, ce qui est le cas pour les CANs. L'écart de la deuxième zone traduit le biais introduit par la distorsion harmonique dont la contribution n'est plus négligeable. Ainsi, le fait de négliger l'effet de la distorsion harmonique lors de l'estimation de la pulsation normalisée est complètement justifié. Tenir compte de la distorsion augmentera le temps et la complexité des calculs sans amélioration de la précision.



**Figure 4.6:** Effet du rapport signal à bruit sur la variance de la pulsation normalisée calculée sur 200 réalisations. Paramètres:  $N = 16$ ;  $M = 8192$ ;  $THD = -74.5\text{dB}$ ;  $F_{in} = 10.49\text{MHz}$ ;  $F_h = 50\text{MHz}$ ;  $Jitter = 1\text{ps}$ .

L'effet du nombre de périodes sur les performances des différents algorithmes est montré sur la figure (4-7). On voit que la méthode de Gauss-Newton opère efficacement jusqu'à la limite Nyquist. Par contre, la méthode de Newton n'est performante que pour un nombre de périodes supérieur à 5 et loin de la limite de Nyquist. Notons, cependant, que ces situations sont rares en pratique.



**Figure 4-7:** Effet de nombre de périodes sur la variance de la pulsation normalisée. (a)  $J = [1.01, 10.11]$ ; (b)  $J = [11, 3652.09]$ ; (c)  $J = [4067.85, 4095.35]$ . Paramètres:  $N = 12$ ;  $M = 8192$ ;  $F_h = 50\text{MHz}$ ;  $SFSR = -1\text{dB}$ ;  $THD = -75.5\text{dB}$ ;  $SNR = 65.2\text{dB}$ ;  $Jitter = 1\text{ps}$ .

Nous avons utilisé la méthode décrite dans ce chapitre pour estimer les paramètres dynamiques de trois différents CANs réels. Les conditions de test pour chaque CAN sont données sur le tableau (4-1). Notons que les fréquences de test sont calculées de façon à satisfaire la condition de cohérence ce qui se voit bien sur la figure (4-8) où nous avons représenté les spectres des différents CANs.

**Tableau 4-4:** Conditions de test et paramètres des CANs.

CAN	fh [MHz]	fe [Hz]	M	N [Bits]
1	60	9896850.5859375	32768	12
2	20	10499877.9296875	32768	12
3	1000	250518798.8281	32768	8

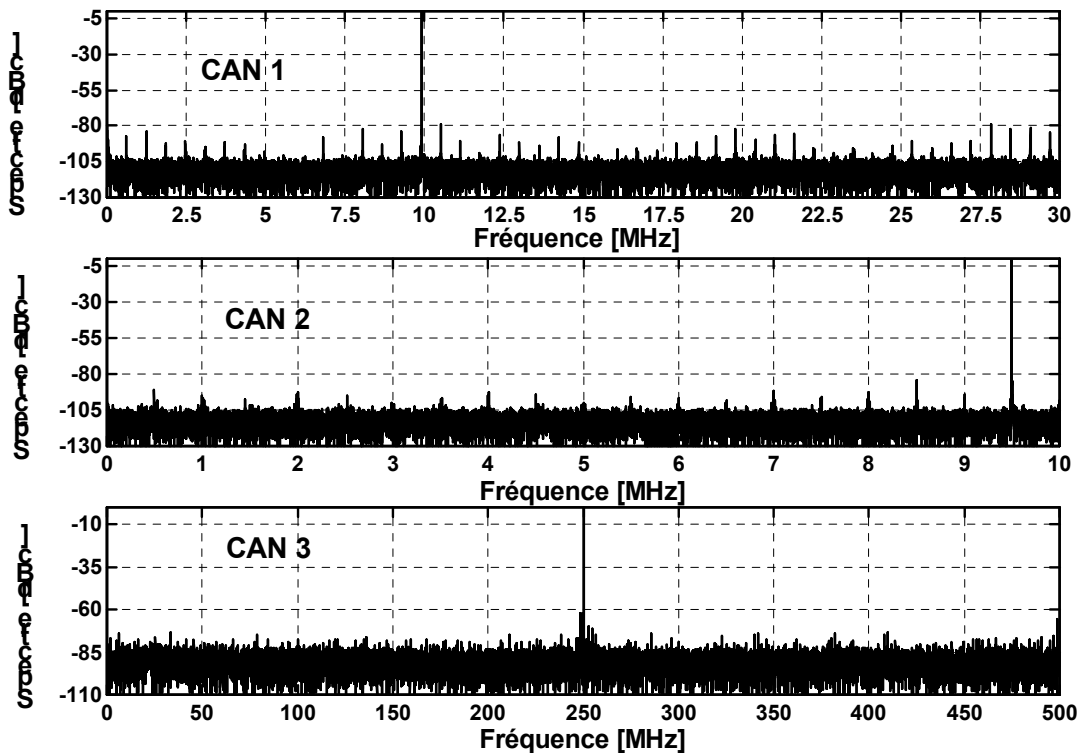


Figure 4-8: Spectres des données des différents CANs.

Pour remplir la condition de cohérence nous avons utilisé un générateur de signaux dont la précision en fréquence est de 0.01Hz. Dans ce cas, on sait que l'analyse spectrale est très précise. C'est pourquoi cette méthode sera utilisée comme référence pour juger les performances de l'analyse temporelle. Nous avons tenu compte des 9 premières harmoniques dans le calcul des différents paramètres. La pulsation initiale a été générée à partir des fréquences d'horloge et de test du tableau (4-1). Les résultats obtenus sont donnés sur le tableau (4-2).

Tableau 4-5 : Résultats des analyses spectrale et temporelles.

CAN	Méthode	SFSR [dBFS]	SINAD [dB]	ENOB [Bits]	SNR [dB]	THD [dB]	DFDR/SFDR [dBFS]	Temps Test[s]
1	Newton	-1.06	65.00	10.68	65.60	-73.88	-78.96/ *	2.64
	Gauss	-1.06	65.00	10.68	65.60	-73.88	-78.96/ *	4.12
	FFT	-1.06	64.04	10.69	65.64	-73.89	-78.97/-78.75	0.16

2	Newton	-1.07	68.10	11.20	68.27	-82.42	-84.09/*	2.75
	Gauss	-1.07	68.10	11.20	68.27	-82.42	-84.09/ *	4.34
	FFT	-1.07	68.30	11.23	68.47	-82.42	-84.09/-84.09	0.27
3	Newton	-0.93	46.40	7.57	46.63	-59.32	-62.17/ *	2.69
	Gauss	-0.93	46.40	7.57	46.63	-59.32	-62.17/ *	4.89
	FFT	-0.97	46.42	7.58	46.64	-59.32	-62.17/-61.97	0.27

On voit que les méthodes de Newton et de Gauss-Newton donnent exactement le même résultat, qui est en très bon accord avec celui de l'analyse spectrale. D'un autre côté, on obtient le résultat attendu à savoir que l'analyse spectrale est plus rapide que l'analyse temporelle. Notons que la méthode de Newton est plus rapide que celle de Gauss-Newton, ce qui s'accorde bien avec la théorie. Compte tenu du nombre d'échantillons et d'harmoniques considérés, le temps de test de l'analyse temporelle est parfaitement adapté au test en laboratoire et à l'étape engineering en test industriel. Cependant, il nécessite une simplification afin d'être utilisé en production.

## 5. Conclusion

Ce chapitre traite de la méthode de séparation des variables appliquée au test des CANs en mode single-tone. C'est une méthode qui a permis de simplifier le problème d'estimation des paramètres du modèle, ainsi que le problème d'estimation des valeurs initiales. Cette complexité a été encore réduite en négligeant l'effet de la distorsion harmonique. Les simulations de Monte-Carlo ont permis de justifier cette approximation.

La détermination de l'incertitude tolérée sur la pulsation initiale par différents algorithmes nous a permis de développer une méthode robuste à l'égard des valeurs initiales, dont les performances ont été validées par simulation. Enfin, les simulations de Monte-Carlo et des données réelles ont montré que les algorithmes présentés atteignent les Bornes de Cramer-Rao et sont en très bon accord avec l'analyse spectrale.

**CHAPITRE 5**

**ANALYSE TEMPORELLE EN MODE DUAL-TONE**



## CHAPITRE 5

### ANALYSE TEMPORELLE EN MODE DUAL-TONE

#### 1. Introduction

Les paramètres à estimer en mode dual-tone nécessitent la connaissance des paramètres du signal de test et des produits d'intermodulation (cf. chapitre 1). Le signal de test qui consiste en la combinaison de deux sinusoïdes peut s'écrire sous la forme matricielle suivante:

$$\mathbf{s}(\boldsymbol{\Omega}, \boldsymbol{\theta}) = a_1 \cos(\omega_1 \mathbf{t}) + b_1 \sin(\omega_1 \mathbf{t}) + a_2 \cos(\omega_2 \mathbf{t}) + b_2 \sin(\omega_2 \mathbf{t}) + C \cdot \mathbf{1} \\ = \mathbf{H}(\boldsymbol{\Omega}) \boldsymbol{\theta}$$

avec :

$$\boldsymbol{\theta} = [a_1 \quad b_1 \quad a_2 \quad b_2 \quad C]^T \quad (5-1)$$

$$a_i = A_i \cos(\varphi_i), \quad b_i = -A_i \sin(\varphi_i); \quad \boldsymbol{\Omega} = [\omega_1 \quad \omega_2]^T$$

$$\mathbf{H}(\boldsymbol{\omega}) = [\cos(\omega_1 \mathbf{t}) \quad \sin(\omega_1 \mathbf{t}) \quad \cos(\omega_2 \mathbf{t}) \quad \sin(\omega_2 \mathbf{t}) \quad \mathbf{1}]$$

$$\mathbf{t} = [0, \dots, M-1]^T; \quad \mathbf{1} = [1, \dots, 1]^T$$

La fonction de coût à minimiser afin d'estimer les paramètres du signal de test est donnée par:

$$r(\boldsymbol{\Omega}, \boldsymbol{\theta}) = (\mathbf{y} - \mathbf{H}(\boldsymbol{\Omega}) \boldsymbol{\theta})^T (\mathbf{y} - \mathbf{H}(\boldsymbol{\Omega}) \boldsymbol{\theta}) \quad (5-2)$$

L'objet de ce chapitre est de montrer comment minimiser la fonction (5-2) en utilisant la méthode de séparation des variables décrite dans le chapitre 2. Ensuite, connaissant les paramètres du signal de test, on développera une méthode permettant l'estimation des paramètres des produits d'intermodulation. Enfin, les performances de cette méthode seront validées à travers des simulations de Monte-Carlo et des données réelles.



## 2. Estimation des paramètres du signal de test

En appliquant la méthode de séparation des variables, la minimisation de (5-2) se fait en deux étapes:

1. Estimation des paramètres non linéaires  $\Omega$  en minimisant la fonction de coût modifiée suivante:

$$\begin{aligned} r(\Omega) &= \left[ \mathbf{y} - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y} \right]^T \left[ \mathbf{y} - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y} \right] \\ &= \|\mathbf{y}\|^2 - \mathbf{v}^T(\omega) \mathbf{E}^{-1}(\omega) \mathbf{v}(\omega) \end{aligned} \quad (5-3)$$

avec  $\mathbf{v} = \mathbf{H}^T \mathbf{y}$ ,  $\mathbf{E} = \mathbf{H}^T \mathbf{H}$

On peut montrer que la matrice  $\mathbf{E}$  est inversible. Les éléments de  $\mathbf{E}$  et  $\mathbf{v}$  sont donnés par les expressions suivantes:

$$\mathbf{E} = \sum_{n=0}^{M-1} \begin{vmatrix} \cos(n\omega_1)^2 & \cos(n\omega_1)\sin(n\omega_1) & \cos(n\omega_1)\cos(n\omega_2) & \cos(n\omega_1)\sin(n\omega_2) & \cos(n\omega_1) \\ * & \sin(n\omega_1)^2 & \sin(n\omega_1)\cos(n\omega_2) & \sin(n\omega_1)\sin(n\omega_2) & \sin(n\omega_1) \\ * & * & \cos(n\omega_2)^2 & \cos(n\omega_2)\sin(n\omega_2) & \cos(n\omega_2) \\ * & * & * & \sin(n\omega_2)^2 & \sin(n\omega_2) \\ * & * & * & * & 1 \end{vmatrix} \quad (5-4)$$

$$\mathbf{v} = \sum_{n=0}^{M-1} [y_n \cos(n\omega_1) \quad y_n \sin(n\omega_1) \quad y_n \cos(n\omega_2) \quad y_n \sin(n\omega_2) \quad 1]^T$$

Les éléments désignés par (\*) s'obtiennent par le fait que  $\mathbf{E}$  est symétrique.

2. Estimation des paramètres linéaires  $\theta$  en résolvant le système linéaire suivant:

$$\mathbf{E}(\hat{\Omega}) \theta = \mathbf{v}(\hat{\Omega}) \quad (5-5)$$

où  $\hat{\Omega}$  est une estimation de  $\Omega$ . La matrice  $\mathbf{E}$  étant symétrique et inversible, ce système peut être résolu efficacement à l'aide de la décomposition de Cholesky.

Dans le paragraphe suivant nous détaillons l'estimation des paramètres non linéaires du signal de test.

## 2.1. Estimation des paramètres non linéaires

La méthode de Gauss-Newton appliquée à (5-3) se traduit par le processus itératif suivant:

$$\begin{aligned}\boldsymbol{\Omega}_{k+1} &= \boldsymbol{\Omega}_k + \Delta\boldsymbol{\Omega}_k \\ \Delta\boldsymbol{\Omega}_k &= [\Delta\omega_{1,k} \quad \Delta\omega_{2,k}]^T \\ (\mathbf{J}^T(\boldsymbol{\Omega}_k)\mathbf{J}(\boldsymbol{\Omega}_k))\Delta\boldsymbol{\Omega}_k &= -\mathbf{J}^T(\boldsymbol{\Omega}_k)\mathbf{y}\end{aligned}\quad (5-6)$$

La matrice  $\mathbf{J}$  de taille  $M \times 2$ , rappelons-le, est la matrice jacobienne du résiduel  $\mathbf{r} = \mathbf{y} - \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y} = \mathbf{y} - \mathbf{H}\mathbf{E}^{-1}\mathbf{v}$ . Elle est donnée par l'expression suivante:

$$\begin{aligned}\mathbf{J} &= \begin{bmatrix} \frac{\partial \mathbf{r}}{\partial \omega_1} = -\frac{\partial(\mathbf{H}\mathbf{E}^{-1}\mathbf{v})}{\partial \omega_1} & \frac{\partial \mathbf{r}}{\partial \omega_2} = -\frac{\partial(\mathbf{H}\mathbf{E}^{-1}\mathbf{v})}{\partial \omega_2} \end{bmatrix} \\ \text{avec } \frac{\partial \mathbf{E}^{-1}}{\partial \omega_i} &= -\mathbf{E}^{-1} \frac{\partial \mathbf{E}}{\partial \omega_i} \mathbf{E}^{-1}; \quad i = 1, 2\end{aligned}\quad (5-7)$$

Notons que le calcul de  $\mathbf{J}$  ne nécessite que le calcul des dérivées de  $\mathbf{E}$ ,  $\mathbf{H}$ ,  $\mathbf{v}$  qui s'obtiennent facilement à partir des équations (5-1) et (5-4).

Quant à la méthode de Newton elle conduit au schéma itératif suivant:

$$\begin{aligned}\Delta\omega_{1,k} &= \frac{\left(\frac{\partial^2 \mathbf{r}}{\partial \omega_1 \partial \omega_2} \frac{\partial \mathbf{r}}{\partial \omega_2}\right) - \left(\frac{\partial^2 \mathbf{r}}{\partial \omega_2^2} \frac{\partial \mathbf{r}}{\partial \omega_1}\right)}{\left(\frac{\partial^2 \mathbf{r}}{\partial \omega_1^2} \frac{\partial^2 \mathbf{r}}{\partial \omega_2^2}\right) - \left(\frac{\partial^2 \mathbf{r}}{\partial \omega_1 \partial \omega_2}\right)^2}; \quad \Delta\omega_{2,k} = \frac{\left(\frac{\partial^2 \mathbf{r}}{\partial \omega_1 \partial \omega_2} \frac{\partial \mathbf{r}}{\partial \omega_1}\right) - \left(\frac{\partial^2 \mathbf{r}}{\partial \omega_1^2} \frac{\partial \mathbf{r}}{\partial \omega_2}\right)}{\left(\frac{\partial^2 \mathbf{r}}{\partial \omega_1^2} \frac{\partial^2 \mathbf{r}}{\partial \omega_2^2}\right) - \left(\frac{\partial^2 \mathbf{r}}{\partial \omega_1 \partial \omega_2}\right)^2} \\ \boldsymbol{\Omega}_{k+1} &= \boldsymbol{\Omega}_k + \Delta\boldsymbol{\Omega}_k\end{aligned}\quad (5-8)$$

L'évaluation des différentes dérivées se fait facilement en se référant à l'expression de chaque variable.

Les processus (5-6) et (5-8) étant itératifs, ils nécessitent alors des valeurs initiales dont le choix est déterminant pour la convergence. Ce problème sera l'objet du paragraphe suivant.

### 2.1.1. Problème des valeurs initiales

La figure (5-1) montre un exemple de la fonction de coût (5-3) autour du minimum global. Comme pour le mode single-tone, on voit bien qu'en plus du minimum

global, il y a des minima locaux vers lesquels les algorithmes convergeront dans le cas où les valeurs initiales ne seraient pas suffisamment précises.

La méthode plus simple est bien évidemment de générer les valeurs initiales à partir des fréquences des générateurs d'horloge et du signal de test. Une alternative, lorsque les générateurs ne sont suffisamment précis, est d'utiliser la méthode basée sur la transformée de Fourier discrète et la fenêtre de pondération de Hanning décrite dans le chapitre 3. C'est une technique simple, rapide et précise.

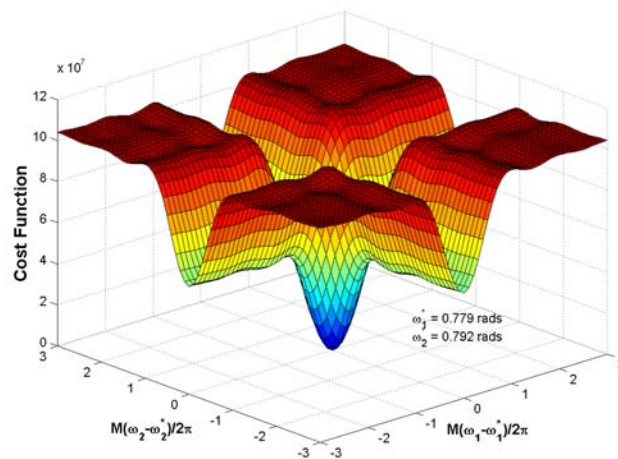


Figure 5-1: Allure de la fonction de coût au voisinage du minimum global.

