HAL open science

Synthèse et implémentation de filtres à bande passante variable à faible complexité à l'aide des bancs de filtres

Arnaud Le Pladec, Gilles Burel, Christian Jacquemont

To cite this version:

Arnaud Le Pladec, Gilles Burel, Christian Jacquemont. Synthèse et implémentation de filtres à bande passante variable à faible complexité à l'aide des bancs de filtres. Traitement du Signal, 2002, 19 (3), pp.217-233. hal-03222763

HAL Id: hal-03222763

https://hal.univ-brest.fr/hal-03222763v1

Submitted on 27 Feb 2023

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.

Distributed under a Creative Commons Attribution 4.0 International License

Synthèse et implémentation de filtres à bande passante variable à faible complexité à l'aide des bancs de filtres.
Low complexity implementation of variable band filters using filters banks.

Arnaud Le Pladec*, Gilles Burel**, Christian Jacquemont*

(*) CELAR, Route de Laillé, BP 7419, 35174 BRUZ Cedex
Lepladec@celar.fr
(**) LEST - UMR CNRS 6165, 6 av. Le Gorgeu, BP 809, 29285 BREST Cedex
Gilles.Burel@univ-brest.fr

8 juillet 2002

Résumé

Nous présentons dans cet article une méthode originale de synthèse de filtres à bande passante variable à l'aide des bancs de filtres qui permettent une faible complexité d'implémentation et un changement rapide de la bande passante. Cette méthode, que nous appellerons Rcos, s'appuie sur une analyse approfondie de la méthode TFD (Transformée de Fourier Discrète) et de la méthode OLS (Overlap Save). En introduisant une pondération lors de la sélection des bandes de fréquence élémentaires composant le filtre souhaité, il est possible de réduire l'effet de repliement dû au comportement variant dans le temps du filtre, et d'obtenir ainsi une méthode simple et rapide de synthèse de filtre à bandes passantes variables qui possède de plus une complexité inférieure à l'OLS. La contribution de cet article consiste donc en l'analyse, le développement et la comparaison des différentes méthodes.

Abstract

The paper introduces an innovative method for variable passband filters synthesis using filters banks. This technique allows a fast filter passband modification with a low complexity implementation. This method, called Rcos, has been designed by deeply analysing the DFT (Discrete Fourier Transform) method, and the OLS (Overlap Save) method. By introducing an appropriate weighting during the selection of the elementary frequency bands which belong to the desired filter, it is possible to reduce the aliasing effect due to the time varying behavior of the filter, and to obtain a simple and efficient method which has a lower complexity than that of OLS. The paper's contribution is the analysis, the development and the comparison between the different methods.

Mots clés : Filtres à bande passante variable, Traitement multicadence, Bancs de filtres, Filtres linéaires variant dans le temps.

Keywords : Variable passband filters, Multirate processing, Filter Banks, Linear time-varying filter.

Présentation des auteurs

Arnaud Le Pladec a obtenu le diplôme d'ingénieur de l'Ecole Nationale Supérieure des Ingénieurs des Etudes et Techniques d'Armement (ENSIETA) et le Diplôme d'Etudes Approfondies (DEA) d'électronique et optronique à l'université de Brest en 1999. Il débute ensuite sa carrière à la Délégation Générale de l'Armement (DGA) au Centre Electronique de L'Armement (CELAR) à Bruz, où il assure maintenant la responsabilité de l'activité du Traitement du Signal des Télécommunications.

Gilles Burel a obtenu le diplôme d'ingénieur de l'Ecole Supérieure d'Electricité (Supélec) en 1988, le doctorat de l'Université de Bretagne Occidentale en décembre 1991 et l'Habilitation à Diriger des Recherches en avril 1996. Il est auteur de 19 brevets et de 70 articles. Après avoir débuté sa carrière à Thomson CSF en 1988, dans le domaine du traitement d'images, il a ensuite rejoint Thomson Multimédia. Depuis septembre 1997, il est Professeur à l'Université de Bretagne Occidentale, où il assure la responsabilité de l'une des trois équipes de recherche du Laboratoire d'Electronique et Systèmes de Télécommunications (UMR CNRS 6165), l'équipe "Traitement du Signal pour les Télécommunications". Ses nouvelles activités de recherche se situent essentiellement dans le domaine des communications numériques (interception et analyse, systèmes MIMO, transmissions furtives).

Christian Jacquemont est diplômé de l'Ecole Polytechnique (1994) de l'Ecole Nationale Supérieure des Techniques Avancées (1996). Il a également obtenu un Master of Science en aéronautique au Massachussetts Institute of Technology (M.I.T.) en 1997. Il a ensuite commencé sa carrière à la Délégation Générale de l'Armement (DGA) au Centre Electronique de L'Armement (CELAR) à Bruz, sur le programme Syracuse 3, futur système de communications par satellite pour les armées françaises.

1 Introduction

Avec le développement des systèmes de télécommunications, que ce soit dans le domaine terrestre ou satellite, on voit croître le besoin en méthodes rapides de filtrage. Le filtrage permet par exemple d'isoler des sous canaux fréquentiels et de les commuter. A la contrainte de faible complexité peuvent alors se rajouter des contraintes supplémentaires, telle que la souplesse.

L'objectif de cet article est de présenter une méthode de synthèse de filtres à bande passante variable à moindre complexité de calcul et simple d'utilisation afin de pouvoir changer facilement la bande passante du filtre. Pour un gabarit de filtre défini, nous allons analyser plusieurs méthodes de synthèse possibles en soulignant les contraintes de complexité de calcul (exprimées en Opérations Réelles Par Echantillon Complexe : orpec), mais aussi la facilité à changer la (les) bande(s) passante(s) de ce filtre. En effet, une application typique de ce type de système concerne la commutation fréquentielle de sous canaux : il est alors nécessaire de pouvoir modifier simplement et rapidement les bandes passantes du filtre.

Les bancs de filtres sont adaptés à ce genre d'opérations car le parallélisme engendré par la décomposition en sous-bandes élémentaires et la décimation de la fréquence d'échantillonnage entraînent une très faible complexité de calcul. Toujours dans un objectif de faible complexité et de simplicité, nous analyserons en particulier les méthodes basées sur l'utilisation de la Transformée de Fourier Discrète (TFD) car elles permettent de tirer profit d'algorithmes rapides de type FFT.

L'article est organisé comme suit. Dans la section 2, nous présentons quelques outils mathématiques pour l'analyse des bancs de filtres. Ces outils nous servent notamment à introduire la notion de distorsion de repliement et à fournir un moyen de calculer cette distorsion. Dans la section 3, nous présentons le principe des méthodes à base de TFD. Nous examinerons ensuite trois méthodes (section 4). Ces méthodes se comportent, à une exception près, comme des filtres linéaires variant périodiquement dans le temps. Les variations de la réponse impulsionnelle introduisent de la distorsion de repliement (que nous exprimons dans la section 2). Ces méthodes sont :

- L'OLS (Overlap Save) est une méthode bien connue ([OS89]) qui permet de faire le calcul exact et rapide d'une convolution linéaire à l'aide d'une FFT ;
- La méthode TFD/TFD^{-1} permet d'effectuer la synthèse du filtre directement dans le domaine fréquentiel, mais a le défaut d'introduire une importante distorsion de repliement ;
- La méthode Rcos, que nous proposons, est une amélioration de la méthode précédente qui permet de réduire la distorsion de repliement.

Nous proposons ensuite (section 5) une méthode qui permet, pour chaque approche, de calculer la taille optimale de la FFT. Enfin, nous présentons différents résultats expérimentaux dans la section 6. La contribution originale de l'article consiste en :

- l'analyse des défauts des méthodes OLS et TFD/TFD^{-1}, ceci conduisant au développement d'une méthode originale : Rcos**.**
- La détermination de la taille optimale de la FFT.
- la comparaison des trois méthodes.

La figure 13 donne un exemple de cahier des charges. Une bande élémentaire est définie comme une bande de fréquence de largeur minimale (cette largeur minimale est donnée par les spécifications, et est appelée "résolution"). Dans un contexte de communications, une communication donnée circule dans la bande de fréquence qui lui est allouée, et on utilisera également le terme "canal" pour désigner une telle bande de fréquence. Par rapport à un canal donné, on distingue les deux canaux voisins (canaux adjacents) et les autres canaux. Les spécifications seront notées comme suit (pour les largeurs de bande, l'unité est la fréquence d'échantillonnage) :

B_u : largeur de la bande utile (minimale) ;

B_g : largeur de la bande de garde ;
O_{dB} : ondulation crête à crête dans la bande (en dB) ;
$\widetilde{R}_{dB}$: réjection des canaux adjacents (en dB) ;
R_{dB} : réjection des autres canaux (en dB).

2 Outils mathématiques

2.1 Les bancs de filtres

Un banc de filtres est par définition une série de filtres ayant soit une entrée, soit une sortie en commun. On parle dans le premier cas (resp. second) de banc de filtres d'analyse (resp. de synthèse). Les méthodes précitées mettent en relation un banc de filtres d'analyse et un banc de filtres de synthèse au travers d'une matrice de commutation G (cf. fig 1). La définition de G dépend de la méthode employée. Nous considérerons ici des matrices G diagonales et nous noterons g_k les éléments de la diagonale.

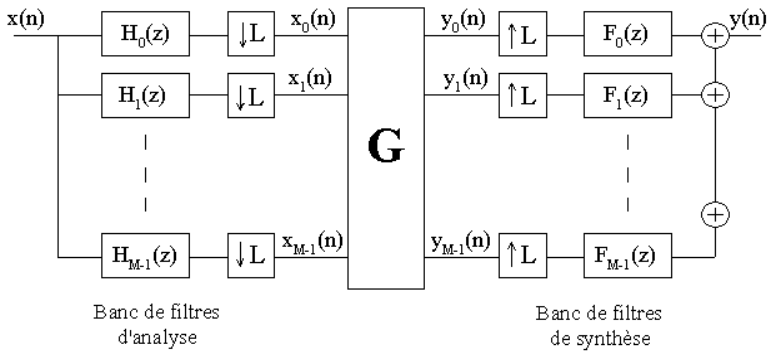

FIG. 1: Bancs de filtres analyse/synthèse

On trouve dans la littérature de nombreux ouvrages et articles traitant des bancs de filtres. Cependant, en général, la matrice G est inexistante et le facteur de décimation L est égal au nombre de branches M. Nous verrons plus loin que dans le cadre de notre application, nous devrons prendre $L < M$ pour limiter la distorsion de repliement.

Reprenons donc, en les adaptant au cas où G est présente et où $L \neq M$, les calculs de ([VAI93], p.224). Rappelons tout d'abord que si $q(n)$ désigne la sortie d'un décimateur d'ordre L en entrée duquel on présente un signal $p(n)$, alors on a :

$$Q(z) = \frac{1}{L}\sum_{k=0}^{L-1} P(z^{1/L} w_L^k) \tag{1}$$

avec $w_L = e^{-j2\pi/L}$. De la même manière, si $q(n)$ désigne la sortie d'un élévateur d'ordre L en entrée duquel on présente un signal $p(n)$, alors on a :

$$Q(z) = P(z^L) \tag{2}$$

D'après la figure 1, on peut écrire :

$$y_i(n) = g_i . x_i(n)$$

Après décimation par un facteur L, on obtient :

$$X_k(z) = \frac{1}{L}\sum_{l=0}^{L-1} H_k(z^{1/L}.w_L^l).X(z^{1/L}.w_L^l)$$

En sommant les sorties des filtres de synthèse on a :

$$Y(z) = \sum_{k=0}^{M-1} F_k(z).g_k.X_k(z^L) \tag{3}$$

Ce qui peut s'écrire :

$$Y(z) = A_0(z).X(z) + \sum_{l=1}^{L-1} A_l(z).X(z.w_L^l) \tag{4}$$

avec :

$$A_l(z) = \frac{1}{L}\sum_{k=0}^{M-1} g_k.H_k(z.w_L^l).F_k(z) \quad \text{pour } l \in \{0, 1, .., L-1\} \tag{5}$$

$A_0(z)$ est la fonction de transfert résultante du banc de filtres et les $A_l(z)$ (pour $l \geq 1$) sont appelées les composantes de repliement. **Si l'on souhaite que le banc de filtres fonctionne comme un filtre classique LIT (Linéaire Invariant dans le Temps), défini par sa fonction de transfert $gab(z)$, on cherchera à ce que :**

- $A_0(z)$ **soit le plus proche possible de** $gab(z)$
- les $A_l(z)$ **pour** $l \in \{1, .., L-1\}$ **soient les plus proches possibles de** 0.

2.2 Représentation Polyphase

Nous allons voir dans cette partie comment, à partir d'une structure telle que celle de la figure 1, nous pouvons obtenir une structure à faible complexité de calcul.

2.2.1 Décomposition de type I et de type II

L'idée est de décomposer un filtre $H(z) \triangleq \sum h(n).z^{-n}$ de la manière suivante ([VAI93] p. 120) :

$$H(z) = \sum_{l=0}^{L-1} z^{-l}.E_l(z^L) \qquad \text{(Type 1 Polyphase)} \tag{6}$$

$$\text{Avec } E_l(z) = \sum_{n=-\infty}^{+\infty} h(nL+l).z^{-n}\ ,\ 0 \leq l \leq L-1$$

Cette décomposition est appelée polyphase de type 1 d'ordre L. On peut aussi écrire $H(z)$ sous la forme :

$$H(z) = \sum_{l=0}^{L-1} z^{-(L-1-l)}.R_l(z^L) \qquad \text{(Type 2 Polyphase)} \tag{7}$$

$$\text{Où } R_l(z) = E_{L-1-l}(z)$$

Cette décomposition est appelée polyphase de type 2 d'ordre L.

2.2.2 Représentations matricielles

La notion de représentation polyphase [VAI93] permet d'élaborer des équivalences qui nous seront utiles par la suite. Ces équivalences sont représentées sur la figure 2.

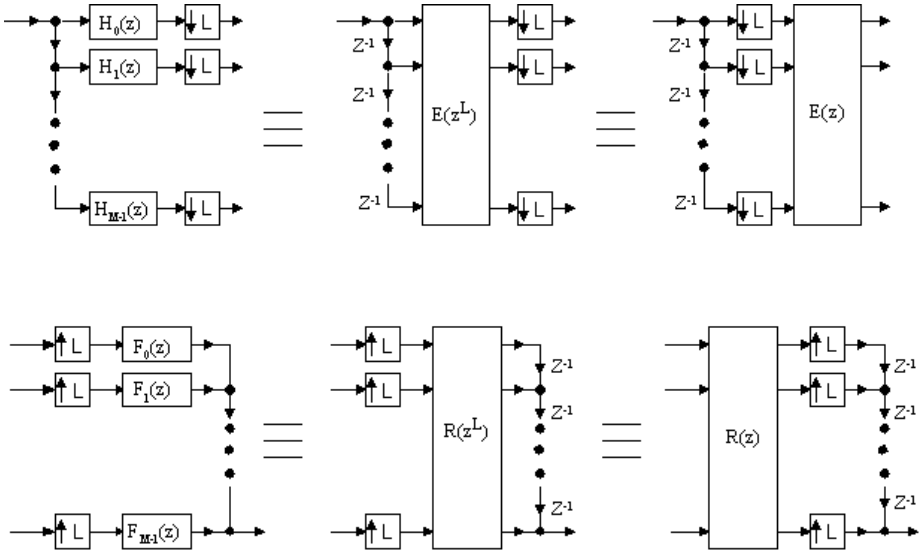

FIG. 2: Représentations polyphases

Sur cette figure, on trouve, en haut, trois schémas équivalents pour la partie "analyse" et, en bas, trois schémas équivalents pour la partie "synthèse". Les liens entre les filtres élémentaires et les éléments des matrices **E** et **R** sont donnés par les équations matricielles suivantes :

$$\mathbf{H}(z) = \begin{bmatrix} H_0(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = \begin{bmatrix} E_{0,0}(z^L) & \cdots & E_{0,L-1}(z^L) \\ \vdots & & \vdots \\ E_{M-1,0}(z^L) & \cdots & E_{M-1,L-1}(z^L) \end{bmatrix} . \begin{bmatrix} 1 \\ \vdots \\ z^{-(L-1)} \end{bmatrix} \tag{8}$$

$$\mathbf{H}(z) = \mathbf{E}(z^L). \begin{bmatrix} 1 \\ \vdots \\ z^{-(L-1)} \end{bmatrix}$$