### 3. Estimation des produits d'intermodulation

Le signal de test étant estimé, il est soustrait des données pour obtenir le résiduel  $\mathbf{r} = \mathbf{y} - \mathbf{s}$ . Les pulsations normalisées des raies d'intermodulation sont données par:

$$\omega_{i,j}^{\pm} = |\omega_1 \pm j\omega_2| \quad \text{avec} \quad (5-9)$$

$$2 \leq i+j \leq q; \quad i > 0, \quad j > 0$$

où  $q$  représente l'ordre d'intermodulation le plus élevé. En pratique ce paramètre est typiquement pris égal à 5. Ainsi, pour un couple d'entiers  $(i, j)$  donné correspond deux raies d'intermodulation  $\omega_{i,j}^+$  et  $\omega_{i,j}^-$ .

La méthode la plus simple et la plus rapide pour estimer les paramètres restants des produits d'intermodulation consiste à extraire une raie à la fois. Une raie de pulsation  $\omega_{i,j}$  peut s'écrire sous la forme:

$$a_{i,j} \cos(\omega_{i,j} \mathbf{t}) + b_{i,j} \sin(\omega_{i,j} \mathbf{t}) \quad (5-10)$$

D'après l'équation (3-25) (cf. chapitre 3), les paramètres de cette raie sont obtenus en résolvant le système linéaire suivant:

$$\begin{bmatrix} 0.5(M + S_{c,0}(2\omega_{i,j})) & -0.5S_{s,0}(2\omega_{i,j}) \\ -0.5S_{s,0}(2\omega_{i,j}) & 0.5(M - S_{c,0}(2\omega_{i,j})) \end{bmatrix} \begin{bmatrix} a_{i,j} \\ b_{i,j} \end{bmatrix} = \sum_{n=0}^{M-1} \begin{bmatrix} r_{i,j;n} \cos(n\omega_{i,j}) \\ r_{i,j;n} \sin(n\omega_{i,j}) \end{bmatrix} \quad (5-11)$$

où les fonctions  $S_{c,0}$   $S_{s,0}$  sont données par l'équation (3-7) et  $r_{i,j}$  correspond au résiduel  $\mathbf{r}$  auquel on a soustrait toutes les raies qui viennent d'être estimées précédemment.

#### 4. Simulations de Monte-Carlo et résultats réels

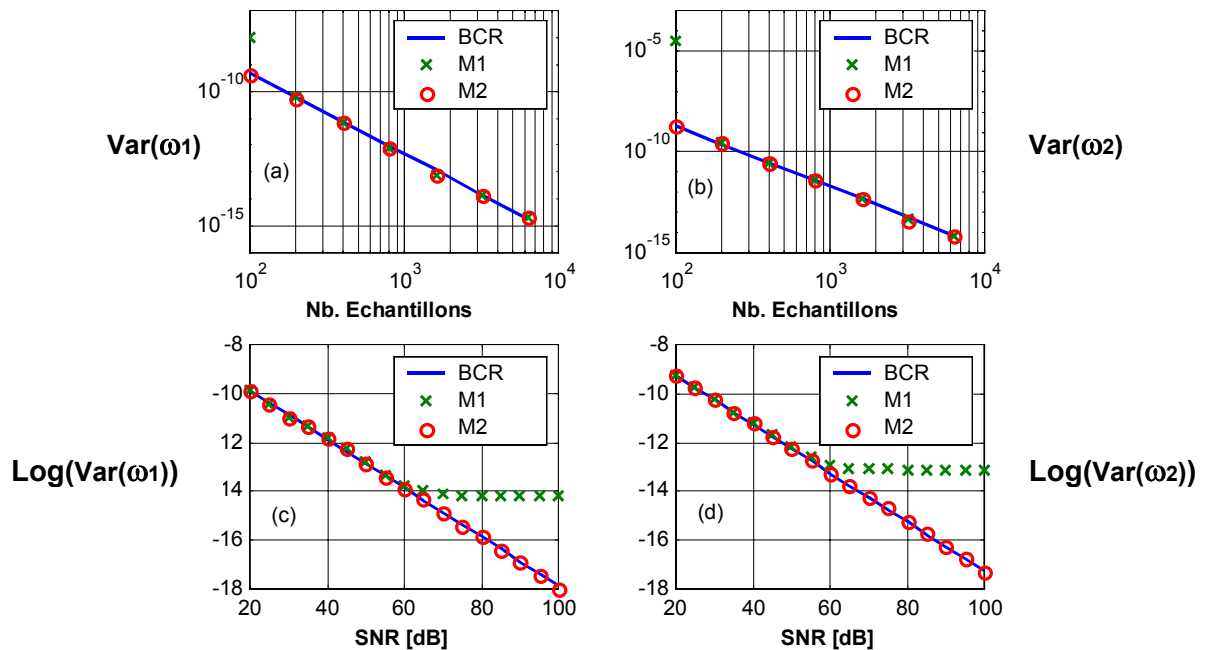
Pour évaluer les performances de la méthode d'estimation des paramètres non linéaires en mode dual-tone, nous avons effectué des simulations de Monte-Carlo sur 100 réalisations indépendantes. Comme pour le mode single-tone, la variance de chaque pulsation normalisée est calculée par rapport à la valeur exacte et comparée à la borne de Cramer-Rao. Les résultats obtenus sont donnés sur la courbe (5-2).

Trois méthodes ont été étudiées:

1. La référence, bien sûr, qui est les bornes de Cramer-Rao (BCR) (cf. annexe (A.1.2)).
2. La méthode (M1) décrite dans ce chapitre et qui néglige l'effet de la distorsion harmonique et d'intermodulation. Nous avons trouvé que les méthodes de Newton et de Gauss-Newton donnent exactement le

même résultat, c'est pourquoi nous avons affiché uniquement les résultats d'une seule méthode.

- Enfin, la méthode (M2) qui tient compte de l'effet de la distorsion dans l'estimation des pulsations normalisées.



**Figure 5-2:** Variation des variances des pulsations normalisées en fonction (a, b) du nombre d'échantillons et (c, d) du rapport signal à bruit

Le CAN a été simulé de manière similaire à celui du mode single-tone, en rajoutant des produits d'intermodulation. Les paramètres du modèle sont les suivants:  $\omega_{1,e} = 0.3$  rads,  $\omega_{2,e} = 1.3$  rads, THD = IMD = -50 dB; SNR = 45dB pour les courbes (a, b) et  $M = 1024$  pour les courbes (c, d). Notons que les courbes (a, b) sont en échelle log-log.

A partir des courbes (a, b), on voit que la méthode décrite ici (M1) atteint les bornes de Cramer-rao dès que l'on dépasse une centaine d'échantillons, ce qui est toujours le cas dans le test des CANs. La méthode (M2) opère efficacement même en deçà de cette valeur. Les courbes (c, d) montre que la méthode (M2) est efficace sur toute la plage du rapport signal à bruit, alors que la méthode (M1) n'atteint les bornes de Cramer-Rao que sur la plage où le SNR est inférieur à la

distorsion. Or, c'est toujours le cas pour les CANs. En conclusion, étant donné les propriétés de notre problème, la méthode décrite dans ce chapitre atteint les bornes de Cramer-Rao et le fait de négliger la distorsion dans l'estimation des pulsations normalisées n'introduit aucune dégradation de la précision. Au contraire, cette approximation permet d'améliorer la vitesse et de réduire considérablement la complexité de l'algorithme.

Nous avons aussi utilisé la méthode décrite ci-dessus pour caractériser deux CANs réels dont les paramètres et les conditions de test sont donnés dans le tableau (5-1). Les spectres des données sont donnés sur la figure (5-3). Pour prendre l'analyse spectrale comme référence, nous avons utilisé des générateurs très précis afin de satisfaire la condition de cohérence avec une bonne précision. Les valeurs initiales ont été générées à partir des fréquences des générateurs donnés sur le tableau (5-1). Nous avons tenu compte des produits d'intermodulation jusqu'à l'ordre 5 dans l'estimation des paramètres.

**Tableau 5-1:** Conditions de test et paramètres des CANs.

<b>CAN</b>	<b>Fh [MHz]</b>	<b>fe1 [Hz]</b>	<b>fe2 [Hz]</b>	<b>M</b>	<b>N [Bits]</b>
<b>1</b>	20	10698852.5390625	9699096.6796875	32768	12
<b>2</b>	1000	123992919.9219	126068115.2344	32768	8

Les résultats sont résumés dans le tableau (5.2). D'une part, on voit que les méthodes de Newton et de Gauss-Newton donnent exactement les mêmes résultats avec un léger avantage pour la méthode de Newton en terme de vitesse. D'autre part, on remarque un très bon accord entre les analyses temporelle et spectrale. Comme attendu, le temps de test de l'analyse spectrale est largement inférieure à celui de l'analyse temporelle. Cependant, compte tenu du nombre important d'échantillons et de produits d'intermodulation considérés, ce temps de test reste tout à fait raisonnable pour le test en laboratoire et pour l'engineering.

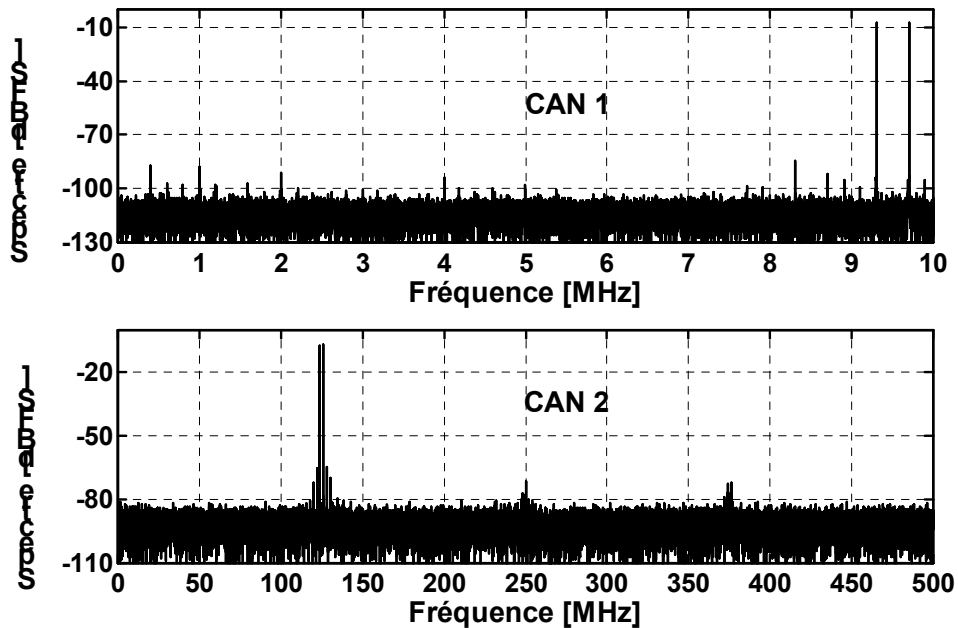


Figure 5-3: Spectres des données issues des différents CANs.

Tableau 5-2 : Résultats des analyses spectrale et temporelles.

CAN	Méthode	SFSR1 [dBFS]	SFSR2 [dBFS]	IMD [dB]	DFDR [dBFS]	SFDR [dBFS]	Temps Test [s]
1	Newton	-7.04	-7.04	-76.00	-84.42		13.95
	Gauss	-7.04	-7.04	-76.00	-84.42		15.71
	FFT	-7.04	-7.04	-76.00	-84.42	-84.42	0.11
2	Newton	-7.17	-7.076	-55.92	-64.48		13.24
	Gauss	-7.17	-7.076	-55.92	-64.48		15.16
	FFT	-7.17	-7.07	-55.92	-64.48	-64.48	0.11

## 5. Conclusion

Dans ce chapitre nous avons étudié l'analyse temporelle en mode dual-tone. Comme pour le mode single-tone, la méthode de séparation des variables nous a permis de réduire considérablement la complexité du problème en séparant l'estimation des paramètres linéaires et non linéaires du modèle. Cette complexité

a été encore réduite en négligeant l'effet de la distorsion lors de l'estimation des pulsations normalisées. Cette approximation a été justifiée par des simulations de Monte-Carlo qui ont montré que la méthode décrite était asymptotiquement efficace dans le sens où elle atteignait les bornes de Cramer-Rao. Des méthodes ont été proposées pour générer des valeurs initiales précises afin d'assurer la convergence globale des algorithmes. Enfin, les résultats réels ont montré un très bon accord entre cette méthode et l'analyse spectrale avec, comme attendu, un grand avantage pour l'analyse spectrale en terme de vitesse. Cependant, le temps de test de l'analyse temporelle reste parfaitement adapté au test en laboratoire et à l'engineering.





**PARTIE II**

**TEST INDUSTRIEL ET  
CONSIDERATIONS INSTRUMENTALES**



**CHAPITRE 6**

**TEST INDUSTRIEL**



## CHAPITRE 6

### TEST INDUSTRIEL

#### 1. Introduction

Les algorithmes décrits dans les chapitres précédents ne sont pas adaptés en terme de temps de test pour le test industriel. Si en caractérisation on privilégie le coût de la solution au détriment du temps de test, c'est l'inverse qui se produit en test industriel. En effet, la simplicité et le gain en temps qu'apporte une instrumentation coûteuse permet de l'amortir rapidement. Cela fait qu'en test industriel on dispose généralement d'une instrumentation qui permet de satisfaire la condition de cohérence (3-10) avec une précision suffisante. Dans ce cas, on verra que les algorithmes se simplifient considérablement, et qu'ils pourraient dans certain cas être utilisés en production.

L'objectif de ce chapitre est de présenter les résultats de l'implantation des analyses temporelle et spectrale en test industriel. L'étude sera faite en modes single-tone et dual-tone, et elle portera sur le temps de test et la précision des algorithmes.

#### 2. L'algorithme

Le modèle général du signal à estimer en modes single-tone et dual-tone peut s'écrire sous la forme suivante:

$$\mathbf{s}(\boldsymbol{\theta}) = \begin{cases} \sum_{k=1}^P A_k \cos(k\omega \mathbf{t} + \varphi_k) + C & \text{single - tone} \\ \sum_{k=1}^P \{A_{1,k} \cos(k\omega_1 \mathbf{t} + \varphi_{1,k}) + A_{2,k} \cos(k\omega_2 \mathbf{t} + \varphi_{2,k})\} + C & \text{dual - tone} \\ \sum_{\substack{i+j=2 \\ i>0, j>0}}^q \{A_{i,j}^+ \cos(\omega_{i,j}^+ \mathbf{t} + \varphi_{i,j}^+) + A_{i,j}^- \cos(\omega_{i,j}^- \mathbf{t} + \varphi_{i,j}^-)\} + C & \end{cases} \quad (6-1)$$

où la définition des différents paramètres du modèle est donnée par les équations (1-2), (1-3), (1-12), et (1-3). Les nouveautés ici sont les produits de distorsion harmonique en mode dual-tone  $\{A_{1,k}, \varphi_{1,k}, A_{2,k}, \varphi_{2,k}\}_{k>1}$ . On suppose que les pulsations angulaires normalisées sont connues avec une bonne précision et que la condition de cohérence (3-10) est remplie. D'après les équations (2-19) et (2-20), l'estimé  $\hat{\boldsymbol{\theta}}$  qui minimise le critère des moindres carrés est obtenu en résolvant le système linéaire suivant:

$$\mathbf{H}^T \mathbf{H} \hat{\boldsymbol{\theta}} = \mathbf{H}^T \mathbf{y} \quad (6-2)$$

avec  $\mathbf{H}$  la matrice du signal définie par:  $\mathbf{s}(\boldsymbol{\theta}) = \mathbf{H}\boldsymbol{\theta}$ , et  $\mathbf{y}$  le vecteur des données. En utilisant les formules (3-7) et le fait que la condition de cohérence est satisfaite, on peut montrer que la matrice  $\mathbf{H}$  est diagonale. A partir de (6-1) et (6-2), on en déduit que les paramètres  $\{a = A \cos(\varphi), b = -A \cos(\varphi)\}$  de la sinusoïde du modèle de pulsation normalisée  $\Omega$ , d'amplitude  $A$  et de phase  $\varphi$ , sont donnés par les relations suivantes:

$$\begin{aligned} a &= \frac{2}{M} \sum_{n=0}^{M-1} y_n \cos(n\Omega) \\ b &= \frac{2}{M} \sum_{n=0}^{M-1} y_n \sin(n\Omega) \end{aligned} \quad (6-3)$$

où  $M$  est le nombre d'échantillons et  $y_n$  la  $n^{\text{ème}}$  donnée. L'offset, quant à lui est donné par l'expression suivante:

$$C = \frac{1}{M} \sum_{n=0}^{M-1} y_n \quad (6-4)$$

A partir de (6-3), les paramètres de toutes les sinusoïdes du modèle peuvent être calculés. Il suffit pour cela de remplacer  $\Omega$  par la pulsation normalisée de la sinusoïde en question. Connaissant les paramètres du modèle, on peut calculer

les paramètres dynamiques du CAN en modes single-tone et dual-tone (cf. chapitre 1).

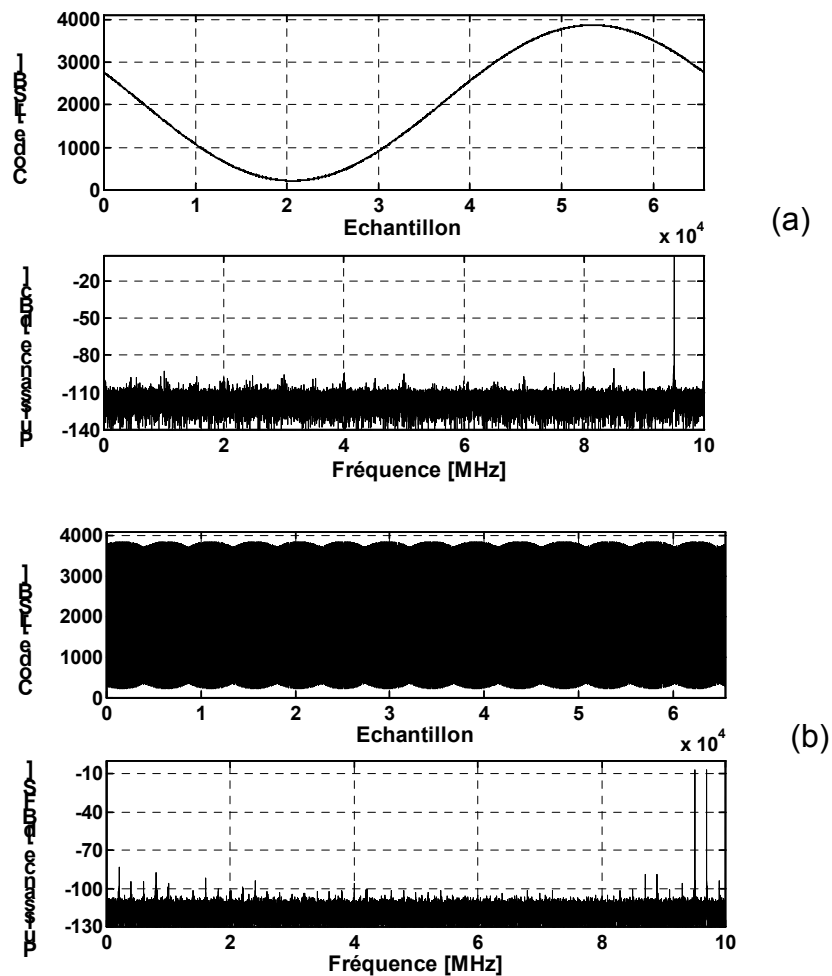
Pour l'analyse spectrale, nous avons utilisé la méthode décrite dans [Dynad, 2001]. Cette méthode est basée sur la transformée de Fourier rapide dont la complexité est de l'ordre de  $M \times \log_2(M)$  [Press et al., 1992]. En revanche, la transformée de Fourier discrète (DFT) est un processus en  $M^2$ . Cette réduction de complexité s'obtient au prix de la contrainte que le nombre d'échantillons soit une puissance de deux. Notons que l'analyse temporelle n'impose aucune condition sur le nombre d'échantillons, si ce n'est celle qu'exige la condition de cohérence, à savoir que les nombres de périodes et d'échantillons doivent être premiers entre eux. Le fait de prendre un nombre d'échantillons qui est une puissance de deux, simplifie la satisfaction de la condition de cohérence, dans la mesure où il suffit alors de prendre un nombre de périodes impair.