Par définition, $\mathbf{E}(z)$ est la matrice polyphase de type 1 du banc de filtres d'analyse $\mathbf{H}(z)$. Notons que la matrice polyphase $\mathbf{E}(z)$ caractérise totalement le banc de filtres d'analyse $\mathbf{H}(z)$. De la même manière pour les filtres de synthèse, on a :

$$\mathbf{F}(z) = \left[\begin{array}{ccc} F_0(z) & \cdots & F_{M-1}(z) \end{array}\right] = \left[\begin{array}{ccc} z^{-(L-1)} & \cdots & 1 \end{array}\right] . \left[\begin{array}{ccc} R_{0,0}(z^L) & \cdots & R_{0,M-1}(z^L) \\ \vdots & & \vdots \\ R_{L-1,0}(z^L) & \cdots & R_{L-1,M-1}(z^L) \end{array}\right]$$

$$\mathbf{F}(z) = \left[\begin{array}{ccc} z^{-(L-1)} & \cdots & 1 \end{array}\right] . \mathbf{R}(z^L) \tag{9}$$

Par définition, $\mathbf{R}(z)$ est la matrice polyphase de type 2 du banc de filtres de synthèse $\mathbf{F}(z)$. Notons que la matrice polyphase $\mathbf{R}(z)$ caractérise totalement le banc de filtres de synthèse $\mathbf{F}(z)$.

Pour chacun des deux cas de figure, la première équivalence est obtenue en appliquant la représentation polyphase et la seconde équivalence est obtenue en appliquant les identités nobles ([VAI93], p.119). Etant placée derrière les décimateurs (resp. devant les élévateurs), la matrice $\mathbf{E}(z)$ (resp. $\mathbf{R}(z)$) travaille à une cadence L fois moins élevée dans la troisième représentation. La complexité du système d'analyse et de synthèse est donc L fois moindre. Ceci est une des justifications de l'intérêt des bancs de filtres pour des systèmes à moindre complexité de calcul.

2.3 Distorsion de repliement

Nous avons pour l'instant étudié les bancs de filtres en utilisant la transformée en Z. Les gabarits de filtres étant définis en fréquentiel, il nous reste donc à convertir les termes de repliement. Ceci se réalise simplement en posant $z = e^{j\omega}$ dans l'équation 5. On peut alors noter les termes obtenus $A_l(\omega)$.

Pour un ω fixé, on définit : $\mathbf{A}(\omega) = [A_0(\omega), A_1(\omega), ..., A_{L-1}(\omega)]$ et $\mathbf{Gab}(\omega) = [gab(\omega), 0, ..., 0]$ le gabarit général avec $gab(\omega)$ le gabarit du filtre linéaire invariant dans le temps (LIT). Ce terme $gab(\omega)$ représente la fonction idéale que l'on cherche à approcher et les zéros qui suivent indiquent qu'idéalement, l'objectif en ce qui concerne les termes de repliement est zéro. On peut définir un critère d'erreur par le calcul de la norme de Frobenius[1] de la différence :

$$D(\omega) = \|\mathbf{A}(\omega) - \mathbf{Gab}(\omega)\|_f \tag{10}$$

$$D(\omega) = \sqrt{|A_0(\omega) - gab(\omega)|^2 + \sum_{l=1}^{L-1} |A_l(\omega)|^2}$$

Les termes $|A_0(\omega) - gab(\omega)|^2$ et $\sum_{l=1}^{L-1} |A_l(\omega)|^2$ sont de natures différents. En effet, le premier terme traduit l'erreur sur la réponse fréquentielle invariante dans le temps, alors que le second terme représente les distorsions dues à la non-stationarité du signal (elle-même due au traitement par blocs). Ce dernier type de distorsion introduit des fréquences qui n'existent pas dans le signal d'entrée.

On peut faire l'amalgame entre ces deux types de distorsions, et considérer la distorsion globale, comme ci-dessus, ou bien, imposer une borne maximale au second type de distortion (par exemple, -40dB hors bande de garde dans certains résultats expérimentaux que nous présenterons plus loin) et étudier ensuite le terme $|A_0(\omega) - gab(\omega)|^2$.

Si on veut uniquement étudier l'influence des termes non LIT ($A_l(\omega)$ pour $l \neq 0$), on considère que $A_0(\omega) = gab(\omega)$ et on a :

[1] Rappel : $\|\mathbf{A}\|_f = \sqrt{Trace\left(\mathbf{A}^{*T}.\mathbf{A}\right)}$

$$Dr(\omega) = \sqrt{\sum_{l=1}^{L-1} |A_l(\omega)|^2} \tag{11}$$

On retrouve ainsi le terme utilisé dans ([VAI93], p.367) nommé “distorsion de repliement” (“Aliasing Distorsion”), terme sur lequel on peut définir une tolérance. L’étude complète du banc de filtres doit donc prendre en compte les spécifications de filtrage classiques, la complexité, mais aussi cette distorsion de repliement qui peut être considérée comme un bruit supplémentaire dans le système (bruit dépendant du signal d’entrée). Nous verrons dans quelle mesure ce terme peut nous renseigner sur la pertinence de notre filtrage par rapport à un filtrage LIT classique.

En règle générale, le niveau acceptable pour la distorsion de repliement est indirectement déterminé par le fait que le terme $D(\omega)$ de l’équation 10 ne doit pas dépasser un niveau donné dans les spécifications, soit directement, soit indirectement en considérant que la distorsion de repliement (au carré) vient s’ajouter à la distorsion sur le terme LIT. Dans ce dernier cas, la somme de ces deux distorsions doit respecter les spécifications.

3 Principe des méthodes à base de transformée de Fourier rapide

3.1 Schéma général

Les trois méthodes que nous étudions dans cet article sont basées sur l’utilisation de la transformée de Fourier rapide (FFT). La figure 3 illustre le principe général de ces méthodes. Le signal d’entrée est découpé en blocs de longueur M. D’un bloc au suivant on se décale de L échantillons ($L \leq M$). Il peut donc y avoir du recouvrement entre les blocs d’entrée. Pour un bloc donné, on calcule sa Transformée de Fourier Discrète sur M points (ceci peut être réalisé de manière rapide en utilisant la FFT). Mathématiquement, la TFD est représentée par une matrice W de taille $M \times M$. L’élément à la ligne k et la colonne l dans cette matrice est w_M^{kl} avec $w_M = e^{-j2\pi/M}$. Les M valeurs obtenues sont pondérées : ceci est représenté par une multiplication par une matrice diagonale G. Ensuite, une TFD inverse est appliquée (matrice $\frac{1}{M}W^*$). Parmi les M valeurs obtenues, seules L valeurs seront retenues (sélection). On retrouve donc en sortie un signal de même fréquence d’échantillonnage que le signal d’entrée. Mathématiquement, la sélection peut être représentée par une matrice diagonale S, à valeurs binaires.

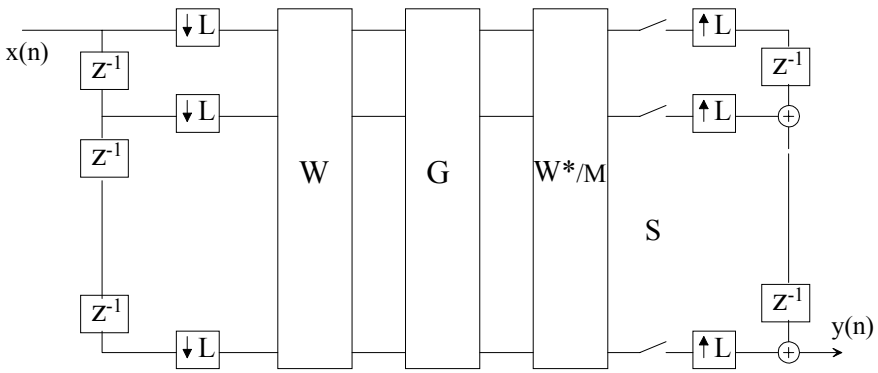

FIG. 3: Principe des méthodes à base de FFT

Tout le problème consiste à choisir au mieux les coefficients de pondération, afin d’atteindre une complexité de calcul aussi faible que possible, sous la contrainte du respect du gabarit de filtrage spécifié.

Nous avons vu, de plus, qu'un autre critère doit être pris en compte : la simplicité des coefficients de pondération. Si ces coefficients sont tous binaires, il s'agit alors de commutateurs et nous avons une simple sélection. Dans le cas contraire, nous avons une pondération qui nécessite des multiplicateurs. Cette opération est plus ou moins aisée selon le nombre et la variation de ces coefficients pour changer de bande passante.

Le schéma de la figure 3 peut être rapproché des schémas de la figure 2, en prenant $E(z) = W$ et $R(z) = \frac{1}{M}SW^*$. Les équivalences de la figure 2 permettent de déplacer les décimateurs et élévateurs : on obtient alors le schéma de la figure 4. Selon les identité nobles, les composantes des matrices polyphases sont : $\left[E(z^L)\right]_{k,l} = w_M^{kl}$ et $\left[R(z^L)\right]_{l,k} = \frac{1}{M}s_l w_M^{-kl}$, où les s_l désignent les éléments diagonaux de la matrice de sélection.

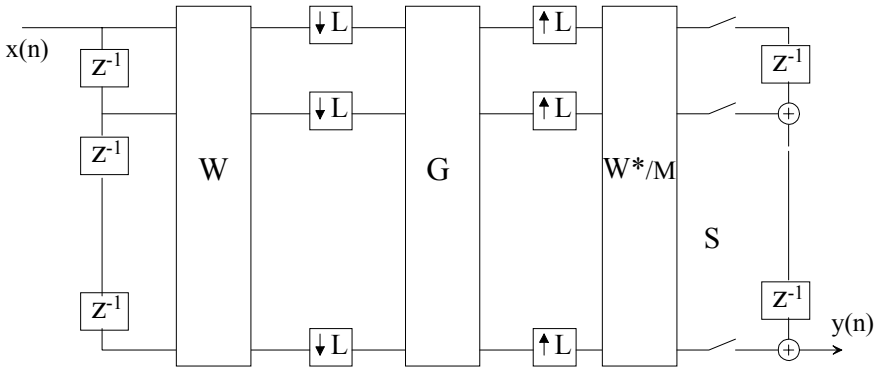

FIG. 4: Représentation polyphase des méthodes à base de FFT

En utilisant les identités de la figure 2 ainsi que les équations 8 et 9, on se ramène finalement au schéma de la figure 1, avec les filtres suivants :

$$H_k(z) = \sum_{l=0}^{M-1} z^{-l} E_{kl}(z^L) = \sum_{l=0}^{M-1} z^{-l} w^{kl} \tag{12}$$

et

$$\begin{aligned} F_k(z) &= \sum_{l=0}^{M-1} z^{-(M-1-l)} R_{lk}(z^L) = \frac{1}{M}\sum_{l=0}^{M-1} z^{-(M-1-l)} s_l w_M^{-kl} \\ &= \frac{1}{M}\sum_{l=0}^{M-1} z^{-l} s_{M-1-l} w_M^{-k(M-1-l)} = w_M^k \sum_{l=0}^{M-1} z^{-l} s_{M-1-l} w_M^{kl} \end{aligned} \tag{13}$$

Ce qui nous donne :

$$\begin{cases} H_0(z) = 1 + z^{-1} + ... + z^{-(M-1)} = \frac{1-z^{-M}}{1-z} \\ H_k(z) = H_0(z.w^{-k}) \end{cases}$$

$$\text{et} \begin{cases} F_0(z) = s_{M-1} + s_{M-2}.z^{-1} + ... + s_0.z^{-(M-1)} \\ F_k(z) = w^k.F_0(z.w^{-k}) \end{cases}$$

$[s_0, ..., s_{M-1}]$ représente la diagonale de la matrice de sélection.

L'intérêt des formules ci-dessus est de nous fournir un moyen de calculer les termes de repliement. Il suffit pour cela d'utiliser les équations 5, 10 et 11, en remplaçant les filtres par les valeurs ci-dessus.

3.2 Complexité de calcul

Dans ce paragraphe, nous allons calculer la complexité de calcul d'une implémentation à partir de bancs de filtres à base de transformée de Fourier rapide, c'est-à-dire le coût de calcul lors du filtrage. Ce critère de comparaison doit être différencié du critère de simplicité d'implémentation et d'utilisation pour des changements rapides de bande(s) passante(s). Cette dernière notion sera abordée lors de la présentation des méthodes.

On peut évaluer la complexité du système (pour une entrée complexe), grâce à l'évaluation de la complexité de la FFT suivante ([MAL92], p.60) :

	Add. réelles / bloc	Mult. réelles / bloc
FFT/FFT^{-1}	$6M(\log_2(M)-1)+8$	$2M(\log_2(M)-3)+8$
Pondération	0	$\beta \quad (0 \leq \beta \leq 2M)$

Cette complexité de la FFT est donnée pour l'algorithme "Split-Radix FFT" qui est la méthode actuelle la moins complexe pour des tailles de FFT en puissance de 2.

Les méthodes étudiées se distinguent par la pondération (matrice G). La complexité de la pondération dépendra du type de méthode utilisée (lorsque les coefficients de pondérations sont binaires, cette complexité est nulle, alors que lorsqu'ils sont quelconques, elle vaut $2M$). Avec des accumulateurs mult-add, on garde la complexité maximale entre une addition et une multiplication pour un type d'opération. On obtient l'expression suivante pour la complexité, exprimée en opérations réelles par échantillon complexe (orpec) :

$$C = \frac{6M.\log_2(M) - bM + 8}{L} \text{ (orpec)} \tag{14}$$

où $b = 6 - \beta$ est compris entre 4 (pondération quelconque) et 6 (pondération binaire). On peut donc remarquer que, pour une longueur de FFT donnée, la formule ci-dessus donne une complexité quasiment indépendante de la méthode (ormis une faible influence du facteur b). Cependant, selon la méthode utilisée, la taille de FFT requise pour atteindre les spécifications pourra fortement varier, et donc entraîner des variations non négligeables de la complexité. Nous reviendrons sur ce point important après avoir présenté les méthodes.

La figure 5 donne quelques valeurs de la complexité pour différentes valeurs de M et différents taux de recouvrement $(M-L)/M$ (pour $b = 6$). Cela permet d'avoir une idée des valeurs de la complexité pour des couples (M, L) typiques.