Nous avons implanté les deux méthodes de test sur le testeur Catalyst™ de chez TERADYNE. Ce testeur possède toute les ressources permettant de tester industriellement les produits mixtes en général et les CANs en particulier:

- Ressources DC: alimentations, ...
- Ressources numériques: pattern, mémoire, comparateurs et drivers, ...
- Horloge faible jitter: c'est une horloge PECL différentielle ayant un jitter RMS typique de 1.5ps.
- Ressources AC: générateur d'ondes arbitraires, générateur RF, mesure de temps, ...
- Synchronisation entre les parties analogique et numérique: indispensable pour satisfaire la condition de cohérence.

Le véhicule de test est un CAN 12 bits rapide, faible consommation de Chez STMicroelectronics [TSA1204, 2001]. La figure (6-1) montre les représentations temporelle et spectrale des données typiques obtenues en modes single-tone et dual-tone. Dans les paragraphes suivants, nous présenterons les résultats obtenus en terme de temps et de précision.





**Figure 6-1:** (a) Signal reconstitué et spectre en puissance en mode single-tone. Paramètres:  $f_s = 20\text{MHz}$ ,  $f_e = 10.5\text{MHz}$ ,  $M = 65536$ . (b) Signal échantillonné et spectre en puissance en mode dual-tone. Paramètres:  $f_s = 20\text{MHz}$ ,  $f_{e1} = 9.7\text{MHz}$ ,  $f_{e2} = 10.5\text{MHz}$ ,  $M = 65536$ .

### 3. Temps de test

Les résultats relatifs aux temps de test des analyses temporelle (FIT) et spectrale (FFT) en fonction du nombre d'échantillons sont résumés dans le tableau (8-1). Dans ce tableau,  $T_{\text{FIT},1\text{-tone}}$  et  $T_{\text{FIT},2\text{-tone}}$  désignent les temps de test de l'analyse

temporelle respectivement en modes single-tone et dual-tone, et  $T_{FFT,1-tone}$  et  $T_{FFT,2-tone}$  ceux de l'analyse spectrale.

Comme attendu, on retrouve bien la supériorité de l'analyse spectrale par rapport à l'analyse temporelle en terme de vitesse. En mode single-tone, on constate un rapport de vitesse d'environ 43, alors qu'en dual-tone ce rapport est doublé pour atteindre 82. On constate également que les temps de test de l'analyse spectrale en mode dual-tone sont pratiquement identiques, alors qu'en mode single-tone on a un rapport d'environ 2. Enfin, on voit que le temps de test de l'analyse temporelle reste raisonnable pour un nombre d'échantillons inférieur à 8192 en mode single-tone et à 4096 en mode dual-tone. En pratique, cela correspond au test d'un CAN de résolution inférieure ou égale à 10 bits. Notons qu'en pratique le nombre d'échantillons choisi est un compromis entre le temps test et la précision. Ce point sera traité dans le paragraphe suivant.

**Tableau 6-6:** Temps de tests des analyses temporelle et spectrale en fonction du nombre d'échantillons en modes single-tone et dual-tone. Paramètres single-tone:  $f_s = 20\text{MHz}$ ,  $f_e = 10.5\text{MHz}$ ,  $M = 65536$ . Paramètres dual-tone:  $f_s = 20\text{MHz}$ ,  $f_{e1} = 9.7\text{MHz}$ ,  $f_{e2} = 10.5\text{MHz}$ ,  $M = 65536$ .

Nb. Points	Temps de Test [ms]					
	$T_{FIT,1-tone}$	$T_{FFT,1-tone}$	$\frac{T_{FIT,1-tone}}{T_{FFT,1-tone}}$	$T_{FIT,2-tone}$	$T_{FFT,2-tone}$	$\frac{T_{FIT,2-tone}}{T_{FFT,2-tone}}$
2048	88.775	1.901	46.70	171.158	2.062	83.00
4096	183.261	3.883	47.20	346.754	4.026	86.13
8192	368.832	8.258	44.66	710.107	8.277	85.80
16384	754.806	17.608	42.87	1448.366	17.598	82.30
32768	1530.266	36.860	41.52	2944.396	36.378	80.94
65536	3073.255	80.013	38.41	5956.314	79.280	75.13
<b>Moyenne</b>			43.56			82.22

Pour évaluer le temps de test pour n'importe quel nombre d'échantillons, nous avons utilisé les modèles suivants:

$$T_{FIT} = a \times M + b$$

$$T_{FFT} = a \times M \times \log_2(M) + b \tag{6-5}$$

Les paramètres a et b, obtenus en minimisant le critère des moindres carrés, sont donnés dans le tableau (6-2). Sur les figures (6-2) et (6-3), nous avons représenté les données relatives aux modèles et aux mesures, et on constate un très bon accord. Ainsi, le temps de test de l'analyse temporelle varie linéairement avec le nombre d'échantillons, alors que celui de l'analyse spectrale varie linéairement avec  $M \times \log_2(M)$ , qui est le terme qui traduit la vitesse de la transformée de Fourier rapide (cf. paragraphe 2).

Tableau 6-7: Paramètres des modèles.

Paramètre	FIT 1-Tone	FIT 2-Tone	FFT 1-Tone	FFT 2-Tone
a [ $\mu$ s]	47.062	91.232	0.0760	0.0751
b [ $\mu$ s]	-12163.144	-32348.472	83.778	227.330

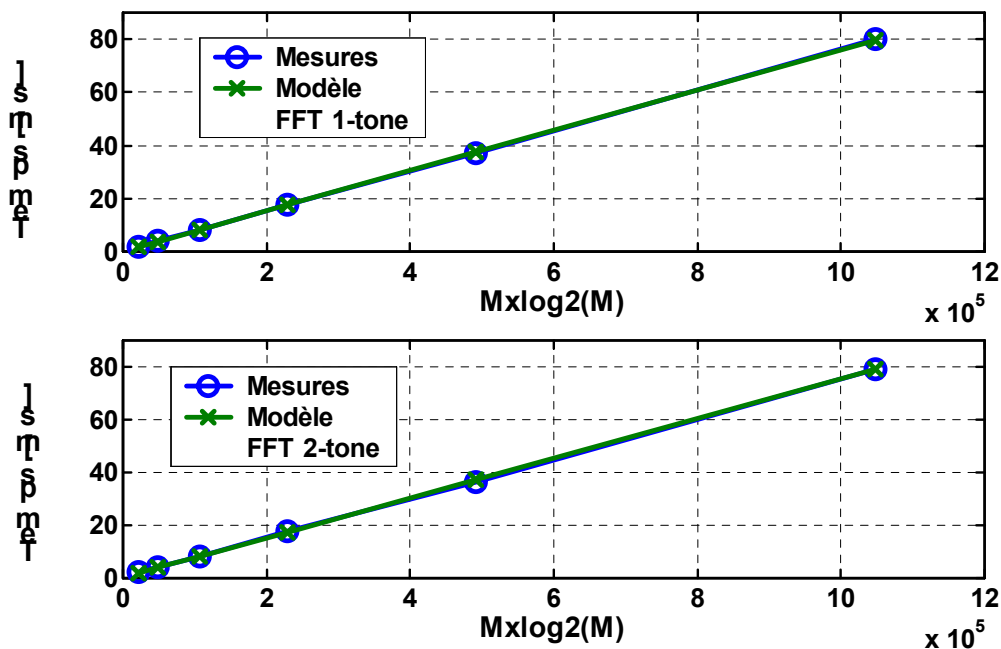


Figure 6-2: Comparaison entre le modèle et les mesures pour l'analyse spectrale.

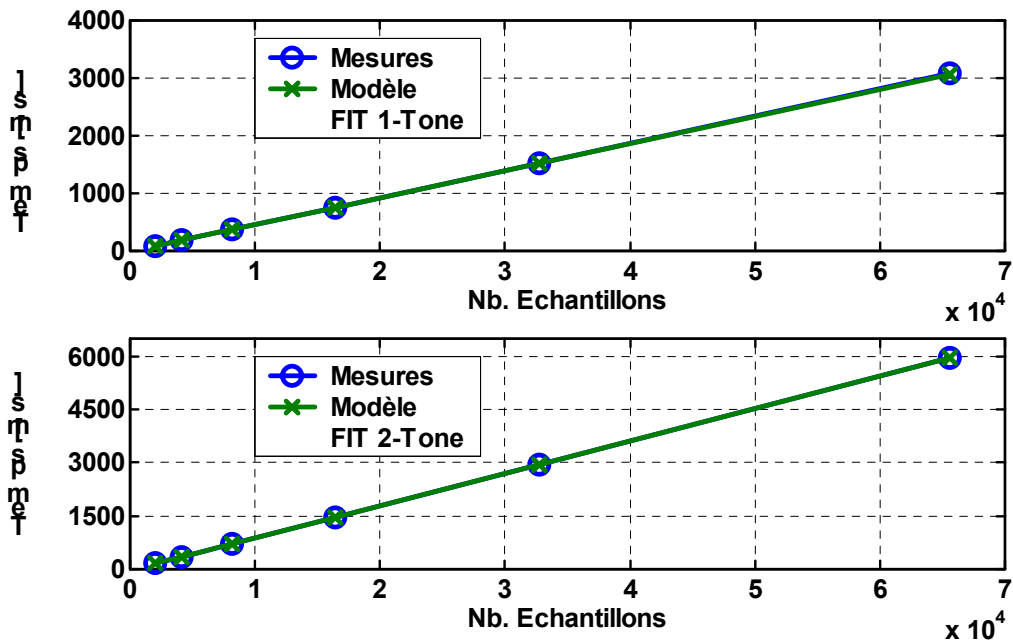


Figure 6-3: Comparaison entre le modèle et les mesures pour l'analyse temporelle.

#### 4. Répétabilité

La répétabilité (précision) est une notion extrêmement importante en test industriel. Pour un paramètre donné, elle se mesure par l'écart type obtenu en testant environ 300 fois la même pièce. Ce paramètre intervient dans le calcul de la marge entre les limites de la datasheet et celles du programme de test en production. Une mauvaise répétabilité élargit les limites de la datasheet et participe donc à la dévalorisation du produit. Ce fait est illustré sur la figure (6-4), où  $\sigma_p$  est l'écart type du process de fabrication, et  $\sigma_t$  celui de la solution de test.

Dans les tableaux (6-3) et (6-4) nous avons donné les résultats statistiques (valeur moyenne, écart type) pour les analyses temporelle et spectrale. On voit un très bon accord entre les deux méthodes en mode single-tone comme en mode dual-tone. Notons que pour l'analyse temporelle le paramètre SFDR est en fait le DFDR.

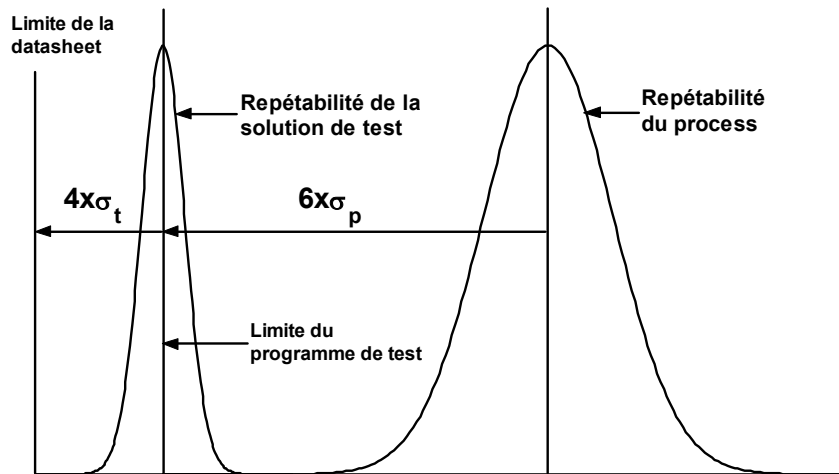


Figure 6-4: Le lien entre la répétabilité et les spécifications.

Tableau 6-8 : Répétabilité en mode dual-tone pour 300 tests et  $M = 32768$ .

Paramètre	FIT		FFT		$\mu_{FIT} - \mu_{FFT}$	$\sigma_{FIT} / \sigma_{FFT}$
	$\mu_{FIT}$	$\sigma_{FIT}$	$\mu_{FFT}$	$\sigma_{FFT}$		
SFSR1 [dBFS]	-7.10	0.0015	-7.10	0.0015	0	1
SFSR2 [dBFS]	-7.00	0.0016	-7.00	0.0016	0	1
IMD [dB]	75.72	0.149	75.72	0.149	0	1
SFDR [dBFS]	83.65	0.237	83.65	0.237	0	1

Tableau 6-9: Répétabilité en mode single-tone pour 300 tests et  $M = 32768$ .

Paramètre	FIT		FFT		$\mu_{FIT} - \mu_{FFT}$	$\sigma_{FIT} / \sigma_{FFT}$
	$\mu_{FIT}$	$\sigma_{FIT}$	$\mu_{FFT}$	$\sigma_{FFT}$		
SFSR [dBFS]	-1.040	0.000851	-1.04	0.000851	0	1
SINAD [dB]	68.26	0.2157	68.25	0.2156	0.01	1
ENOB [Bits]	11.22	0.03583	11.22	0.03582	0	1
SNR [dB]	68.31	0.2182	68.25	0.2156	0.06	1.012
THD [dB]	87.14	0.4081	87.14	0.4081	0	1
SFDR [dBFS]	90.05	0.5437	89.91	0.7306	0.14	0.74

Nous avons vu au deuxième paragraphe l'effet du nombre d'échantillons sur le temps de test. Pour illustrer son effet sur la répétabilité, nous avons donné dans les tableaux (6-5) et (6-6) les résultats statistiques pour un nombre d'échantillons égal à 8192. En comparant ces écarts types avec ceux obtenus avec 32768, on constate que la répétabilité s'améliore avec l'augmentation du nombre d'échantillons. En pratique, le nombre d'échantillons optimal est celui qui réalise un bon compromis entre le temps de test et la répétabilité. Il est déterminé de façon empirique en étudiant le temps de test et la répétabilité en fonction du nombre d'échantillons.

**Tableau 6-10:** Répétabilité en mode dual-tone pour 200 tests et  $M = 8192$ .

Paramètre	FIT		FFT		$\mu_{FIT} - \mu_{FFT}$	$\sigma_{FIT}/\sigma_{FFT}$
	$\mu_{FIT}$	$\sigma_{FIT}$	$\mu_{FFT}$	$\sigma_{FFT}$		
SFSR1 [dBFS]	-7.10	0.0012	-7.10	0.0012	0	1
SFSR2 [dBFS]	-7.00	0.0015	-7.00	0.0015	0	1
IMD [dB]	75.57	0.2791	75.57	0.2791	0	1
SFDR [dBFS]	83.66	0.4219	83.66	0.4219	0	1

**Tableau 6-11:** Répétabilité en mode single-tone pour 200 tests et  $M = 8192$ .

Paramètre	FIT		FFT		$\mu_{FIT} - \mu_{FFT}$	$\sigma_{FIT}/\sigma_{FFT}$
	$\mu_{FIT}$	$\sigma_{FIT}$	$\mu_{FFT}$	$\sigma_{FFT}$		
SFSR [dBFS]	-1.01	0.000802	-1.01	0.000802	0	1
SINAD [dB]	68.46	0.3201	68.46	0.3202	0	0.999
ENOB [Bits]	11.25	0.05317	11.25	0.05319	0	0.999
SNR [dB]	68.52	0.3244	68.46	0.32	0.06	1.013
THD [dB]	-87.15	0.7059	-87.15	0.7059	0	1
SFDR [dBFS]	-90.31	0.9785	-90.27	0.9267	-0.04	1.0559

## 5. Conclusion

Après la simplification de l'analyse temporelle, en se plaçant dans le cas où la condition de cohérence est satisfaite, ce chapitre présente les résultats de l'implantation industrielle des analyses temporelle et spectrale. Comme attendu, nous avons retrouvé la supériorité de l'analyse spectrale en terme de temps de test. Des modèles simples et précis ont été développés pour l'évolution du temps de test en fonction du nombre d'échantillons. Nous avons également constaté que l'analyse temporelle est adaptée au test de CANs de résolution inférieure à 10 bits qui n'exigent pas un nombre d'échantillons élevé. D'un autre coté, il y a un très bon accord entre les deux méthodes en terme de précision.

Les chapitres suivants seront consacrés aux problèmes liés à la carte de test et aux limitations de l'instrumentation.

## **CHAPITRE 7**

### **MINIMISER LES PERTURBATIONS DE LA CARTE DE TEST**





## CHAPITRE 7

### MINIMISER LES PERTURBATIONS DE LA CARTE DE TEST

#### 1. Introduction

L'implantation hardware du test d'un CAN rapide présente des difficultés qui ne sont pas toujours faciles à surmonter. Ces difficultés augmentent avec la vitesse et la résolution du CAN. Cela est dû, d'une part à la nature mixte du produit, et d'autre part à la limitation des performances de l'instrumentation de test. La nature mixte d'un CAN nécessite de faire cohabiter sur la même carte trois éléments antagonistes: une partie analogique sensible, une horloge sensible et bruyante, et une partie numérique bruyante.

Il y a toujours un conditionnement du signal entre la source et le composant à tester, ce qui injecte du bruit dans le système et crée de la distorsion. Ce conditionnement concerne les alimentations, le signal de test, l'horloge, et les données de sortie. Cependant, le conditionnement ne permet pas toujours de surmonter les limitations de l'instrumentation de test, telles que la grande pureté spectrale requise sur le signal de test ou encore le faible jitter sur l'horloge.

Le principe de l'implantation des tests permettant de mesurer les paramètres d'un CAN est décrit dans [dynad, 2001]. Dans ce chapitre, nous reviendrons sur un certain nombre de considérations fondamentales dont il faut tenir compte alors de la réalisation de la carte de test, afin d'optimiser l'extraction des performances du composant à tester.