M	512		1024		2048		4096		8192	
Rec. (%)	L	comp.	L	comp.	L	comp.	L	comp.	L	comp.
0	512	48,02	1024	54,02	2048	60,02	4096	66,02	8192	72,02
5	486	50,6	972	56,9	1946	63,2	3892	69,5	7782	75,8
10	460	53,4	922	60,0	1844	66,7	3686	73,4	7372	80,0
15	436	56,4	870	63,6	1740	70,6	3482	77,7	6964	84,7
20	410	60,0	820	67,5	1638	75,0	3276	82,5	6554	90,0
25	384	64,0	768	72,0	1536	80,0	3072	88,0	6144	96,0
30	358	68,7	716	77,3	1434	85,7	2868	94,3	5734	102,9
35	332	74,1	666	83,1	1332	92,3	2662	101,6	5324	110,8
40	308	79,8	614	90,1	1228	100,1	2458	110,0	4916	120,0

FIG. 5: Quelques valeurs de la complexité (en orpec)

4 Description des trois méthodes

4.1 Principe de la démarche

L'un des problèmes qui nous restent à résoudre est de calculer au mieux les gains g_k afin de minimiser la complexité sous la contrainte du respect des spécifications de filtrage. En vertu des applications visées, on doit aussi prendre en compte la simplicité des coefficients g_k (des valeurs binaires étant idéales). En effet, une application typique de ce type de système concerne la commutation fréquentielle de sous canaux : le fait d'avoir des coefficients g_k non binaires complique beaucoup le système dans ce contexte, notamment lorsque plusieurs sous canaux doivent être sélectionnés simultanément.

La démarche suivie consiste à partir d'une méthode bien connue, l'Overlap Save (OLS), qui a l'avantage de supprimer les termes de repliement mais, on le verra, au prix d'une complexité plus importante, et de valeurs non-binaires des g_k. Nous présentons ensuite une méthode moins connue, mais assez immédiate, que nous nommerons TFD/TFD^{-1}. Dans cette méthode, les coefficients g_k sont binaires, mais le repliement est important. L'analyse des faiblesses de ces deux méthodes nous amènera à proposer une autre approche, que nous nommerons Rcos, dans laquelle les coefficients g_k sont binaires "presque partout", et qui présente l'avantage de limiter le repliement à de faibles valeurs.

4.2 L'Overlap Save (OLS)

4.2.1 Descriptif

La méthode OLS permet de déterminer des coefficients de pondération g_k qui assurent l'absence de distorsion de repliement. Ces coefficients sont en effet calculés en réalisant la TFD de la réponse impulsionnelle d'un filtre d'ordre limité. Cette technique a été développée pour faire le calcul rapide d'une convolution linéaire ([OS89], p.558). Le principe est simplement de faire une multiplication dans le domaine fréquentiel au lieu d'une convolution dans le domaine temporel [LM96].

La convolution obtenue par cette méthode est une **convolution circulaire**. La sortie obtenue n'est donc pas exactement égale à la sortie souhaitée. Il n'y a qu'une partie des termes qui est correcte. Pour y remédier, du recouvrement est introduit à l'entrée. Pour cela on sélectionne un bloc de longueur M tous les L échantillons ($M > L$). On calcule ainsi les termes qui n'étaient pas bons lors du calcul précédent. **Le principe est ensuite de comparer la convolution circulaire à la convolution linéaire**. L'égalité de ces deux convolutions est assurée dès que :

$$\underbrace{M}_{\text{Taille de la FFT}} \geqslant \underbrace{L}_{\text{décimation}} + \underbrace{H}_{\text{Ordre du filtre}} - 1 \tag{15}$$

Si l'on assure la condition (15), on a alors $y_c(n) = y_l(n)$, où y_c (resp. y_l) représente le résultat de la convolution circulaire (resp. linéaire). On sait de façon classique que pour une convolution linéaire $Y_l(z) = H(z).X(z)$, d'où $Y_c(z) = H(z).X(z)$. En identifiant avec l'équation (4), on observe que dans ce cas tous les $A_l(z)$ sont nuls pour $l \neq 0$. On se retrouve dans le cas d'un système LIT : **identifier la convolution circulaire à la convolution linéaire revient à annuler les composantes de repliement**.

Détaillons à présent la condition sur l'ordre du filtre. On sélectionne en entrée un bloc de M échantillons $x(n)$. La longueur du filtre $h(n)$ étant notée H, on sait que la longueur de $y_l(n)$, sortie de la convolution linéaire, est $M + H - 1$ (*Cf fig. 6 a,b,c*). Mais la convolution obtenue par cette méthode est une convolution circulaire. Or faire une M-convolution circulaire équivaut à prendre une période de longueur M de la convolution linéaire périodisée par M. (*Cf fig. 6 d*). En examinant les termes de 0 à $(M - 1)$, on observe alors du repliement temporel sur les $H - 1$ premiers termes de la convolution

circulaire. En revanche, les L derniers termes sont égaux à ceux de la convolution linéaire. L'idée de l'OLS est donc d'introduire du recouvrement à l'entrée, c'est à dire que nous allons prendre un bloc de longueur M tous les L échantillons, où L vérifie la condition d'OLS (15). Il suffit ensuite d'éliminer les $H-1$ premiers termes et de garder les L derniers termes pour reconstituer le signal de sortie.

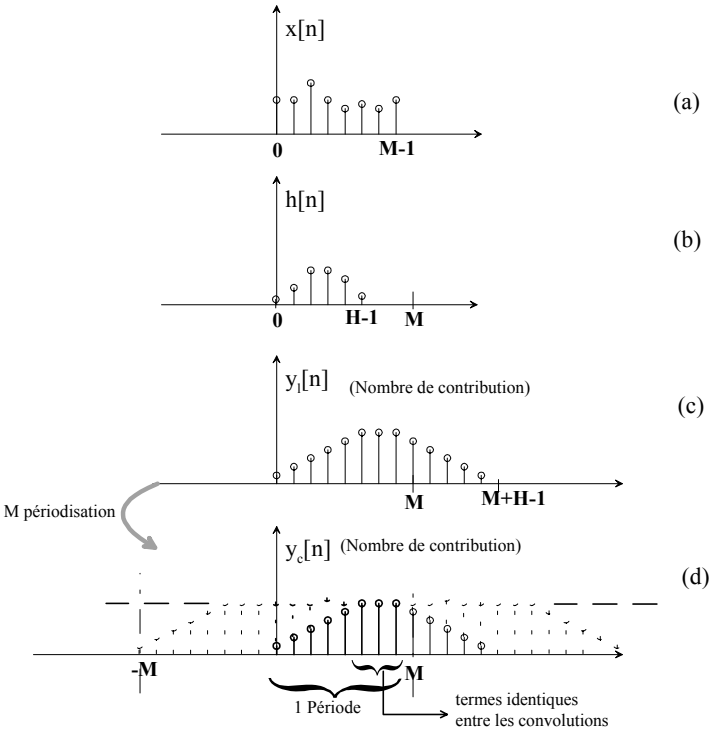

FIG. 6: Analyse temporelle de l'OLS

4.2.2 Méthodologie

Pour mettre en oeuvre cette méthode, nous devons tout d'abord établir la réponse impulsionnelle du filtre à partir du gabarit fréquentiel désiré (ondulation, réjection,...), représenté sur la figure 7. On peut par exemple utiliser la méthode de Parks-MacClellan [PMC73] pour synthétiser un filtre à ondulation constante ("equiripple"). Ces filtres sont optimaux en terme de longueur de filtre RIF (Réponse Impulsionnelle Finie), étant donné un gabarit. Or, dans le cadre de l'OLS, diminuer la longueur du filtre diminue la complexité. On calcule ensuite les M coefficients TFD de ce filtre que l'on introduit dans la pondération en sortie de la FFT (nous verrons, dans la section suivante, comment calculer M). La formule de Bellanger ([VAI93], p.57) nous permet d'évaluer l'ordre H d'un filtre RIF à ondulation constante connaissant O_{dB}, R_{dB} et B_g qui représentent respectivement l'ondulation dans la bande passante, la réjection des autres canaux, et la largeur de la bande de garde.