## 2. Le bruit des sources d'alimentation

L'effet des alimentations est généralement spécifié dans les datasheets par le paramètre PSRR (Power Supply Rejection Ratio). Ce paramètre donne la variation d'un paramètre donné, tel que le gain ou l'erreur d'offset, par rapport à une variation de la tension d'alimentation, et le tout étant exprimé en dB:

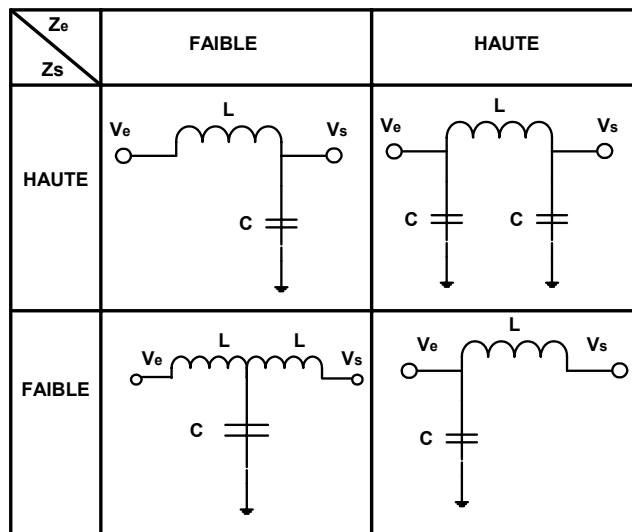
$$\text{PSRR [dB]} = 20 \log \left( \frac{\Delta \text{OE}}{\Delta V_{\text{cc}}} \right) \quad (7-1)$$

Ainsi, si une variation  $\Delta V_{\text{cc}} = 250\text{mV}$  de la tension d'alimentation cause une variation de l'erreur d'offset de  $\Delta \text{OE} = 0.5\text{mV}$ , on obtiendra alors un PSRR de  $-54\text{dB}$ .

Cependant, ce paramètre ne reflète pas tout à fait la réalité où le bruit d'alimentation est dynamique. En effet, on constate que les performances du composant en terme de rejection d'un bruit d'alimentation dynamique (AC PSRR) sont inférieures à celles données par le paramètre PSSR statique. Cela s'explique par le fait qu'un bruit dynamique se couple vers les circuits actifs du composant via les capacités parasites des étages de polarisation, et dégrade ainsi les performances du composant en terme de bruit. Par conséquent, c'est la variation du rapport signal à bruit en fonction d'un signal dynamique sur la tension d'alimentation qui révèle le mieux la sensibilité du composant vis-à-vis du bruit d'alimentation.

Pour minimiser ce bruit, les signaux d'alimentation doivent être filtrés avant de les appliquer au composant. Les filtres utilisés sont généralement de type LC et agissent par perte par insertion. Leur topologie dépend des impédances d'entrée et de sortie (cf. figure (7-1)). Le choix des types comme des valeurs des éléments du filtre est extrêmement important. Typiquement, on utilisera une ferrite d'impédance environ égale à  $50\Omega$  pour  $100\text{MHz}$ , et une capacité polarisée de  $10\mu\text{F}$  (pour les basses fréquences) en parallèle avec une capacité céramique

d'environ 100nF (pour les hautes fréquences). Ce filtre doit être placé sur la carte au plus près du connecteur d'entrée de l'alimentation.



**Figure 7-1:** Filtres d'alimentation en fonction des impédances d'entrée ( $Z_e$ ) et de sortie ( $Z_s$ ).

De plus, les broches d'alimentation du composant doivent être découplées séparément par une capacité céramique haute fréquence, typiquement de l'ordre de 100nF. Le placement de cette capacité ainsi que son routage sont critiques, notamment pour la partie numérique. En effet, un condensateur possède une inductance série (ESL) intrinsèque qui limite son efficacité en haute fréquence. Il ne faut donc pas augmenter la valeur de cette inductance lors du routage des signaux liés à ce condensateur.

Sur la figure (7-2), nous avons représenté cette situation. Sur la figure (a) les vias d'alimentation et de masse sont connectés au condensateur par des pistes dont l'inductance s'ajoute à l'inductance intrinsèque. Cette configuration augmente aussi la boucle de courant, ce qui soulève également le risque d'interférence électromagnétique (EMI), représenté par le champ magnétique **B**. Une meilleure configuration est donnée dans la figure (b) où les vias sont directement connectés aux broches du condensateur, cela minimise et l'inductance et la surface de la boucle. Cependant, cette solution n'est pas toujours applicable en pratique à cause des contraintes de fabrication. Mais dans tous les cas, on s'efforcera de

minimiser la surface de la boucle et la longueur des pistes. L'inductance série pourra aussi être diminuée en mettant plusieurs condensateurs en parallèles.

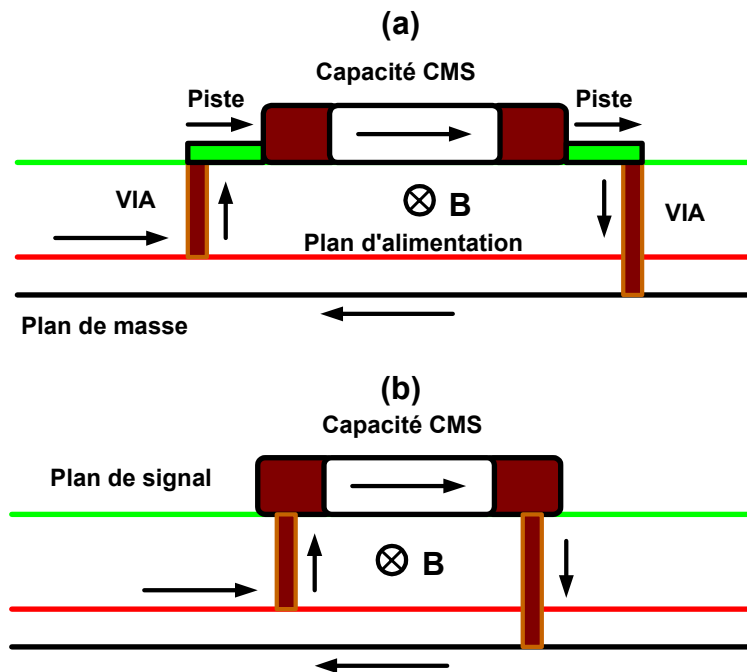


Figure 7-2: Placement et routage de la capacité de découplage.

### 3. La pureté spectrale du signal test

La pureté spectrale d'un synthétiseur de signaux est définie par trois caractéristiques: la distorsion harmonique, le bruit de phase et le niveau des raies parasites. Pour les générateurs RF, les deux dernières spécifications sont en général satisfaisantes, alors que la distorsion harmonique n'est que de l'ordre de  $-30\text{dBc}$ . Or, la distorsion du signal de test doit être meilleure que celle du CAN à tester d'environ  $20\text{dB}$  [dynad, 2001]. Pour répondre à cette exigence, le signal de test est filtré par un filtre passe bande passif, qui permet également de filtrer le bruit situé hors de la bande passante.

Le choix du filtre est critique. Les spécifications importantes sont les suivantes:

- **L'atténuation minimale [dB]:** les fabricants réalisent des filtres dédiés pouvant atteindre une atténuation de 80dB. Cette valeur doit être atteinte de la fréquence de la deuxième harmonique jusqu'à environ 1GHz. En effet, la bande passante en amplitude des CANs est en général assez large, ce qui laisse passer les signaux hautes fréquences. Ces signaux se retrouveront sur le spectre des données par le phénomène de repliement spectral.
- **La largeur de la bande passante à -3dB [%]:** elle est exprimée en pourcentage de la fréquence centrale, et elle est typiquement de l'ordre de 20%.
- **La perte par insertion [dB]:** le choix de cette valeur dépend de la puissance du synthétiseur et de la dynamique du CAN. Notons aussi qu'il y a toujours un compromis à trouver entre la bande passante et la perte par insertion. En effet, une bande passante étroite s'obtient au prix d'une plus grande perte par insertion.
- **Les impédances d'entrée et de sortie:** l'impédance d'entrée est celle du synthétiseur, qui est généralement égale à  $50\Omega$ . Pour transférer le maximum d'énergie, l'impédance de sortie doit être aussi prise égale à  $50\Omega$ .
- **La puissance d'entrée admise par le filtre [dBm]:** cette valeur doit être égale à la puissance du synthétiseur, afin d'exploiter toute sa dynamique et de compenser les inévitables pertes par insertion. En pratique, les filtres commerciaux admettent une puissance de 13 à 16dBm.

Sur la figure (7-3) nous avons représenté le paramètre  $S_{12}$  d'un filtre passe bande de fréquence centrale de 24MHz, sur une bande 300KHz-1GHz. On voit qu'il a une perte par insertion de  $-0.76\text{dB}$  et une bande passante d'environ 40%. L'atténuation à la deuxième harmonique est de l'ordre de 90dB.

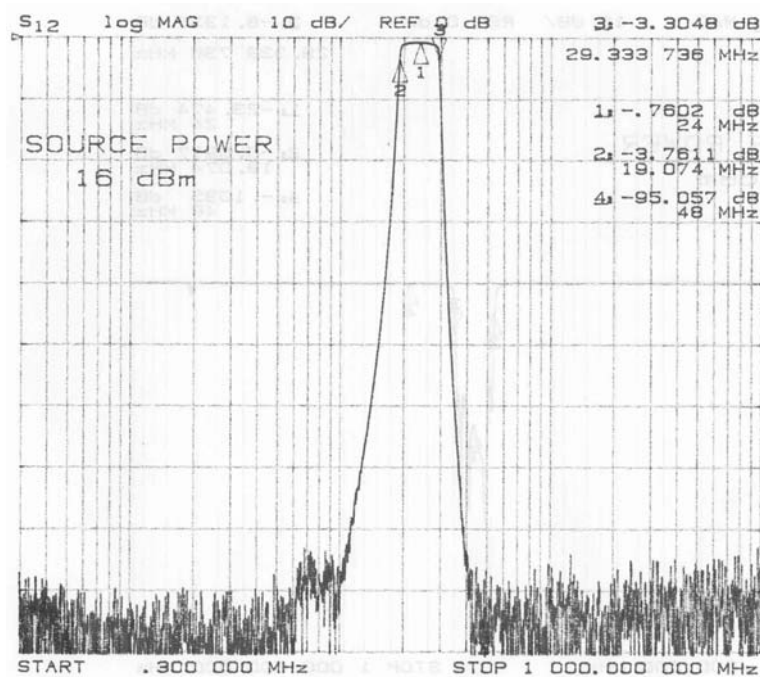


Figure 7-3: Paramètre S12 d'un filtre passe bande.

#### 4. Bruit de commutation et diaphonie

La figure (7-4) représente un modèle simplifié d'un CAN. L'étage de sortie est constitué essentiellement d'inverseurs (transistors  $T_1$  et  $T_2$ ). Les connexions de bonding et du package sont modélisées par une inductance ( $L_p$ ). Les alimentations et les masses de la partie analogique et de l'étage de sortie sont tenues isolées à l'intérieur de la puce pour éviter un couplage via le substrat. Cependant, il persiste un couplage capacitif modélisé par une capacité de fuite  $C_f$ . La charge au niveau des sorties est modélisée par une capacité  $C_s$ , et la capacité de découplage de l'alimentation numérique est notée  $C_d$ . Enfin,  $R_{on}$  désigne la résistance de contact du transistor qui est de l'ordre d'une dizaine d'Ohms.

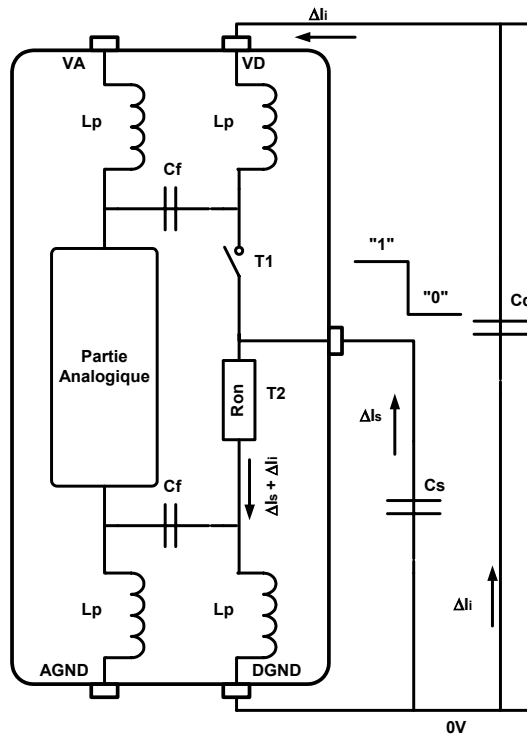


Figure 7-4: Modèle simplifié d'un CAN.

La partie numérique consomme trois types de courant. Le premier est celui consommé par les circuits internes et qui ne pose pas de problème particulier. Le deuxième ( $\Delta I_s$ ) est celui de la charge et de la décharge de la capacité  $C_s$ . Enfin, le troisième est le courant de surconsommation interne ( $\Delta I_i$ ). La source de ce dernier courant est la suivante: à chaque transition, les deux transistors haut et bas se trouvent simultanément fermés pendant une durée de temps très brève. Ainsi, une connexion directe s'établit entre l'alimentation et la masse donnant naissance à une forte transition de courant.

Les courants  $\Delta I_s$  et  $\Delta I_i$  provoquent une chute de tension aux bornes des inductances  $L_p$  donnée par la relation:  $L_p \frac{\Delta I}{\Delta t}$ , où  $\Delta t$  est la durée de transition des courants. Cette chute de tension provoque un bruit d'alimentation lorsque la sortie commute de l'état bas à l'état haut, et un bruit de masse (ground bounce) lors d'un basculement d'un état haut à un état bas. L'effet du bruit d'alimentation est atténué



grâce à la capacité de découplage, en revanche le bruit de masse ne peut être minimisé que par l'optimisation des différentes grandeurs qui interviennent dans l'expression de la chute de tension, à savoir:

- **L'inductance**  $L_p$ : une partie de cette inductance est figée par le choix du package et du bonding. Une deuxième partie dépend des pistes connectant les broches de masse sur la carte de test. Ce sont les longueurs de ces pistes qu'il faut minimiser pour réduire leur inductance.
- **La variation du courant**  $\Delta I$ : le courant  $\Delta I_i$  dépend du design et ne peut donc être minimisé, par contre le courant  $\Delta I_s$  peut être réduit en baissant autant que possible la capacité de la charge  $C_s$ .
- **La durée de transition**  $\Delta t$ : également cette valeur ne peut pas être augmentée car elle dépend du design.

Le bruit de masse se couple inévitablement vers la partie analogique via la capacité de fuite du substrat. Il peut également se coupler par impédance commune sur la carte de test, c'est pourquoi il faut isoler physiquement la partie analogique de la partie numérique sur la carte de test. Contrairement à ce que certains constructeurs recommandent, il n'est pas nécessaire de séparer les plans de masse pour pallier à ce problème. Au contraire, le fait de séparer les plans de masse risque d'augmenter le bruit de masse, et du coup le couplage via le substrat.

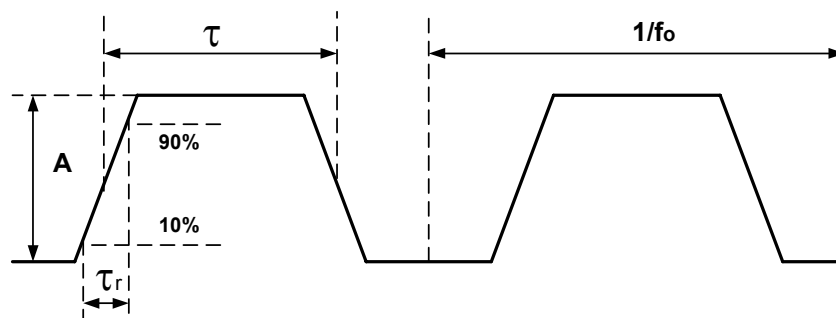


Figure 7-5: Signal numérique trapézoïdal .

Avec l'utilisation de technologies de plus en plus avancées, les temps de montée des signaux numériques ne cessent de baisser, et les problèmes d'intégrité de signal (SI) deviennent de plus en plus aigus. Ce lien entre les temps de montée et l'intégrité de signal peut être illustré à l'aide d'un signal périodique trapézoïdal (cf. figure 7-5) d'amplitude  $A$ , de largeur d'impulsion  $\tau$  et de temps de montée  $\tau_r$ . La décomposition en série de Fourier de ce signal permet de l'écrire sous la forme analytique suivante:

$$s(t) = \sum_{n=-\infty}^{\infty} C_n e^{j2n\pi\left(f_0 t - \frac{\tau + \tau_r}{2}\right)}; \quad C_n(f_0) = A f_0 \tau \operatorname{sinc}(n\pi\tau f_0) \operatorname{sinc}(n\pi\tau_r f_0) \quad (7-2)$$

avec :  $f_0$  = fréquence du signal, et  $\operatorname{sinc}(x) = \frac{\sin(x)}{x}$

On peut montrer que l'enveloppe des coefficients  $C_n$  admet trois asymptotes, comme le montre la figure (7-6), en fonction de la fréquence continue  $f = n f_0$  :

- Une asymptote de pente 0dB/décade dans l'intervalle  $\left[ f_0, \frac{1}{\pi\tau} \right]$  qui n'apparaît que lorsque  $\frac{1}{\pi\tau} > f_0$ .
- Une asymptote de pente -20dB/décade dans l'intervalle  $\left[ \frac{1}{\pi\tau}, \frac{1}{\pi\tau_r} \right]$ .
- Enfin, une asymptote de pente -40dB/décade pour  $f > \frac{1}{\pi\tau_r}$ .

Ainsi, la vitesse d'un signal numérique est définie non pas par sa fréquence mais par son temps de montée. Le contenu spectral s'étend jusqu'à la fréquence effective:

$$f_e \text{ [GHz]} = \frac{1}{\pi\tau_r \text{ [ns]}} \quad (7-3)$$

Les phénomènes de diaphonie, réflexion et d'interférence électromagnétique, principales causes de la dégradation de l'intégrité des signaux, augmentent avec la fréquence des signaux, et donc avec la diminution des temps de montée. Les pistes numériques doivent être considérées comme des lignes de transmission dès que leur longueur dépasse la longueur critique suivante:

$$d_c = 0.5\tau_r \frac{c}{\sqrt{\epsilon_r}} \quad (7-4)$$

où  $c$  est la vitesse de la lumière dans le vide ( $3 \times 10^8 \text{ m} \cdot \text{s}^{-1}$ ), et  $\epsilon_r$  la permittivité relative du diélectrique ( $\epsilon_r = 4.5$  pour un diélectrique FR-4). Si on désigne par  $e$  la distance entre deux pistes et par  $h$  leur hauteur par rapport au plan de masse (cf. figure (7-7)), alors la diaphonie (inductive et capacitive) décroît rapidement dès que  $e > 3h$  [Charoy, 2000].