$$H = \frac{2\log_{10}(\frac{1}{10.\delta_1.\delta_2})}{3.B_g} \tag{16}$$

avec

$$\delta_1 = 10^{O_{dB}/40} - 1 \tag{17}$$

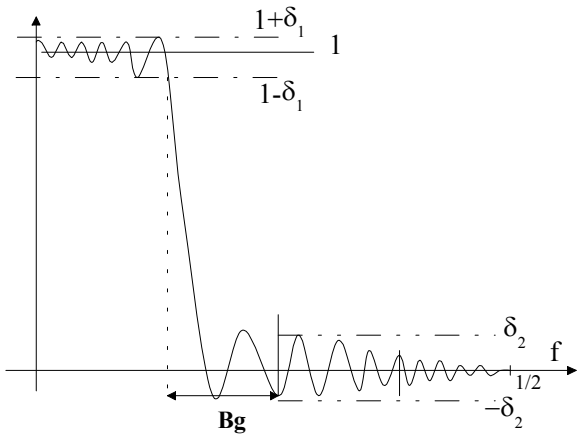

FIG. 7: Paramètres de la formule de Bellanger

et :

$$\delta_2 = 10^{-R_{dB}/20} \tag{18}$$

L'avantage de cette méthode est qu'elle permet de maîtriser les caractéristiques telles que l'ondulation, la réjection et la bande de garde. Par contre, si l'on souhaite changer la bande passante, il faut recalculer tous les coefficients g_k, ce qui est un inconvénient important pour le type d'application que nous visons.

4.3 Méthode TFD/TFD^{-1}

4.3.1 Descriptif

Cette méthode est la plus simple d'utilisation, mais elle contient des composantes de repliement non nulles. Le filtrage ainsi réalisé n'est pas un filtrage LIT. De plus, elle nécessite une taille de FFT grande pour obtenir des résultats satisfaisants. En revanche, la simplicité de cette méthode en fait un système efficace. En effet, la synthèse du filtre se fait directement dans le domaine fréquentiel.

La différence avec le schéma précédent est qu'ici la diagonale de la matrice G ne contient que des 0 et des 1, c'est à dire que l'on bascule des interrupteurs. Ceci correspond à la sélection de bandes de fréquence élémentaires. A la sortie de la FFT^{-1}, nous avons une matrice de sélection qui permet d'éliminer ou de garder les termes en sortie.

Pour cette méthode, nous construisons la réponse fréquentielle du filtre en fonction de la bande de fréquence désirée. Nous ajoutons en plus quelques bandes de fréquence élémentaires choisies dans la bande de garde pour obtenir nos spécifications. Cette méthode est la synthèse dite par "frequency sampling". Comme nous sélectionnons un nombre limité d'échantillons à 1 parmi les M échantillons possibles, la réponse impulsionnelle de ce filtre est de longueur M (ce qui est fini en fréquentiel est infini en temporel).

Si l'on calcule la FFT inverse de la réponse fréquentielle désirée (une porte), on obtient une réponse temporelle qui a l'allure donnée sur la figure 8.

Si l'on reprend l'analyse faite pour l'OLS avec un filtre de longueur M et un mot en entrée de longueur M, on a une convolution linéaire $y_l(n)$ de longueur $2.M - 1$ (fig. 9).

Une interprétation possible de la M-convolution circulaire est de périodiser par M la réponse impulsionnelle du filtre, de faire une convolution linéaire classique et enfin d'extraire une période M de

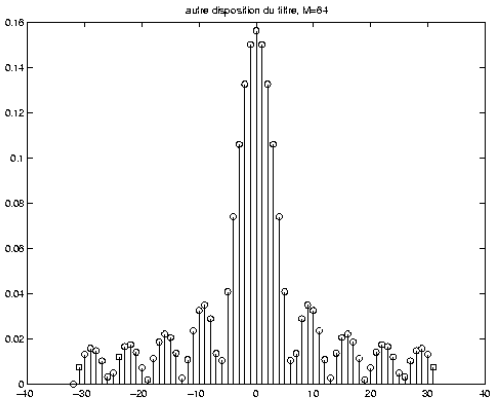

FIG. 8: Réponse impulsionnelle centrée

cette convolution linéaire (*fig. 9*). Après périodisation pour obtenir la convolution circulaire, on observe que tous les termes (sauf un), sont mauvais. Si nous voulions retrouver la convolution linéaire, nous serions amenés à faire une décimation par un, c'est à dire à ne pas faire de décimation. La complexité de calcul serait maximale. Nous pouvons retrouver ce résultat en reprenant la condition d'OLS (15) avec $H = M$. On a alors tout de suite $L = 1$. Nous allons en fait tolérer un écart entre les convolutions, tout en sachant que cet écart introduit des composantes de repliement non nulles. Il est clair que les termes qui contiennent le moins de repliement sont les termes centraux. Nous allons donc choisir la matrice de sélection de manière à conserver les L termes centraux.

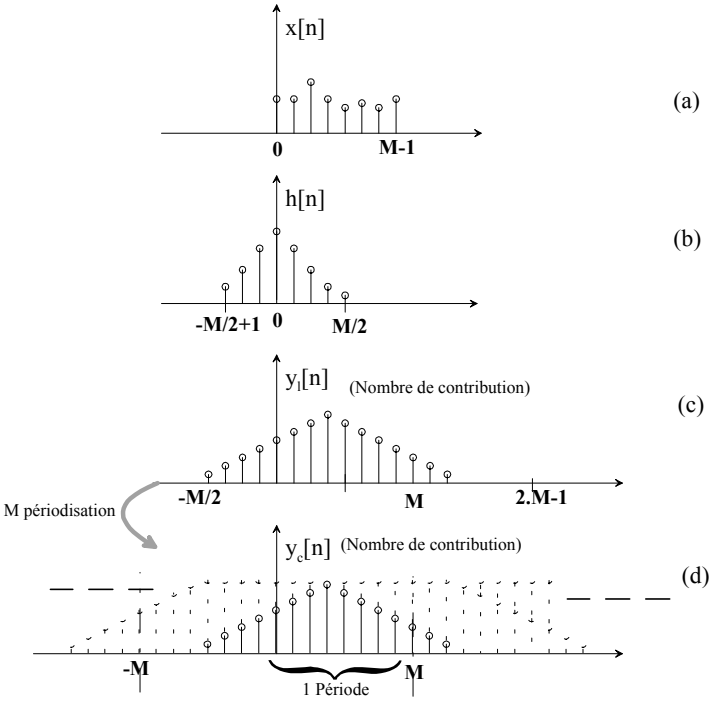

FIG. 9: Analyse temporelle de la méthode TFD/TFD^{-1}

4.3.2 Méthodologie

La procédure suivie est la suivante :

1. Choix d'une valeur de M (puissance de 2)
2. Détermination des coefficients binaires g_k
3. Détermination de la plus grande valeur de L qui permet d'atteindre le gabarit imposé
4. Détermination des valeurs optimales de M et L, en utilisant les résultats de la section 5.

Le choix réalisé dans l'étape (1) est arbitraire, car il sera corrigé dans l'étape (4). On choisira cependant, de préférence, une valeur "raisonnable", par exemple en prenant deux ou quatre fois la valeur de M qui serait trouvée pour l'OLS avec le même gabarit (en effet, on verra plus loin que la méthode TFD/TFD^{-1} nécessite des tailles de FFT plus importantes que l'OLS).

Dans l'étape 2, on sait que l'on doit affecter la valeur 1 à la bande passante et la valeur 0 à la bande de réjection. Reste la bande de garde : les coefficients binaires affectés à la bande de garde ont une forte influence sur le résultat. Plus il y a de 1, meilleure est l'ondulation crête à crête mais plus le niveau dans la bande de réjection est important. Pour avoir un niveau acceptable tant en qui ce concerne la bande de réjection que la bande passante, il faut envisager un nombre égal de 1 et de 0 dans la bande de garde.

Les coefficients g_k étant ainsi fixés, nous devons rechercher la plus grande valeur de L qui permet d'atteindre le gabarit imposé (étape 3). Nous devons tout d'abord obtenir le terme $A_0(z)$ suivant le gabarit désiré, en laissant une marge suffisante, car les termes de repliement vont venir dégrader le résultat. Nous devons vérifier l'ondulation crête à crête et la suffisance de la réjection. Une fois le gabarit obtenu, nous devons essayer d'obtenir une distorsion de repliement Dr (eq. 11) suffisamment faible pour ne pas détériorer le terme LIT au delà de la marge adoptée. En effet, l'équation 10 montre que la distorsion de repliement (au carré) s'ajoute à la distorsion sur le terme LIT. La somme de ces deux distorsions ne doit pas nous faire sortir du gabarit donné par les spécifications.

La valeur de L étant déterminée, et la valeur de M étant celle qui a été choisie dans l'étape (1), la méthode proposée dans la section 5 permet d'en déduire la couple (M, L) optimal. Cette méthode utilisant cependant quelques approximations, un dernier ajustement de la valeur de L peut être utile.

4.3.3 Intérêt supplémentaire de la méthode TFD/TFD^{-1} dans un contexte de commutation fréquentielle

Dans de nombreuses applications de télécommunications, chaque utilisateur se voit allouer (de manière statique ou dynamique) une bande de fréquence, souvent appelée "canal". Au sein d'un relais, qui peut être par exemple un satellite, on a souvent à réaliser de la commutation, c'est à dire rediriger chaque canal entrant sur le bon canal sortant afin de connecter les utilisateurs qui souhaitent communiquer. Par exemple, pour permettre à l'utilisateur qui écoute sur le canal numéro 1 et recevoir les données transmises par l'utilisateur qui transmet sur le canal numéro 3, il faut que le satellite transpose le signal présent dans la bande de fréquence numéro 3 vers la bande de fréquence numéro 1.

Lorsque l'on vise des applications de commutation fréquentielle de sous-canaux, le fait d'avoir des coefficients de pondération simples et pas ou peu étendus au delà de la (des) bande(s) passante(s) du filtre présente un avantage supplémentaire important, illustré par la figure 10. Sur cette figure est représenté le contenu de la matrice G pour les méthodes TFD/TFD^{-1} et OLS, sans commutation (à gauche) et avec commutation (à droite). Nous pouvons voir ainsi la simplicité de la commutation dans le cas de la méthode TFD/TFD^{-1}.

On a ici $M = 16$ et trois canaux fréquentiels (ces valeurs sont faibles par rapport aux valeurs que l'on aurait en pratique, mais ont l'avantage de simplifier l'illustration). Sur les images de droite, on commute les deux premiers canaux (le canal 1 en entrée est dirigé vers le canal 2 en sortie, et inversement). Avec

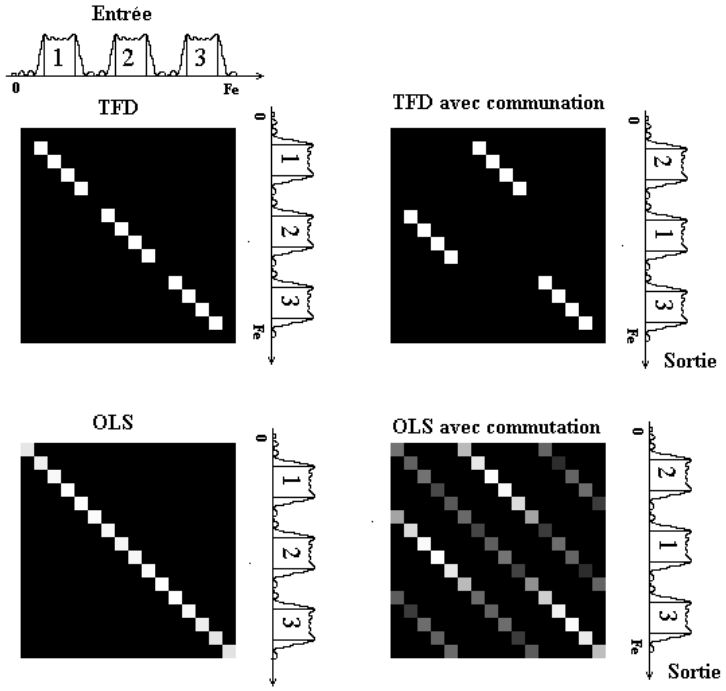

FIG. 10: Matrice G pour les méthodes TFD et OLS (3 canaux fréquentiels). A gauche : sans commutation. A droite : avec commutation

la méthode TFD/TFD^{-1}, les coefficients de la matrice G qui servent à isoler tel ou tel sous-canal sont binaires et confinés à la bande passante du canal considéré. Le nombre total de coefficients non nuls dans la matrice G reste donc le même, qu'il y ait commutation ou pas. Avec la méthode OLS ce n'est pas le cas, car, pour assurer l'absence de repliement, le filtrage de chaque sous canal nécessite M coefficients non nuls. Le nombre total de coefficients non nuls en cas de commutation est donc $3M$ dans notre exemple, ce qui augmente encore la complexité. Dans un cas général, le nombre de coefficients non nuls serait M multiplié par le nombre de sous-canaux, ce qui peut conduire à une augmentation considérable de la complexité.

4.4 Méthode en Cosinus surélevé (Rcos)

L'analyse sous l'aspect temporel (comme pour l'OLS) de la méthode TFD/TFD^{-1} avec du recouvrement, nous a permis de concevoir une amélioration.

4.4.1 Descriptif.

Le principe est de pondérer les points FFT de la bande de garde par des coefficients calculés selon un cosinus surélevé. Cette pondération semble adaptée à notre utilisation : les expérimentations décrites plus loin montrent que d'autres pondérations classiques sont légèrement moins performantes que le cosinus surélevé.

Cette méthode est en quelque sorte un compromis entre la méthode radicale TFD/TFD^{-1}, où l'on utilise uniquement des 0 et des 1, et la méthode de l'OLS, où l'on doit pondérer tous les points FFT pour dessiner la réponse temporelle souhaitée. L'avantage est qu'ici la pondération est unique et ne concerne

que la bande de garde. On peut imaginer un système qui récupère et calcule la bande de garde une fois connue la(les) bande(s) passante(s). Il faut noter qu'il est préférable de ne pas utiliser le dernier point de la bande de garde pour s'assurer une réjection suffisante. Pour cette méthode, nous retrouvons les mêmes variables que pour la méthode TFD/TFD^{-1}.

D'autre part, du fait du confinement des coefficients non nuls à la bande passante et à son voisinage immédiat, la méthode proposée présente en cas de commutation fréquentielle le même avantage supplémentaire que la méthode TFD/TFD^{-1} sur la méthode OLS (cf. section 4.3.3).

4.4.2 Analyse.

Si l'on regarde d'un peu plus près la construction des termes de repliement dans l'analyse de la méthode TFD/TFD^{-1}, on voit que ceux-ci sont construits à partir des termes extrêmes de la réponse impulsionnelle. L'idée est de diminuer leurs valeurs afin de diminuer la contribution des termes de repliement. On sait que ces termes sont importants à cause de la rapide transition de la réponse fréquentielle. L'idée est alors d'utiliser la bande de garde pour diminuer ces termes en donnant une transition moins brutale. Une bande de garde en cosinus surélevé ([PRO95], p. 546) semble adaptée à cette opération. Nous obtenons une réponse fréquentielle du type indiqué sur la figure 11.