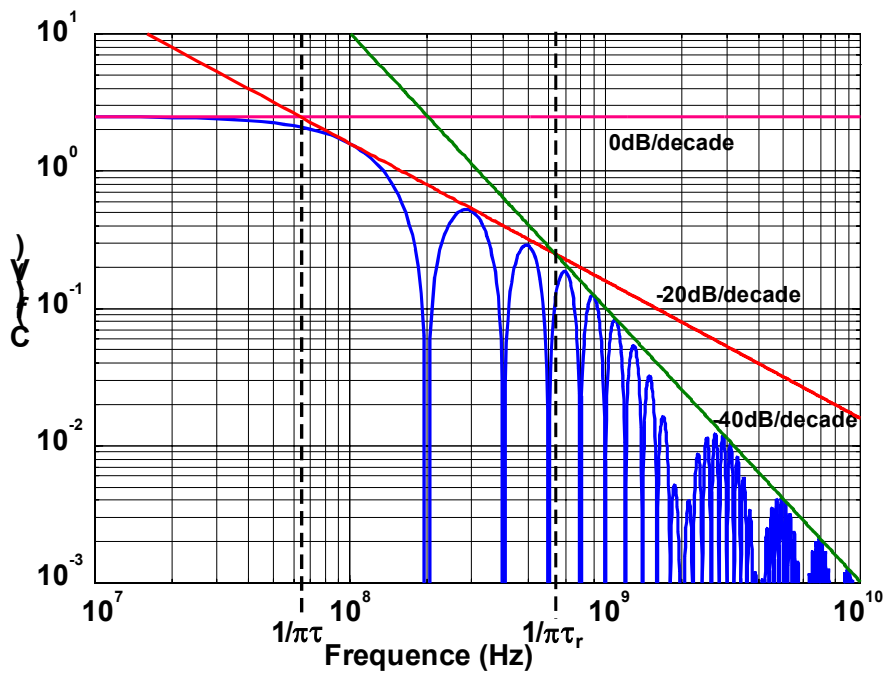


Figure 7-6: Spectre d'amplitude d'un signal trapézoïdal avec:

$A=2.5V$ ,  $f_0=100MHz$ ,  $\tau_r=0.5ns$ ,  $\tau=5ns$ .

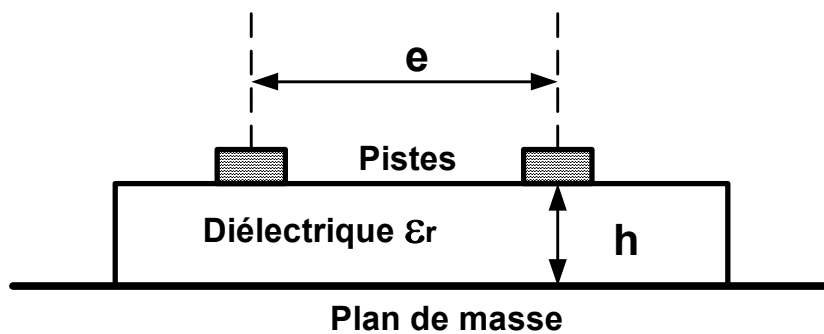


Figure 7-7: Deux pistes numériques au-dessus d'un plan de masse.

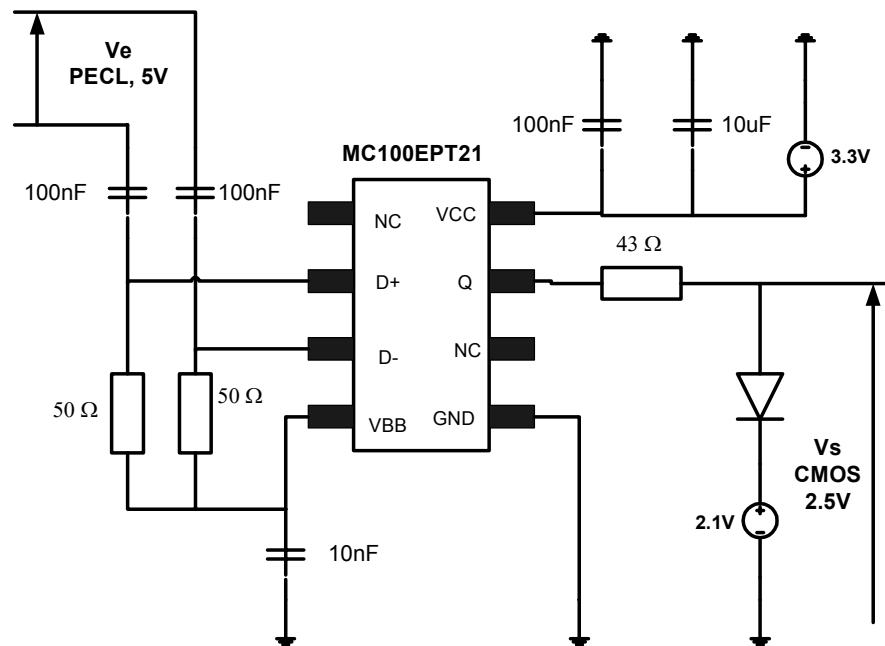
En utilisant la théorie des lignes de transmission, on peut montrer que la diaphonie est proportionnelle à la longueur des pistes et à la vitesse (temps de montée) des signaux. L'utilisation d'un plan de masse est indispensable pour réduire la diaphonie dans une carte rapide.

## 5. Le Signal d'horloge

Le signal d'horloge est un signal à la fois sensible et perturbateur. Il faut donc l'isoler des parties numérique et analogique. De plus, comme c'est un signal numérique, les considérations décrites ci-dessus en terme de contenu spectral et de ligne de transmission s'appliquent également.

Les perturbations sur le signal d'horloge se traduisent par une augmentation du jitter qui dégrade le rapport signal à bruit du composant à tester, et une déformation de la forme du signal. L'adaptation d'impédance permet de préserver la forme du signal lorsque la longueur de la piste d'horloge dépasse la distance critique (7-4).

En pratique, le générateur d'horloge est généralement fixe, alors que l'entrée d'horloge d'un CAN peut avoir plusieurs formats et niveaux: simple, différentiel, 2.5V, 5V, ... Ainsi, un conditionnement du signal d'horloge est souvent nécessaire. Il faut veiller à ce que le jitter supplémentaire introduit par cette interface reste faible. Pour cela, il faut privilégier les interfaces passives. Par exemple, la conversion d'un signal simple en un signal différentiel peut se faire à l'aide d'un transformateur RF large bande. Lorsque le générateur d'horloge est de format ECL, il existe sur le marché des composants actifs à faible jitter permettant de convertir dans les autres formats, tels que le CMOS et TTL.



**Figure 7-8:** Convertisseur d'horloge PECL 5V en horloge CMOS 2.5V.

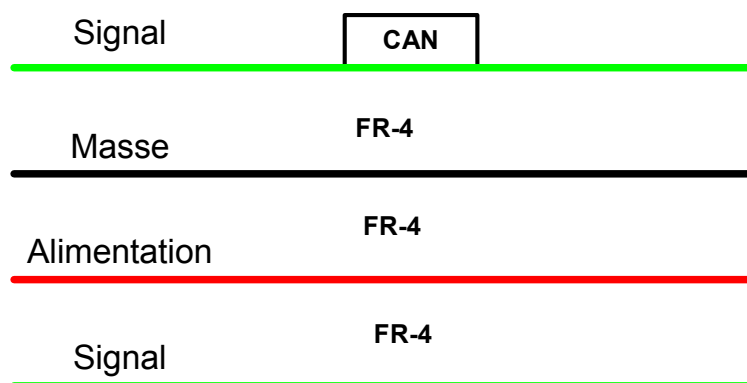
La figure (7-8) montre une interface permettant de convertir une horloge PECL différentielle 5V en une horloge CMOS 2.5V. Ce circuit est basé sur le composant MC100EPT21 de chez OnSemi [MC100EPT21] qui a un jitter inférieur à 1ps. Ainsi, si le générateur d'horloge possède un jitter RMS de 1.5ps, le jitter RMS à la sortie du circuit vaut  $\sqrt{1^2 + 1.5^2} = 1.8\text{ps}$ . Les condensateurs d'entrée permettent de bloquer la composante continue qui est de 3.5V, et les deux résistances 50Ω jouent un double rôle: d'une part réduire l'excursion du signal d'entrée de 1.6V à 800mV pour être compatible avec le composant MC100EPT21, et d'autre part terminer les deux lignes de transmission. La sortie VBB, qui est de 2V, permet de re-polariser le signal d'entrée. A la sortie, la résistance 43Ω permet, avec la résistance série du MC100EPT1 qui de l'ordre de 7Ω, de terminer la ligne en série. Plus proche de l'entrée d'horloge du CAN, on retrouve une diode de Schottky rapide de tension de seuil d'environ 400mV. Cette diode, connectée à une tension 2.1V, permet d'écarter le signal 3.3V du MC100EPT21 pour donner un signal d'amplitude 2.5V. Elle joue également le rôle d'une terminaison non linéaire pour la ligne. Enfin, la tension d'alimentation est découplée en haute et basse fréquence par deux condensateurs en parallèles. Cette opération est extrêmement

importante, puisque tout bruit sur l'alimentation affecte directement le jitter du composant. Finalement, ce circuit est une excellente solution pour générer un signal d'horloge carré à partir d'un signal d'entrée sinusoïdal. Pour cela, il suffit connecter le générateur sinusoïdal à l'entrée (+) du circuit et de laisser l'entrée complémentaire connectée uniquement à la tension VBB.

## 6. Empilement des couches et choix du diélectrique

Les cartes de test des CANs rapides possèdent quatre couches, voire plus. L'empilement de ces couches est critique. Pour une carte quatre couches, l'empilement de la figure (7-9) est bien adapté.

Cette structure offre deux plans pour les signaux, un plan de masse, et un plan d'alimentation. Ses avantages sont: l'accessibilité des signaux pour le debug, la possibilité de réaliser des lignes de transmission, et le découplage haute fréquence par la proximité des plans de masse et d'alimentation. Cependant, cette structure ne possède pas de couche signal interne, et ne permet donc pas de réaliser des pistes "stripline". Pour cela, il faut augmenter le nombre de couches et donc le coût de la carte de test.



**Figure 7-9:** Structure d'une carte 4 couches avec un diélectrique FR-4.

Le diélectrique FR-4 est largement satisfaisant pour les cartes de test destinées aux CANs. En outre, le coût de ce diélectrique reste raisonnable.

## **7. Conclusion**

Le CAN est un composant très sensible aux perturbations du matériel de test. Ce chapitre traite d'un certain nombre de considérations dont il faut tenir lors de la réalisation des circuits de conditionnement des signaux destinés aux différentes parties du CAN, à savoir: le filtrage des alimentations et du signal de test, et l'intégrité du signal d'horloge et des données numériques. Un exemple complet de conditionnement d'horloge à faible jitter a été traité. Enfin, nous avons vu qu'une carte quatre couches avec un diélectrique FR-4 constitue une solution adaptée en termes de performance et de coût pour le test des CANs.

Dans le chapitre suivant, nous reviendrons sur les limitations de l'instrumentation de test, et plus précisément nous proposerons une méthode permettant de compenser le bruit de phase inhérent aux générateurs de signaux.

**CHAPITRE 8**

**COMPENSATION DU BRUIT DE  
PHASE DU GENERATEUR DE TEST**





## CHAPITRE 8

### COMPENSATION DU BRUIT DE PHASE DU GENERATEUR DE TEST

#### 1. Introduction

La pureté spectrale du signal de test requise par les CANs rapides haute résolution, se situe à la limite des performances des générateurs de signaux disponible sur le marché. Nous avons vu dans le chapitre précédent que la distorsion harmonique et le bruit hors bande passante peuvent être filtrés par un filtre passe bande passif. La largeur de la bande passante de ces filtres étant large, le bruit de phase du générateur, qui est très proche de la porteuse, reste infiltré. Cela pose des problèmes lors du test de produits hautes performances tel qu'un 14 bits, voire un 12 bits.

Un certain nombre de méthodes ont été développées dans la littérature afin de pallier au problème de la limitation de l'instrumentation de test. Dans [Setty et al., 1990] un algorithme permettant de compenser l'effet de la distorsion du générateur du signal de test a été décrit. Une alternative à cette méthode a été proposée dans [Benkais et al., 1995], qui consiste à estimer les paramètres single-tone en se plaçant en mode dual-tone. Cependant, ces deux méthodes font un certain nombre d'hypothèses sur la fonction de transfert du CAN, qu'il faut satisfaire pour espérer avoir des résultats corrects. Enfin, une technique permettant de compenser l'effet du jitter introduit par l'environnement de test a été proposée dans [Cauvet et al., 2001]. Son inconvénient est la complexité hardware

supplémentaire qu'elle rajoute, dans la mesure où elle est basée sur le test simultané de deux CANs.

Dans ce chapitre, nous proposons une méthode de traitement du signal permettant de compenser le bruit de phase du générateur du signal de test.

## 2. Le bruit de phase

Le signal sinusoïdal généré par un synthétiseur de signaux peut s'écrire sous la forme analytique suivante:

$$s(t) = (A_0 + \varepsilon(t))\cos(2\pi f_0 t + \phi(t)) \quad (8-1)$$

avec  $\varepsilon(t)$  et  $\phi(t)$  les fluctuations d'amplitude et de phase (bruit de phase). L'effet du bruit d'amplitude est négligeable. En revanche, le bruit de phase est significatif, notamment au voisinage immédiat de la porteuse. Pour définir le bruit de phase, on considère une fluctuation de phase contenant une seule raie de fréquence  $f_m$ :

$$\phi(t) = \Delta\varphi \sin(2\pi f_m t) \quad (8-2)$$

où  $\Delta\varphi$  est la variation maximale de phase. En utilisant un développement du premier ordre des fonctions de Bessel de premier espèce, on obtient:

$$s(t) \approx A_0 \cos(2\pi f_0 t) + \frac{A_0 \Delta\varphi}{2} \cos(2\pi(f_0 + f_m)t) - \frac{A_0 \Delta\varphi}{2} \cos(2\pi(f_0 - f_m)t) \quad (8-3)$$

Ainsi, nous avons trois termes: le premier terme correspond à la porteuse, et les deux derniers traduisent la contribution du bruit de phase. Les deux raies du bruit de phase ont la même amplitude et se situent à une fréquence d'offset  $f_m$  par rapport à la porteuse (cf. figure (8-1)).

Le bruit de phase en bande latérale unique (single side bande (SSB)) ramené dans une bande de 1 Hz est défini par la fonction:

$$L(f_m) \left[ \frac{\text{dBc}}{\text{Hz}} \right] = 10 \log \left( \frac{P_{\text{SSB}}}{P_0} \right) = 10 \log \left( \frac{\left( \frac{A_0 \Delta\varphi}{2} \right)^2}{A_0^2} \right) = 10 \log \left( \left( \frac{\Delta\varphi(f_m)}{2} \right)^2 \right) \quad (8-4)$$

où  $P_{SSB}$  est la puissance du bruit en bande latérale unique, et  $P_0$  la puissance de la porteuse.

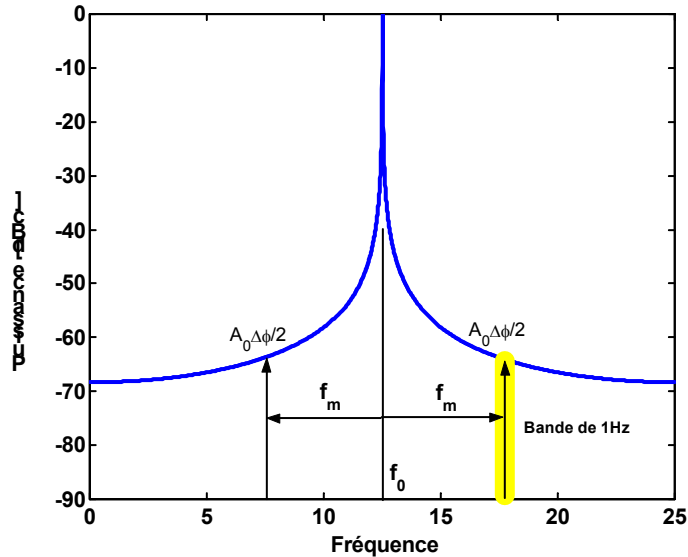


Figure 8-1: Définition du bruit de phase.

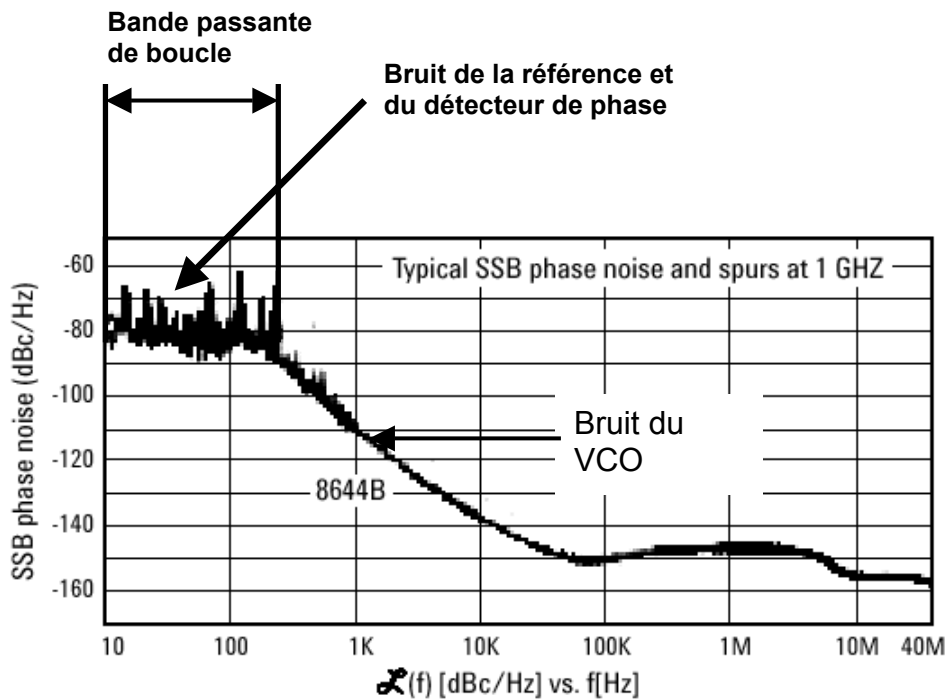


Figure 8-2: Bruit de phase du générateur HP8644B de chez Agilent.

La figure (8-2) montre le bruit de phase du générateur HP8644B en fonction de la fréquence d'offset par rapport à une porteuse de 1GHz. On verra que ce bruit pourrait dégrader les performances dynamiques de CANs hautes performances. Le bruit issu de l'oscillateur est divisé par la chaîne des diviseurs et apparaît sur le détecteur de phase. Les perturbations qu'il génère sont relativement faibles. Lorsque le bruit atteint le filtre de boucle, seules les fréquences inférieures à la fréquence de coupure apparaissent sur la ligne de contrôle de l'oscillateur asservi en fréquence. Il peut être atténué ou même éliminé. Le bruit hors bande reste par contre inchangé. En revanche, le bruit de l'oscillateur de référence entre dans le détecteur de phase, atteint le filtre de boucle et s'additionne au signal de sortie. Le bruit généré par la référence et le détecteur de phase est généralement faible mais sa contribution est critique pour les performances de la boucle.

### 3. Effet du bruit de phase sur les performances d'un CAN

La résolution en fréquence du spectre des données d'un CAN est égale à  $f_s/M$ , où  $f_s$  est la fréquence d'échantillonnage et  $M$  le nombre d'échantillons. Pour  $f_s = 20\text{MHz}$  et  $M = 65536$ , on obtient une résolution d'environ 305Hz. D'après la figure (8-2), cela correspond à un niveau de bruit d'environ 90dBc. Ce bruit peut facilement dégrader les performances d'un CAN 14 bits voire 12 bits, notamment le SFDR, le rapport signal-à-bruit, et la précision des mesures.

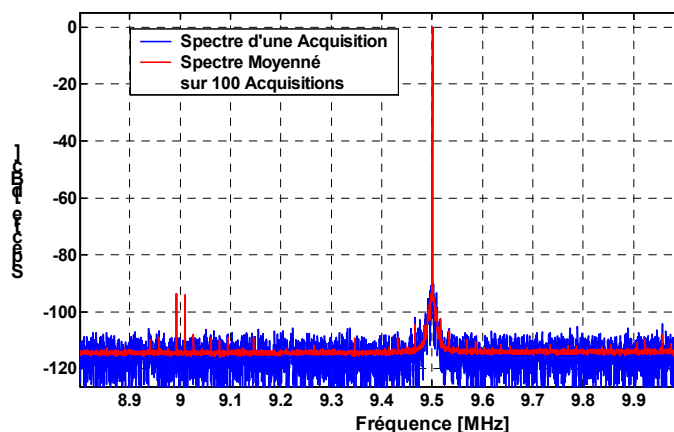


Figure 8-3: Spectres d'un CAN 12 bits réel avec:  $f_s=20\text{MHz}$ ,  $f_c=10.5\text{MHz}$ ,  $M=65536$ .

Ce phénomène est illustré sur la figure (8-3) où nous donnons le spectre autour de la porteuse d'un CAN 12 bits réel, testé par le générateur HP8644B avec un signal de fréquence 10.5MHz. Comme  $f_s = 20\text{MHz}$  et  $M = 65536$ , la résolution en fréquence est alors de 305Hz. On fait bien apparaître le bruit de phase du générateur qui limite le SFDR.

On pourrait légitimement penser qu'il suffit de ne pas tenir compte des raies situées dans la bande étroite autour de la porteuse, pour éliminer l'effet de ce bruit sur les paramètres dynamiques du CAN. Mais, cette zone pourrait contenir des raies de distorsion ou parasites intrinsèques au CAN, et procéder de cette façon reviendrait à surestimer les performances du CAN.

Dans le paragraphe suivant, nous décrivons une méthode permettant de compenser efficacement ce bruit.

#### **4. Compensation du bruit de phase**

Pour pouvoir compenser le bruit de phase il faudrait l'estimer avec précision. Pour ce faire, on procède en deux étapes:

1. La dispersion du bruit additif est minimisée en moyennant les spectres de puissance sur 100 acquisitions.
2. Le spectre en puissance moyenné est ensuite lissé par filtrage numérique.

Comme le montre la figure (8-3), le fait de moyenner le spectre réduit considérablement la variance du bruit. C'est pourquoi cette technique est souvent utilisée pour améliorer la précision d'estimation de la distorsion harmonique. Cependant, il persiste sur le spectre des raies de distorsion et des raies parasites ainsi qu'une certaine dispersion du bruit. Le filtrage numérique permet d'éliminer ces effets.

#### 4.1. Spectre en puissance moyenné

Cette méthode est connue dans la littérature sous le nom de périodogramme moyenné [Kay, 1988]. Il est montré dans cette référence que le fait de moyenné sur K acquisitions indépendantes réduit la variance du bruit d'un facteur K. Plus précisément, le spectre possède  $\beta\%$  de chance de se trouver dans l'intervalle:

$$\left[ \frac{Y_m(f)}{1+\gamma} \quad \frac{Y_m(f)}{1-\gamma} \right] \text{ avec } \gamma = \sqrt{\frac{2}{K}} \operatorname{erf}^{-1}(\beta) \quad (8-5)$$

où  $Y_m$  est le spectre en puissance moyenné,  $f$  la fréquence, et  $\operatorname{erf}^{-1}$  la fonction inverse de la fonction erreur définie par la relation suivante:

$$\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (8-6)$$

Pour  $\beta = 95\%$  et  $K = 30$ , on obtient en décibels l'intervalle de confiance  $[Y_m(f) - 1.33 \quad Y_m(f) + 1.92]$ . Ainsi, les pics de plus de quelques dB doivent être considérés comme faisant partie du spectre et non le résultat des fluctuations statistiques.

Dans [Durnerin, 1999], plusieurs techniques, plus ou moins complexes, d'estimation de fond de spectre ont été étudiées et comparées. Nous avons trouvé qu'un simple filtre médian suffisait largement pour la résolution de notre problème.

#### 4.2. Le filtrage médian

Le principe d'un filtre médian est très simple. Si on note  $Y_m$  le spectre en puissance moyenné et  $n$  l'ordre du filtre, alors le spectre filtré  $Y_f$  est donné par la relation suivante:

$$\{Y_f(k)\}_{0 \leq k \leq M/2-1} = \text{médiane} \begin{cases} \left[ Y_m\left(k - \frac{(n-1)}{2}\right), \dots, Y_m\left(k + \frac{(n-1)}{2}\right) \right] & \text{si } n \text{ impair} \\ \left[ Y_m\left(k - \frac{n}{2}\right), \dots, Y_m\left(k + \frac{n}{2} - 1\right) \right] & \text{si } n \text{ pair} \end{cases} \quad (8-7)$$

Notons que, puisque le spectre est symétrique, il suffit d'effectuer le filtrage sur la moitié des points.

Nous avons représenté sur la figure (8-3) le résultat du filtrage du spectre moyenné avec un filtre médian d'ordre 11. On voit bien la qualité d'estimation du fond de spectre par comparaison avec le spectre d'une seule acquisition. Pour la compensation nous aurons besoin de la valeur du plancher de bruit (NF). Le calcul de ce dernier s'effectue, comme le montre la figure, sur une fenêtre suffisamment large, mais pas nécessairement sur tout le spectre. Cela permettra de ne pas tenir compte des éventuels défauts locaux persistants, comme le pic qui apparaît entre 0 et 2MHz.

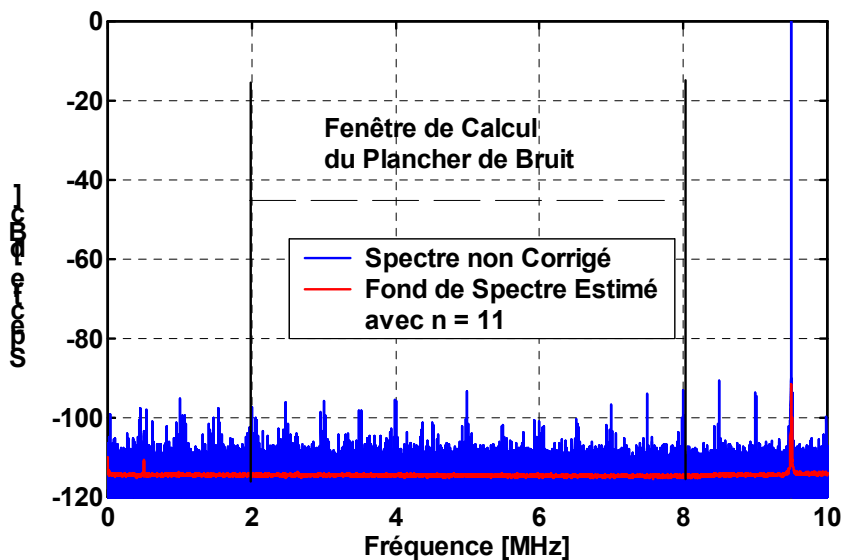


Figure 8-4: Résultat d'un filtre médian d'ordre 11.

Nous avons étudié l'effet de l'ordre du filtre médian sur l'estimation du bruit de phase et du plancher de bruit. Les résultats obtenus sont donnés dans le tableau (8-1) et la figure (8-5). Nous avons également donné le SNR déduit du plancher de bruit par la relation:  $SNR = -NF - 10\log\left(\frac{M}{2}\right)$ , NF étant le plancher de bruit. On constate que l'ordre du filtre n'a pas d'effet significatif sur la valeur du plancher de bruit. Pour l'estimation du bruit de phase, on voit que la qualité du lissage augmente avec l'ordre du filtre, pour se dégrader au voisinage immédiat de la porteuse pour des ordres élevés. En conclusion, des filtres d'ordre compris entre 7 et 13 conviennent parfaitement.



Tableau 8-12 : Valeurs du plancher de bruit en fonction de l'ordre du filtre médian.

Ordre du filtre (n)	5	7	9	11	13	15
Plancher de Bruit (NF) [dBc]	-114.491	-114.494	-114.495	-114.496	-114.496	-114.497
SNR [dB]	69.336	69.339	69.341	69.342	69.342	69.342

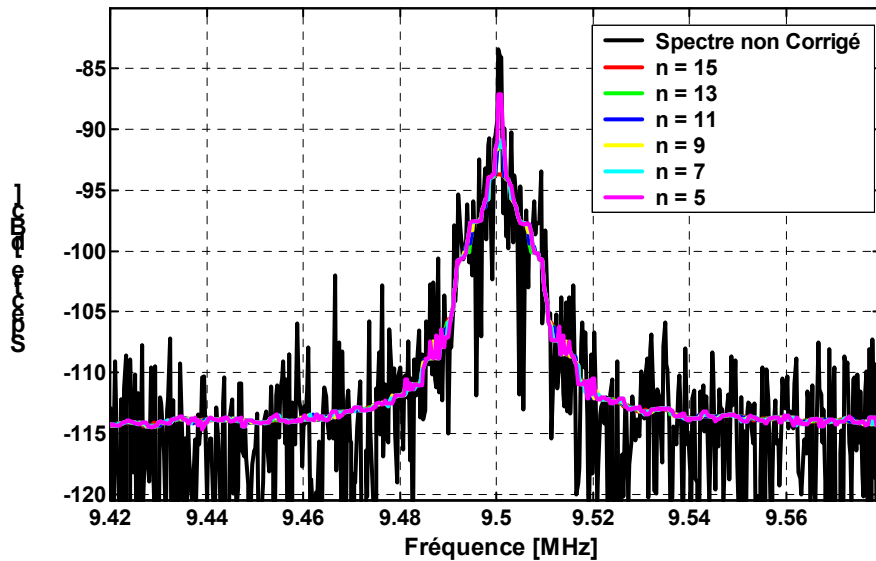


Figure 8-5: Effet de l'ordre du filtre médian sur l'estimation du bruit de phase.

### 4.3. Compensation

Le plancher de bruit et le bruit de phase étant estimés, on peut maintenant les utiliser pour corriger le spectre en puissance  $Y$  d'une seule acquisition au voisinage de la porteuse. Cette opération s'effectue de la façon suivante:

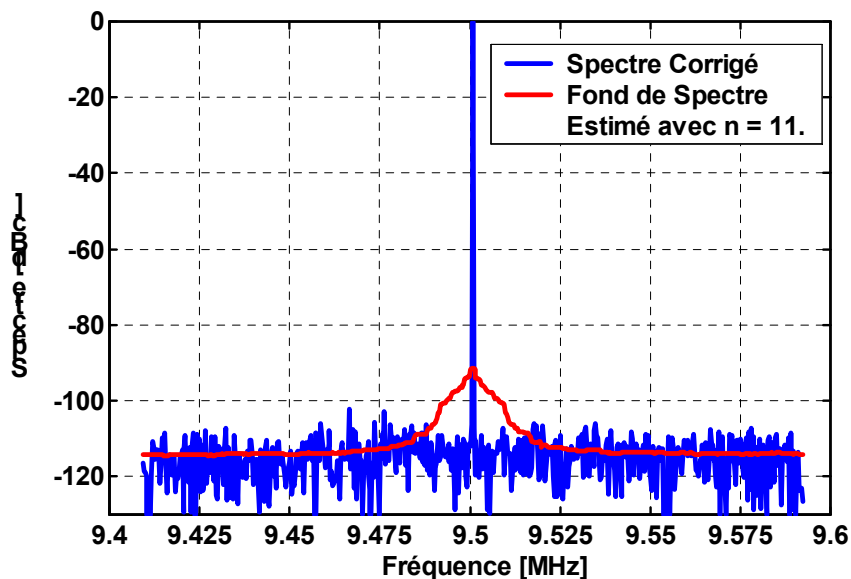
$$Y[k] = \frac{NF \times Y[k]}{Y_f[k]} \quad n_f - \Delta n \leq k \leq n_f + \Delta n \quad (8-8)$$

où  $n_f$  est l'indice de la porteuse et  $\Delta n$  la mi-largeur de la bande du bruit de phase. Il faut tenir compte du phénomène du repliement spectral lors du calcul de  $n_f$ . Le paramètre  $\Delta n$  peut être calculé automatiquement, en utilisant le fait que  $Y_f[n_f + \Delta n] = NF$ . Notons que les différentes grandeurs de l'équation (8-8) ne sont pas exprimées en décibel.

En pratique, puisque le bruit de phase est reproductible, il suffit de sauvegarder dans un fichier les valeurs du plancher du bruit et de bruit de phase qui viennent d'être estimées. Ensuite, le programme de test calibre chaque acquisition obtenue dans les mêmes conditions de test, à l'aide de l'équation (8-8). Par conséquent, cette méthode n'introduit aucune complexité supplémentaire, si ce n'est la lecture d'une centaine de données dans un fichier. Cette opération s'effectue une seule fois lors du lancement du programme.

## 5. Résultats

Nous avons appliqué la méthode décrite ci-dessus pour le test industriel d'un CAN 12 bits. Le résultat obtenu en mode single-tone est représenté sur la figure (8-6).



**Figure 8-6:** Bruit de phase estimé et spectre corrigé d'un CAN réel 12 bits.  
Paramètres:  $f_s = 20\text{MHz}$ ,  $f_e = 10.5\text{MHz}$ ,  $M = 65536$ .

On voit bien la qualité d'estimation du bruit de phase et du plancher de bruit, ainsi que celle de la compensation du spectre d'une acquisition simple. Les résultats avant correction et après l'opération de moyennage sont donnés sur la figure (8-3).

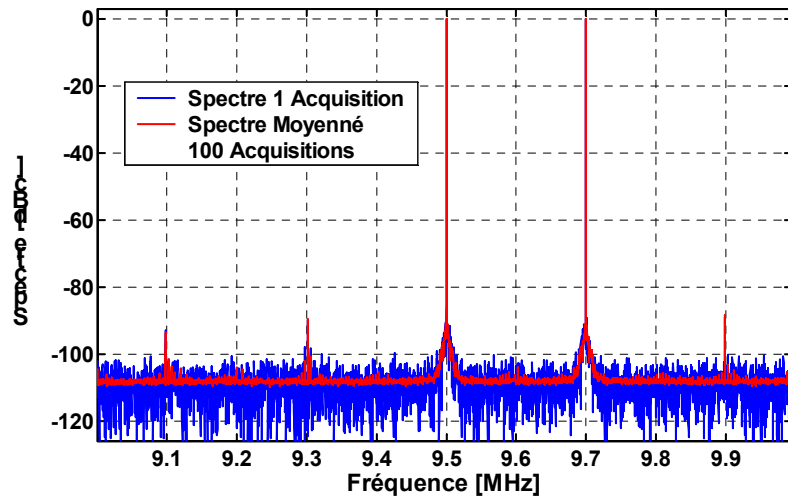


Figure 8-7: Spectres simple et moyenné d'un CAN réel 12 bits en mode dual-tone. Paramètres:  $f_s = 20\text{MHz}$ ,  $f_1 = 9.7\text{MHz}$ ,  $f_2 = 10.5\text{MHz}$ ,  $M = 65536$ .

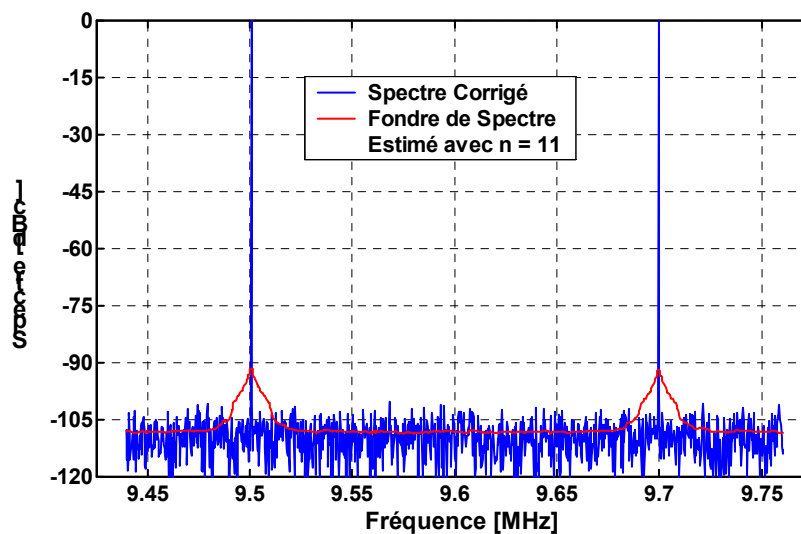


Figure 8-8: Spectres après le filtrage médian et la compensation.

Sur les figures (8-7) et (8-8) nous représentons les résultats obtenus en mode dual-tone. En effet, bien que nous ayons illustré la méthode par des exemples en mode single-tone, le même problème se pose en mode dual-tone. La situation devient critique lorsque les deux fréquences de test sont très proches l'une de l'autre. Dans ce cas, les raies de distorsion d'intermodulation d'ordre trois  $|2f_1 - f_2|$  et  $|2f_2 - f_1|$  se retrouve au voisinage immédiat des porteuses, et elles se

combinent donc directement avec le bruit de phase. Comme en général la contribution de ces raies est importante, les paramètres IMD et SFDR seront forcément grandement entachés d'erreurs.

La figure (8-7) montre une acquisition simple ainsi que le résultat du moyennage sur 100 acquisitions. La figure (8-8) montre le résultat du filtrage médian et le spectre d'une acquisition après compensation du bruit de phase.

## 6. Conclusion

Dans ce chapitre nous avons présenté une méthode simple mais efficace pour la compensation du bruit de phase inhérent aux générateurs de signaux de test. Ce problème se présente lors du test de CANs rapides de résolution supérieure à 12 bits, avec une résolution fréquentielle suffisamment précise pour faire apparaître le bruit de phase autour de la porteuse. Cette méthode n'introduit ni complexité supplémentaire ni augmentation du temps de test. Son efficacité a été validée sur des données réelles, et ce en mode single-tone comme en mode dual-tone.

Cette technique pourrait également être utilisée pour compenser tout phénomène répétitif et à bande spectrale étroite, introduit par la carte ou l'instrumentation de test. Cela évitera de refaire par exemple la carte de test, étant donné le coût que cela implique. C'est aussi une méthode très pratique pour évaluer le bruit de phase du générateur de test. Pour cela, il suffit d'augmenter la résolution fréquentielle, en baissant la fréquence d'échantillonnage et en augmentant le nombre de points. L'information sur le bruit de phase pourra être utilisée pour comparer plusieurs instruments ou encore évaluer le nombre d'échantillons nécessaires à l'analyse statistique [**dynad, 2001**].



**CONCLUSION**



## CONCLUSION

La caractérisation des convertisseurs analogique-numérique par l'analyse temporelle présente un certain nombre d'avantages tels que l'insensibilité au phénomène de leakage ou encore la précision d'estimation. Cependant, de par son principe, cette approche possède deux inconvénients majeurs, à savoir la convergence et la complexité d'implantation. Cela explique pourquoi son utilisation dans la littérature est limitée en général à l'estimation du nombre effectif de bits. Sur le plan instrumentale, la mise en œuvre des différents tests et les limitations de l'instrumentation posent également des problèmes qui ne sont pas toujours faciles à surmonter.

L'objectif de cette étude était double:

- D'une part, étendre la couverture de test de l'analyse temporelle pour estimer la distorsion harmonique, le rapport signal à bruit, la distorsion d'intermodulation, et la dynamique de la distorsion. C'était l'objet de la première partie.
- D'autre part, implanter industriellement les méthodes de test dynamiques et aborder le problème des perturbations introduites par le matériel de test. Ce thème a été abordé lors de la deuxième partie.

La définition des paramètres pouvant être estimés par l'analyse temporelle a été faite au premier chapitre. D'après les équations définissant ces paramètres, leur calcul revient essentiellement à estimer les paramètres d'un modèle d'un certain nombre de sinusoides noyées dans un bruit. Ce modèle est linéaire par rapport à un certain nombre de paramètres et non linéaire par rapport à d'autres. Le modèle est estimé par la méthode des moindres carrés présentée au deuxième chapitre. Il a été montré que ce critère est identique à celui du maximum de vraisemblance. Dans ce chapitre, deux notions fondamentales ont été introduites: les bornes de Cramer-Rao et la méthode de séparation des variables. Les bornes de Cramer-



Rao servent à juger les performances intrinsèques des différents algorithmes étudiés, alors que la méthode de séparation des variables a permis une simplification considérable de la complexité du problème d'estimation, en séparant l'estimation des paramètres linéaires des paramètres non linéaires du modèle.

L'estimation des paramètres en mode single-tone a été discutée dans les chapitres 3 et 4. L'analyse de la convergence des différents algorithmes nous a permis de développer une méthode robuste à l'égard des valeurs initiales. L'efficacité des différents algorithmes a été validée par des simulations de Monte-Carlo. L'application des méthodes développées à la caractérisation de CANs réels a montré un très bon accord avec les résultats de l'analyse spectrale, avec comme attendu, un net avantage pour l'analyse spectrale en termes de temps de test. Cependant, le temps de test de l'analyse temporelle reste parfaitement adapté au test en laboratoire et à l'engineering. Dans le cinquième chapitre, nous avons étendu l'analyse temporelle à l'estimation des paramètres dynamiques en mode dual-tone. Comme en mode single-tone, les simulations de Monte-Carlo et les résultats réels ont montré la précision de la méthode proposée. Son temps de test reste aussi adapté à la caractérisation et au debug.

Les résultats de la première partie de cette étude ont fait l'objet d'un certain nombre de publications [**Haddadi et al., 1999**], [**Haddadi et al., May 2000**], [**Haddadi et al., July 2000**], [**Haddadi et al., Sept. 2000**], [**Haddadi et al., Nov. 2000**].

Le test industriel est extrêmement exigeant sur le temps de test et la répétabilité. Il a fallu donc simplifier l'analyse temporelle pour pouvoir l'implanter en test industriel. Ce travail a été l'objet du sixième chapitre de la deuxième partie. Cette simplification a été possible en se plaçant dans l'hypothèse où la condition de cohérence est satisfaite. En effet, en test industriel on utilise tous les moyens pour simplifier et réduire le temps de l'industrialisation des produits. C'est pourquoi on dispose généralement d'une instrumentation permettant de satisfaire la condition de cohérence avec une précision suffisante. L'étude a permis de quantifier la différence de performances entre les analyses spectrale et temporelle. En termes de temps de test, on obtient, comme attendu, la supériorité de l'analyse spectrale

---

d'un facteur 43 en mode single-tone et 82 en mode-dual-tone. Nous avons constaté que l'analyse temporelle a des temps de test adaptés aux convertisseurs de résolution inférieure ou égale à 10 bits. Enfin, des modèles simples et précis ont été proposés pour quantifier l'évolution du temps de test en fonction du nombre d'échantillons. En ce qui concerne la répétabilité, l'étude a montré un très bon accord entre les deux méthodes.

Les problèmes liés au matériel de test ont été abordés dans les deux derniers chapitres. Avec l'augmentation permanente des performances des produits, la réalisation de la carte de test devient de plus en plus difficile et lente. Dans le septième chapitre, nous avons revu un certain nombre de précautions dont il faut tenir compte lors de la conception des circuits de conditionnement des éléments suivants: alimentations, signal de test, données numériques, et signal d'horloge. Ces considérations permettent de minimiser les perturbations inévitables provenant de la carte de test. L'empilement des couches de la carte ainsi que le choix du diélectrique sont des éléments importants pour les cartes rapides. Nous avons vu qu'une carte quatre couches avec FR-4 comme diélectrique est un bon choix, en terme de coût et de performance pour le test des convertisseurs rapides. Les autres perturbations proviennent des limitations inhérentes à l'instrumentation de test. Ce point a été abordé dans le dernier chapitre, où nous avons proposé une méthode simple et efficace pour compenser le bruit de phase du générateur de test. Ce problème se manifeste lors du test des convertisseurs rapides hautes performances tel qu'un 14 bits. Cette méthode a été implantée avec succès en test industriel. Son efficacité a été illustrée sur un CAN 12 bits hautes performances que ce en mode single-tone ou en mode dual-tone.

Les méthodes de test des convertisseurs analogiques-numériques ont fait l'objet d'une littérature abondante. Les standards [**dynad, 2001**] et [**ieeestd1241, 2001**] constituent à notre sens une synthèse définitive dans ce domaine. Avec l'augmentation des performances des produits, les défis se situent actuellement au niveau de la réalisation de la carte de test et des limitations de l'instrumentation de test.

Les technologies avancées permettent d'améliorer la vitesse et les performances des produits. Le passage d'une technologie à une autre se fait quand l'architecture atteint ses limites ou lorsque l'application l'exige. Or, on sait que les technologies avancées ont des temps de montée faibles, ce qui pose des problèmes d'intégrité de signal qu'il faudra gérer sur la carte de test. Les problèmes liés à l'instrumentation de test sont essentiellement le jitter pour le générateur d'horloge et la pureté spectrale pour le signal de test. Deux autres domaines de recherche sont en plein essor. Le premier concerne la modélisation des différentes architectures des CANs. Son intérêt réside notamment dans la réduction du temps de test industriel [**Cherubal et al.**]. Le deuxième porte sur l'intégration du test au sein même des produits mixtes (BIST) [**Rebai, 2002**]. Contrairement aux systèmes numériques, la tâche n'est pas aisée pour les produits mixtes, et à l'heure actuelle aucun CAN rapide avec du BIST n'est disponible sur le marché. Ces travaux visent à simplifier le test des produits mixtes qui devient de plus en plus difficile et coûteux.

**ANNEXE A**



## ANNEXE A

### A.1. Dérivation des bornes de Cramer-Rao

L'objectif de cette partie de l'annexe est, d'une part, de dériver les expressions asymptotiques des bornes de Cramer-Rao en mode single-tone en considérant un modèle avec offset et, d'autre part, de dériver les bornes de Cramer-Rao en mode dual-tone.

#### A.1.1. Mode Single-tone [Nehorai, 1986] [James et al., 1994]

En mode single-tone le modèle du signal harmonique peut s'écrire sous la forme suivante:

$$\begin{aligned} \{s_n\}_{0 \leq n \leq M-1} &= \sum_{k=1}^P A_k \sin(kn\omega + \varphi_k) + C \\ &= \sum_{k=1}^P a_k \cos(kn\omega) + \sum_{k=1}^P b_k \sin(kn\omega) + C \end{aligned} \quad (\text{A-1})$$

avec

$$\begin{aligned} a_k &= A_k \sin(\varphi_k) \\ b_k &= A_k \cos(\varphi_k) \end{aligned} \quad (\text{A-2})$$

Le vecteur des paramètres du modèle est donné par:

$$\mathbf{x} = [\omega \ a_1 \ \dots \ a_p \ b_1 \ \dots \ b_p \ C]^T \quad (\text{A-3})$$

L'expression générale de l'élément (k, l) de la matrice d'information de Fisher est donnée par:

$$F_{k,l} = \frac{1}{\sigma^2} \sum_{n=0}^{M-1} \frac{\partial s_n}{\partial x_k} \frac{\partial s_n}{\partial x_l} \quad (\text{A-4})$$

On note par  $0(x)$  une quantité asymptotiquement linéaire en  $x$  :

$$\lim_{x \rightarrow \infty} \left| \frac{0(x)}{x} \right| = \text{constante} \quad (\text{A-5})$$

et  $0(1)$  une quantité bornée. Alors, pour tout  $k\omega$  tel que  $k\omega/2\pi$  n'est pas entier et  $M \gg 1/\omega$ , on a :

$$\sum_{n=0}^{M-1} n^i \sin(kn\omega) = 0(M^i); \quad \sum_{n=0}^{M-1} n^i \cos(kn\omega) = 0(M^i) \quad i = 0, 1, 2, \dots \quad (\text{A-6})$$

On peut montrer que :

$$\sum_{n=0}^{M-1} \left( \frac{\partial \mathbf{s}_n}{\partial \omega} \right)^2 = \frac{M^3}{6} \left[ \sum_{k=1}^P k^2 (a_k^2 + b_k^2) \right] + 0(M^2) \quad (\text{A-7})$$

$$\sum_{n=0}^{M-1} \frac{\partial \mathbf{s}_n}{\partial \omega} \frac{\partial \mathbf{s}_n}{\partial a_l} = \frac{M^2}{4} l b_l + 0(M) \quad (\text{A-8})$$

$$\sum_{n=0}^{M-1} \frac{\partial \mathbf{s}_n}{\partial \omega} \frac{\partial \mathbf{s}_n}{\partial b_l} = -\frac{M^2}{4} l a_l + 0(M) \quad (\text{A-9})$$

$$\sum_{n=0}^{M-1} \left( \frac{\partial \mathbf{s}_n}{\partial a_l} \right)^2 = \frac{M}{2} + 0(1) \quad (\text{A-10})$$

$$\sum_{n=0}^{M-1} \frac{\partial \mathbf{s}_n}{\partial a_l} \frac{\partial \mathbf{s}_n}{\partial a_m} = 0(1) \quad (\text{A-11})$$

$$\sum_{n=0}^{M-1} \frac{\partial \mathbf{s}_n}{\partial a_l} \frac{\partial \mathbf{s}_n}{\partial b_m} = 0(1) \quad (\text{A-12})$$

$$\sum_{n=0}^{M-1} \left( \frac{\partial \mathbf{s}_n}{\partial b_l} \right)^2 = \frac{M}{2} + 0(1) \quad (\text{A-13})$$

$$\sum_{n=0}^{M-1} \frac{\partial \mathbf{s}_n}{\partial b_l} \frac{\partial \mathbf{s}_n}{\partial b_m} = 0(1) \quad (\text{A-14})$$

$$\sum_{n=0}^{M-1} \left( \frac{\partial \mathbf{s}_n}{\partial \mathbf{C}} \right)^2 = M \quad (\text{A-15})$$

$$\sum_{n=0}^{M-1} \frac{\partial \mathbf{s}_n}{\partial \omega} \frac{\partial \mathbf{s}_n}{\partial \mathbf{C}} = 0(M) \quad (\text{A-16})$$

$$\sum_{n=0}^{M-1} \frac{\partial \mathbf{s}_n}{\partial a_l} \frac{\partial \mathbf{s}_n}{\partial \mathbf{C}} = 0(1) \quad (\text{A-17})$$

$$\sum_{n=0}^{M-1} \frac{\partial s_n}{\partial b_1} \frac{\partial s_n}{\partial C} = 0(1) \quad (\text{A-18})$$

Ainsi, la matrice d'information de Fisher peut s'écrire sous la forme suivante:

$$\mathbf{F} = \frac{1}{2\sigma^2} (\mathbf{F}_1 + \mathbf{F}_2) \quad (\text{A-19})$$

avec

$$\mathbf{F}_1 = \begin{bmatrix} \gamma \frac{M^3}{3} & \frac{M^2}{2} \boldsymbol{\beta}^T & -\frac{M^2}{2} \boldsymbol{\alpha}^T & 0 \\ \frac{M^2}{2} \boldsymbol{\beta} & M\mathbf{I}_P & \mathbf{0}_P & \mathbf{0} \\ -\frac{M^2}{2} \boldsymbol{\alpha} & \mathbf{0}_P & M\mathbf{I}_P & \mathbf{0} \\ 0 & \mathbf{0}^T & \mathbf{0}^T & 2M \end{bmatrix} \quad (\text{A-20})$$

et

$$\mathbf{F}_2 = \begin{bmatrix} 0(M^2) & 0(M)\mathbf{1}^T & 0(M)\mathbf{1}^T & 0(M) \\ 0(M)\mathbf{1} & 0(1)\mathbf{I}_P & 0(1)\mathbf{1}_P & 0(1)\mathbf{1} \\ 0(M)\mathbf{1} & 0(1)\mathbf{1}_P & 0(1)\mathbf{I}_P & 0(1)\mathbf{1} \\ 0(M) & 0(1)\mathbf{1}^T & 0(1)\mathbf{1}^T & 0 \end{bmatrix} \quad (\text{A-21})$$

avec

$$\begin{aligned} \boldsymbol{\alpha} &= [a_1 \quad 2a_2 \quad \dots \quad Pa_P]^T \\ \boldsymbol{\beta} &= [b_1 \quad 2b_2 \quad \dots \quad Pb_P]^T \\ \gamma &= \sum_{k=1}^P k^2 (a_k^2 + b_k^2) \\ \mathbf{1} &= \underbrace{[1 \quad 1 \quad \dots \quad 1]^T}_{P \text{ éléments}} \\ \mathbf{0} &= \underbrace{[0 \quad 0 \quad \dots \quad 0]^T}_{P \text{ éléments}} \end{aligned} \quad (\text{A-22})$$

$\mathbf{I}_P$  la matrice identité de taille  $P \times P$  et  $\mathbf{1}_P$  la matrice unité de taille  $P \times P$ .

On peut montrer que l'inverse de la matrice  $\mathbf{F}_1$  est donnée par l'expression suivante:



$$\mathbf{F}_1^{-1} = \begin{bmatrix} \frac{12}{\gamma M^3} & -\frac{6}{\gamma M^2} \boldsymbol{\beta}^\top & \frac{6}{\gamma M^2} \boldsymbol{\alpha}^\top & 0 \\ -\frac{6}{\gamma M^2} \boldsymbol{\beta} & \frac{1}{M} \left[ \mathbf{I}_P + \frac{3}{\gamma} \boldsymbol{\beta} \boldsymbol{\beta}^\top \right] & -\frac{3}{\gamma M} \boldsymbol{\beta} \boldsymbol{\alpha}^\top & \mathbf{0} \\ \frac{6}{\gamma M^2} \boldsymbol{\alpha} & -\frac{3}{\gamma M} \boldsymbol{\alpha} \boldsymbol{\beta}^\top & \frac{1}{M} \left[ \mathbf{I}_P + \frac{3}{\gamma} \boldsymbol{\alpha} \boldsymbol{\alpha}^\top \right] & \mathbf{0} \\ 0 & \mathbf{0}^\top & \mathbf{0}^\top & \frac{1}{2M} \end{bmatrix} \quad (\text{A-23})$$

A partir de (A-19), la matrice d'information de Fisher peut s'écrire sous la forme suivante [James et al., 1994]:

$$\mathbf{F}^{-1} = 2\sigma^2 \mathbf{F}_1^{-1} + 2\sigma^2 \sum_{k=1}^{\infty} (-1)^k \mathbf{F}_1^{-1} (\mathbf{F}_2 \mathbf{F}_1^{-1})^k \quad (\text{A-24})$$

On peut montrer à partir de (A-21), (A-23) et (A-24) que, pour M suffisamment grand, le deuxième terme de (A-24) est négligeable devant son premier terme [Nehorai, 1986] [James et al., 1994], et donc:

$$\mathbf{F}^{-1} \approx 2\sigma^2 \mathbf{F}_1^{-1} \quad (\text{A-25})$$

Les bornes de Cramer-Rao sont définies comme étant les éléments diagonaux de l'inverse de la matrice d'information de Fisher, et elles se déduisent facilement à partir des équations (A-23) et (A-25). La nouveauté ici par rapport à la référence [Nehorai, 1986] est la borne de Cramer-Rao de l'offset qui est donnée par l'expression suivante:

$$\text{var}(C) \geq \frac{\sigma^2}{M} \quad (\text{A-26})$$

On en déduit que la borne de l'offset ne dépend que du bruit et du nombre d'échantillons, et que sa présence n'a aucun effet sur les autres bornes données par les expressions suivantes:

$$\begin{cases} \text{var}(\omega) \geq \frac{24\sigma^2}{\gamma M^3} \\ \text{var}(a_k) \geq \frac{2\sigma^2}{M} \left( 1 + \frac{3}{\gamma} k^2 b_k^2 \right) \\ \text{var}(b_k) \geq \frac{2\sigma^2}{M} \left( 1 + \frac{3}{\gamma} k^2 a_k^2 \right) \end{cases} \quad 1 \leq k \leq P \quad (\text{A-27})$$

Pour évaluer les bornes de Cramer-Rao en terme d'amplitudes et de phases au lieu des paramètres  $a_k$  et  $b_k$ , on utilisera la relation générale suivante:

$$\mathbf{G}^{-1}(\mathbf{z}) = \left[ \frac{\partial \mathbf{z}}{\partial \mathbf{x}} \right] \mathbf{F}^{-1}(\mathbf{x}) \left[ \frac{\partial \mathbf{z}}{\partial \mathbf{x}} \right]^T \quad (\text{A-28})$$

donnant l'inverse de la matrice d'information de Fisher correspondant au vecteur de paramètres  $\mathbf{z} = [\omega \ A_1 \ \dots \ A_P \ \varphi_1 \ \dots \ \varphi_P \ C]^T$ , où l'élément  $(i, j)$  de la matrice jacobienne  $\left[ \frac{\partial \mathbf{z}}{\partial \mathbf{x}} \right]$  est  $\frac{\partial z_i}{\partial x_j}$ . Il est montré dans [Nehorai, 1986], après de simples mais longs calculs, que les bornes de Cramer-Rao pour les amplitudes et les phases sont données par les formules suivantes:

$$\begin{cases} \text{var}(A_k) \geq \frac{2\sigma^2}{M} \\ \text{var}(\varphi_k) \geq \frac{2\sigma^2}{M} \left( \frac{1}{A_k^2} + \frac{3k^2}{\gamma} \right) \end{cases} \quad (\text{A-29})$$

### A.1.2. Mode Dual-tone

L'équation (6-1) donne l'expression du modèle du signal à estimer en mode dual-tone. Ce modèle peut s'écrire après linéarisation sous la forme matricielle suivante:

$$\mathbf{s}(\boldsymbol{\Omega}) = \mathbf{H}(\boldsymbol{\Omega})\boldsymbol{\theta} \quad (\text{A-30})$$

où  $\boldsymbol{\Omega} = [\omega_1 \ \omega_2]^T$ ,  $\boldsymbol{\theta}$  le vecteur des paramètres linéaires du modèle et  $\mathbf{H}$  la matrice du signal. D'après (A-4), la matrice d'information de Fisher correspondant au vecteur  $\mathbf{x} = [\boldsymbol{\theta}^T \ \boldsymbol{\Omega}^T]^T$  est donnée par:

$$\mathbf{F}(\mathbf{x}) = \frac{1}{\sigma^2} \left( \frac{\partial \mathbf{s}}{\partial \mathbf{x}^T} \right)^T \left( \frac{\partial \mathbf{s}}{\partial \mathbf{x}^T} \right) \quad (\text{A-31})$$

Comme

$$\frac{\partial \mathbf{s}}{\partial \mathbf{x}^T} = [\mathbf{H} \ \mathbf{D}] \quad \text{avec} \quad \mathbf{D} = \begin{bmatrix} \frac{\partial \mathbf{H}}{\partial \omega_1} \boldsymbol{\theta} & \frac{\partial \mathbf{H}}{\partial \omega_2} \boldsymbol{\theta} \end{bmatrix} \quad (\text{A-32})$$

L'équation (A-31) devient:

$$\mathbf{F}(\mathbf{x}) = \frac{1}{\sigma^2} \left[ \begin{array}{c|c} \mathbf{H}^T \mathbf{H} & \mathbf{H}^T \mathbf{D} \\ \hline \mathbf{D}^T \mathbf{H} & \mathbf{D}^T \mathbf{D} \end{array} \right] \quad (\text{A-33})$$

L'inverse de  $\mathbf{F}$  s'obtient facilement en utilisant la formule d'inversion par bloc d'une matrice:

$$\mathbf{F}^{-1} = \sigma^2 \left[ \begin{array}{c|c} [\mathbf{H}^T (\mathbf{I} - \mathbf{P}_D) \mathbf{H}]^{-1} & \mathbf{Q} \\ \hline \mathbf{Q} & [\mathbf{D}^T (\mathbf{I} - \mathbf{P}_H) \mathbf{D}]^{-1} \end{array} \right] \quad (\text{A-34})$$

avec

$$\mathbf{P}_H = \mathbf{H} \mathbf{H}^+ ; \quad \mathbf{P}_D = \mathbf{D} \mathbf{D}^+$$

$$\mathbf{Q} = -[\mathbf{H}^T (\mathbf{I} - \mathbf{P}_D) \mathbf{H}]^{-1} (\mathbf{D}^+ \mathbf{H})^T$$

On en déduit que les bornes de Cramer-Rao des paramètres  $\boldsymbol{\Omega}$  et  $\boldsymbol{\theta}$  sont les éléments diagonaux des matrices suivantes:

$$\begin{aligned} \boldsymbol{\Omega} &\Rightarrow \sigma^2 [\mathbf{D}^T (\mathbf{I} - \mathbf{P}_H) \mathbf{D}]^{-1} \\ \boldsymbol{\theta} &\Rightarrow \sigma^2 [\mathbf{H}^T (\mathbf{I} - \mathbf{P}_D) \mathbf{H}]^{-1} \end{aligned} \quad (\text{A-35})$$

En simulation, on dispose des valeurs exactes des paramètres du modèle, ce qui permet d'évaluer les bornes de Cramer-Rao. En pratique, on obtient une bonne approximation des bornes de Cramer-Rao en remplaçant les paramètres du modèle par les paramètres estimés. Les bornes ainsi obtenues constituent une bonne estimation de la précision des méthodes d'estimation lorsque celles-ci sont asymptotiquement efficaces.

## A.2. Détermination de la densité de probabilité

Pour déterminer de façon graphique la forme de la densité probabilité du résiduel  $\mathbf{r}$ , on peut procéder de la façon suivante. Si l'on appelle  $r_n$  le centre de l'intervalle  $n$ ,  $\delta$  la largeur des intervalles,  $M_n$  le nombre de fois où l'amplitude appartient à l'intervalle  $n$  et  $M$  le nombre de points contenu dans le vecteur résiduel, alors la valeur:

$$\frac{M_n}{M\delta} \quad (\text{A-34})$$

est une approximation de la valeur de la densité de probabilité au point  $r_n$ .

Cela peut se justifier de la façon suivante. La probabilité que le résiduel appartient à l'intervalle  $n$  de largeur  $\delta$  et centré sur  $r_n$  est égale à:

$$P_n(\delta) = \int_{r_n - \frac{\delta}{2}}^{r_n + \frac{\delta}{2}} p(r) dr = \int_0^{r_n + \frac{\delta}{2}} p(r) dr - \int_0^{r_n - \frac{\delta}{2}} p(r) dr \quad (\text{A-35})$$

On effectue un développement limité du premier ordre de chacune de ces intégrales fonctions de leur borne supérieure:

$$P_n(\delta) = \int_0^{r_n} p(r) dr + \frac{\delta}{2} p(r_n) - \int_0^{r_n} p(r) dr + \frac{\delta}{2} p(r_n) = \delta p(r_n) \quad (\text{A-36})$$

Ainsi la densité de probabilité s'obtient en divisant la probabilité par la largeur  $\delta$  de l'intervalle considérée. Comme  $M_n/M$  est une bonne approximation de la probabilité que l'amplitude du résiduel appartienne à l'intervalle  $n$ , alors (A-34) représente bien la densité de probabilité au point  $r_n$ .

### A.3. La méthode "Line-search"

A chaque itération de la méthode "line-search", une direction  $\mathbf{p}_k$  est calculée, ensuite l'amplitude optimale  $\alpha_k$  de cette direction est déterminée. L'itération est donnée par:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \quad (\text{A-37})$$

Un compromis doit être trouvé lors du calcul de  $\alpha_k$ . D'une part, une réduction substantielle de la fonction de coût  $r$  doit être obtenue. D'autre part, le temps nécessaire à cette opération doit rester raisonnable. Le choix idéal est la valeur qui minimise la fonction à une variable suivante:

$$\phi(\alpha_k) = r(\mathbf{x}_k + \alpha_k \mathbf{p}_k) \quad (\text{A-38})$$

Cependant, le temps nécessaire à la détermination de ce minimum est relativement important.

Une méthode plus simple consiste à exiger que le taux de décroissance de la fonction  $r$  soit au moins égal à une fraction  $\rho$  du taux de décroissance initial  $\phi'(0) = \nabla r \mathbf{p}_k$  :

$$\phi(\alpha_k) \leq \phi(0) + \rho \alpha_k \phi'(0) \quad (\text{A-39})$$

Notons que  $0 < \alpha_k < 1$  et prendre  $\rho = 10^{-4}$  est un bon choix. Le premier choix est le pas de Newton, à savoir  $\alpha_k = 1$ . Si la condition (A-39) n'est pas remplie,  $\phi(\alpha)$  est modélisée alors par la fonction quadratique :

$$\phi_q(\alpha) = (\phi(1) - \phi(0) - \phi'(0))\alpha^2 + \phi'(0)\alpha + \phi(0) \quad (\text{A-40})$$

Cette fonction est construite de façon à satisfaire les conditions d'interpolation suivante:  $\phi_q(0) = \phi(0)$ ,  $\phi'_q(0) = \phi'(0)$ ,  $\phi_q(1) = \phi(1)$ . La nouvelle valeur de  $\alpha_k$  est celle qui minimise  $\phi_q(\alpha)$  :

$$\alpha_1 = \frac{\phi'(0)}{2[\phi(1) - \phi(0) - \phi'(0)]} \quad (\text{A-41})$$

La recherche s'arrête à ce stade si cette valeur permet de satisfaire la condition (A-39). Sinon, on construit une fonction cubique à partir de l'interpolation des valeurs  $\phi(0)$ ,  $\phi'(0)$ ,  $\phi(1)$ ,  $\phi(\alpha_1)$  :

$$\phi_c(\alpha) = a\alpha^3 + b\alpha^2 + \phi'(0)\alpha + \phi(0) \quad (\text{A-42})$$

avec

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{\alpha_1^2(\alpha_1 - 1)} \begin{bmatrix} 1 & -\alpha_1^2 \\ -1 & \alpha_1^3 \end{bmatrix} \begin{bmatrix} \phi(\alpha_1) - \phi(0) - \alpha_1\phi'(0) \\ \phi(1) - \phi(0) - \phi'(0) \end{bmatrix} \quad (\text{A-43})$$

La dérivation de  $\phi_c$  montre que le minimum est donné par l'expression suivante :

$$\alpha_2 = \frac{-b + \sqrt{b^2 - 3a\phi'(0)}}{3a} \quad (\text{A-44})$$

Notons que  $\alpha_2$  appartient à l'intervalle  $[0, \alpha_1]$ .

**REFERENCES BIBLIOGRAPHIQUES**



## REFERENCES BIBLIOGRAPHIQUES

- [Andria et al., 1989]** Gregorio Andria, et al. (1989) "Windows and Interpolation algorithms to improve electrical measurement accuracy", *IEEE Transactions on Instrumentation and Measurement*, 38(4):856—863.
- [Benkais, 1993]** M. Benkais (1993), "*Méthodologie de caractérisation des circuits de conversion de données: application aux convertisseurs analogiques-numériques à facteur de mérite élevé. Mise en oeuvre dans le système CANTEST*", Thèse de Doctorat, Université de Bordeaux I.
- [Benkais et al., 1995]** M. Benkais, S. L. Masson, P. Marchegay (1995), "A/D converters characterization by spectral analysis in dual-tone mode", *IEEE Transactions on Instrumentation and Measurement*, IM-44:940--944.
- [Bertocco et al., 1999]** M. Bertocco et al. (1999), "Robust and accurate Real-time Estimation of sensors signal parameters by a DSP approach", *Proceedings IMTC*, 2: 697--702.
- [Blair, 1994]** J. Blair (1994), "Histogram measurement of ADC nonlinearities using sine wave", *IEEE Transactions on Instrumentation and Measurement*, 43:373--383.



- 
- [Brannon]** Brad Brannon, "Aperture Uncertainty and ADC System Performance", <http://www.analog.com>, AN-501.
- [Cauvet et al., 2001]** Philippe Cauvet, Loïc Hamonou (2001), "Improving the dynamic measurements of ADC's using the 2-ADC method", *Teradyne User Group*.
- [Charoy, 2000]** Alain Charoy (2000), "*Compatibilité Electromagnétique*", DUNOD, Paris.
- [Cherubal et al.]** Sasikumar Cherubal, Abhijit Chatterjee, "FASTest: A Tool for Fast Linearity Testing of ADCs", <http://www.ardext.com>.
- [Dallet, 1995]** D.Dallet (1995), "*Contribution à la caractérisation des Convertisseurs Analogiques-Numériques : Evaluation des méthodes et Mises en Œuvre de Nouveaux procédés*", Thèse de Doctorat, université de Bordeaux I.
- [Dennis et al., 1996]** Dennis, R. Shnabel (1996), "*Numerical Methods for Unconstrained Optimization and Nonlinear Equations*", SIAM, Philadelphia.
- [Dufflos et al., 2000]** E. Dufflos, Ph. Vanheeghe (2000), "*Estimation Prédiction*", TECHNIP.
- [Durnerin, 1999]** M. Durnerin (1999), "*Une stratégie pour l'interprétation en analyse spectrale. Détection et caractérisation des composantes d'un spectre*", Thèse de Doctorat, Institut National Polytechnique de Grenoble.

- [Dynad, 2001]** European Project DYNAD, "Methods and draft standards for the DYNamic characterisation and testing of Analogue-to-Digital converters", published in the Internet at <http://www.fe.upt.pt/~hsm/dynad>.
- [Golub et al. , 1973]** G. H. Golub, V. Pereyra (1973), "The differentiation of pseudo-inverse and nonlinear least squares problems whose variables separate", *SIAM. J. Numer. Anal.*, 10: 413--432.
- [Grandke, 1983]** T. Grandke (1983), "Interpolation algorithms for discrete fourier transform weighted signals", *IEEE Transactions on Instrumentation and Measurement*, IM-32(2):350--355.
- [Haddadi et al., 1999]** D. Haddadi, D. Dallet, P. Marchegay (Sept. 1999), "A globally convergent Newton method for characterizing A/D converters," In *Proceedings of the 4<sup>th</sup> International Workshop on ADC Modelling and Testing, IWADC'99*, Bordeaux, France, pp. 72-77.
- [Haddadi et al., May 2000]** D. Haddadi, D. Dallet, P. Marchegay (May 2000), "The initial frequency uncertainty tolerated by the Levenberg-Marquardt method applied to dynamic testing of A/D converters," In *Proceedings of the 17<sup>th</sup> IEEE Instrumentation and Measurement Technology Conference, IMTC2000*, Baltimore, USA, pp.103-107.
- [Haddadi et al., July 2000]** D. Haddadi, D. Dallet and P. Marchegay (July 2000), "Hybrid Newton/Gauss-Newton algorithm for time-domain analysis of A/D converters," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 47(7): 655-659.

- [Haddadi et al., Sept. 2000]** D. Haddadi, D. Dallet and P. Marchegay (Sept. 2000), "Time-domain testing of A/D Converters," In *Proceedings of the 5<sup>th</sup> International Workshop on ADC Modelling and Testing, IWADC'2000*, Vienna, pp. 141-146.
- [Haddadi et al., Nov. 2000]** D. Haddadi, D. Dallet and P. Marchegay (Nov. 2000), "Application of the variable projection method to dynamic testing of A/D converters in dual-tone mode," *Proceedings of the 15<sup>th</sup> Design of Circuits and Integrated Systems Conference*, Montpellier.
- [Harris, 1978]** Harris, F. J. (1978), "On the use of windows for harmonic analysis with the discrete fourier transform", *Proceedings of the IEEE*, 66(1):51--83.
- [Hejn et al., 1992]** K. W. Hejn, R. C. S. Morling (1992), "A Semifixed Frequency Method for Evaluating the Effective Resolution of A/D Converters", *IEEE Transactions on Instrumentation and Measurement*, 41(2): 212--217.
- [Ieeeestd1057, 1994]** IEEE Std 1057 (1994), "IEEE Standard for Digitizing Waveform Recorders".
- [Ieeeestd1241, 2001]** IEEE Std 1241 (2001), "IEEE Standard for terminology and test methods for Analog-to-Digital Converters".
- [James et al., 1994]** B. James, B. D. O. Anderson, R. C. Williamson (1994), "Conditional Mean and Maximum Likelihood Approches to Multiharmonic Frequency Estimation", *IEEE Trans. on Signal Processing*, 42(6).

- [Kay, 1988]** Stephen M. Kay (1988), "*MODERN SPECTRAL ESTIMATION: Theory and Application*", PRENTICE .HALL, New Jersey.
- [Leroux, 1989]** B.C. Le Roux, T.R. McComb, J. Kuffel (Février 1989), "A Comparative Evaluation of Some Practical Algorithms Used in the Effective Bits Test of Waveform Recorders", *IEEE Trans. Instrum. Meas.*, 38(1): 37--42.
- [Levenberg, 1944]** K. Levenberg (1944), "A Method for the Solution of Certain Problems in Least Squares", *Quart. Appl. Math.*, 2: 164--168.
- [MC100EPT21]** Datasheet, "MC100EPT21: LVPECL/LVTTL Translator", [www.onsemi.com](http://www.onsemi.com).
- [Malewski et al., 1987]** R. Malewski, J. Kuffel, T.R. McComb (Juin 1987), "Comparative Evaluation of Computer Methods for Calculating the Best-Fit Sinusoid to the Digital Record of a High-Purity", *IEEE Trans. Instrum. Meas.*, 36(2): 418--422.
- [Marquardt, 1963]** D. Marquardt (1963), "An Algorithm for Least Squares Estimation of Nonlinear Parameters", *SIAM J. Appl. Math.*, 11: 431--441.
- [Morandi et al., 1994]** Morandi, C. and Niccolai, L. (1994), "An improved code density test for the dynamic characterization of flash A/D converters", *IEEE Transactions on Instrumentation and Measurement*, 43:384--388.

- [Morandi et al., 2000]** C. Morandi, G. Chiorboli, D. Dallet, D. Haddadi, S. Mazzoleni, J. Machado da Silva, H. Pernull, P.Y. Roy (Juin 2000), "DYNAD: a framework IV SMT project addressed to the development of dynamic test techniques for analog-to-digital converters", *Journal of Computer Standards & Interfaces*, Elsevier Science Pub, 22: 113--119.
- [Nehorai et al., 1986]** Arye Nehorai, Boaz Porat (1986), "Adaptive Comb Filtering for Harmonic Signal Enhancement", *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-34 (5): 1124--1138.
- [Nocedal et al., 1999]** Jorge Nocedal, Stephen J. Wright (1999), "*Numerical Optimization*", Springer.
- [Nuttall, 1981]** Nuttall, A. H. (1981), "Some windows with very good side lobe behavior", *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-29(1):84--91.
- [Papoulis, 1991]** Athanasios Papoulis (1991), "*Probability, Random Variables and Stochastic Processes*", McGraw-Hill.
- [Pintelon et al., 1996]** R. Pintelon, J. Schoukens (Avril 1992), "An improved sine-wave fitting procedure for characterizing data acquisition channels", *IEEE Transactions on Instrumentation and Measurement*, 45(2): 588--593.
- [Press et al., 1992]** Press, W., Teukolsky, S., Vetterling, W., and Flannery, B. (1992), "*Numerical recipes in C: the art of scientific computing*", Cambridge University Press.

- [Rebai, 2002]** Chiheb Rebai (2002), "*Contribution à la caractérisation des convertisseurs analogiques-numériques haute performances: Mise en œuvre de nouveaux systèmes de traitement du signal pour le test in-situ*", Thèse de Doctorat, université de Bordeaux I.
- [Reeder, 2002]** Rob Reeder (2002), "Compare aperture delay between ADCs", *Test & Measurement World*.
- [Rife et al., 1976]** D. C. Rife, R. R. Boorstyn (1976), "Multiple Tone Parameter estimation from Discrete-Time Observations", *The Bell System Technical Journal*, 55(9).
- [Setty et al., 1990]** P. A. Setty, P. M. VanPeteghem, R. L. Geiger (1999), "An Algorithm to Compensate for Source Spectral Impurity in Dynamic Range Measurements", *IEEE*.
- [TSA1204, 2001]** Datasheet (2001) "TSA1204: dual-channel, 12-bit, 20MSPS A/D converter", [www.st.com](http://www.st.com).







## **Résumé**

L'évaluation des performances dynamiques des convertisseurs analogiques-numériques (CANs) nécessite des algorithmes précis et un matériel de test performant. La première partie de cette étude traite de l'application de l'analyse temporelle au test dynamique des CANs en modes single-tone et dual-tone. La deuxième partie aborde, d'une part, l'implantation industrielle des analyses temporelle et spectrale et, d'autre part, les problèmes liés aux limitations de la carte et de l'instrumentation de test. En particulier, une méthode permettant de compenser le bruit de phase inhérent aux générateurs de test a été proposée.

## **Abstract**

Accurate algorithms and high performance hardware are necessary in order to evaluate the dynamic performances of analogue-to-digital converters (ADCs). The first part of this thesis work deals with dynamic testing of ADCs with time-domain analysis in single-tone and dual-tone modes. The second part discusses an industrial implementation of the time-domain and spectral analyses and addresses the limitations of the test hardware. In particular, a signal processing technique is proposed in order to compensate the phase noise of the test generators.

## **Mots Clés**

Convertisseurs analogiques-numériques, analyse temporelle, analyse spectrale, estimateur des moindres carrés non linéaires, bornes de Cramer-Rao, test industriel, carte de test, bruit de phase.