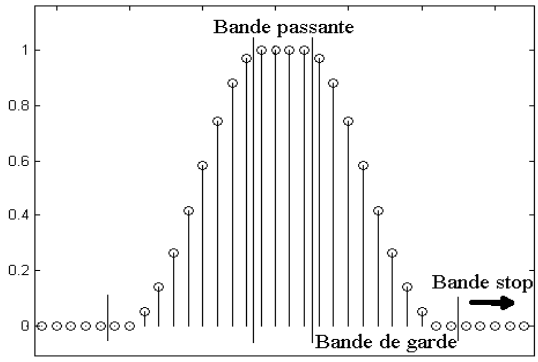

FIG. 11: Pondération typique avec la méthode Rcos

Le critère important pour le choix de la pondération de la bande de garde est la décroissance de la réponse temporelle associée. Il existe d'autres pondérations possibles (Blackman, Hamming,...), mais nous assurons déjà une décroissance en $o\left(1/x^3\right)$, qui est optimale pour ce genre de masque.

4.4.3 Méthodologie

La procédure suivie est la même que pour la méthode TFD/TFD^{-1}. Les seules différences sont les suivantes :

1. Dans l'étape (1), on peut choisir une valeur de M plus faible (on verra, en effet, que cette méthode nécessite des tailles de FFT moins importantes que la méthode TFD/TFD^{-1}). On peut choisir la valeur trouvée pour l'OLS.
2. Dans l'étape (2), on calcule les coefficients de pondération g_k selon la méthode que nous venons d'indiquer.

5 Estimation de la taille de la TFD

5.1 Principe de la méthode

Dans cette section, nous proposons une méthode permettant d'estimer la taille optimale de la TFD en fonction des spécifications, ceci pour les différentes approches présentées (Overlap-Save, TFD et Rcos). Il s'agit d'un problème d'optimisation sous contrainte : le critère à optimiser est la complexité, sous la contrainte du respect du gabarit du filtre. En posant $H = M - L + 1$, la complexité s'écrit :

$$C = \frac{6M.\log_2(M) - bM + 8}{M - H + 1} \; (Orpec) \tag{19}$$

où $b = 4$ pour l'OLS et $b = 6$ pour la méthode TFD/TFD^{-1}. Pour la méthode Rcos, on a $b = 6-2B_g$, mais nous prendrons $b = 6$, car B_g est relativement faible en pratique.

Deux cas vont se présenter :

- Pour l'overlap-save, la contrainte (respect du gabarit) se traduit par une valeur numérique de H. Il suffit ensuite de déterminer la valeur optimale de M pour cette valeur de H.
- Pour les méthodes TFD/TFD^{-1} et Rcos, la situation est plus complexe : la contrainte se traduit par une équation non-linéaire entre M et H. Il faut ensuite optimiser la complexité, sous la contrainte du respect de cette équation non-linéaire. La résolution directe de ce problème d'optimisation sous contrainte est difficile, mais la méthode que nous proposons ci-dessous fournit une bonne estimation de la solution.

La méthode se décompose en deux grandes étapes :

1. A partir des spécifications, on estime la valeur minimale de H permettant d'atteindre les spécifications (pour l'OLS), ou bien une équation non-linéaire entre M et H (pour les méthodes TFD/TFD^{-1} et Rcos).
2. A partir de H, ou de l'équation reliant M et H, on calcule la valeur de M (taille de la TFD) qui minimise la complexité.

5.2 Estimation de M pour l'Overlap-Save

5.2.1 Détermination de H

En ce qui concerne l'overlap-save, H correspond au nombre de coefficients du filtre dans le domaine temporel. Une bonne estimation de ce nombre est donnée par la formule de Bellanger (éq. 16).

5.2.2 Détermination de la valeur optimale de M

Lorsque H a été estimé, il nous reste à calculer la taille optimale de la TFD, c'est-à-dire la valeur de M qui minimise la complexité. La complexité est donnée par l'équation 19, avec $b = 4$.

Lorsque H est donné, la dérivée de la complexité par rapport à M s'écrit :

$$\frac{\partial C}{\partial M} = \frac{6}{\ln 2} \frac{(-\ln M + (2/3)\ln 2 - 1)\, H + M + \ln M + 1 - 2\,\ln 2}{(M - H + 1)^2} \tag{20}$$

Pour obtenir la valeur optimale de M en fonction de H, il suffit d'annuler cette dérivée. L'expression obtenue fait intervenir la fonction W de Lambert [COR96], d'indice -1, que nous notons W_{-1} :

$$M = (1 - H)\, W_{-1} \left\{ \frac{\exp\,(u(H))}{1 - H} \right\} \tag{21}$$

avec :

$$u(H) = \frac{(1 - (2/3)\ln 2)\, H + 2\, \ln 2 - 1}{1 - H} \tag{22}$$

La figure 12 donne la valeur optimale de M en fonction de H, pour $b = 6$

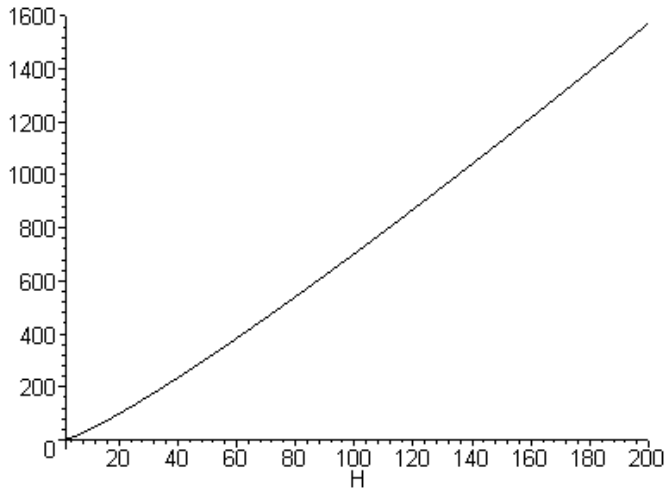

FIG. 12: Valeur optimale de M en fonction de H

Cependant, il ne faut pas oublier que M doit être une puissance de 2 (pour l'efficacité de la FFT, et la validité de la formule donnant la complexité) : on peut évaluer la complexité obtenue avec les deux puissances de 2 qui encadrent la valeur trouvée pour M, et retenir la meilleure. On a alors une taille de TFD qui optimise la complexité suivant l'ordre du filtre :

Ordre Filtre (H)	Taille FFT (M)	Complexité (orpec)
7-10	32	32.3 à 36.5
11-18	64	38.1 à 43.7
19-31	128	44.3 à 49.7
32-55	256	50.1 à 55.8
56-99	512	56 à 61.9
100-181	1024	62 à 68
182-333	2048	68 à 74
334-615	4096	74 à 80

5.3 Estimation de M pour les méthodes TFD et Rcos

5.3.1 Détermination de la relation entre M et H

En annexe, nous montrons que les spécifications imposent une relation non-linéaire entre M et H :

$$H^{a+1} - (M+1)H^a + c = 0 \tag{23}$$

où $a = 2$ pour la méthode TFD et $a = 6$ pour la méthode Rcos. La constante c dépend du gabarit du filtre et de la méthode. Pour déterminer la valeur de cette constante, on peut procéder ainsi :

1. Choisir une valeur arbitraire M_0 de M (puissance de 2).
2. Rechercher la plus grande valeur L_0 de L pour laquelle les spécifications sont respectées.
3. En déduire $H_0 = M_0 - L_0 + 1$
4. Calculer $c = L_0\,(H_0)^a$

5.3.2 Détermination de la valeur optimale de M

Pour déterminer la valeur optimale de M, la procédure proposée consiste à tester les différentes valeurs de M correspondant à des puissances de 2. Pour chaque valeur de M, on procède ainsi :

1. Déterminer la valeur de H en utilisant l'équation 23
2. Calculer la complexité en utilisant l'équation 19, avec $b = 6$

On retient finalement la valeur de M qui donne la plus faible complexité.

L'étape (1) peut donc être réalisée très simplement en testant les valeurs entières de H comprises dans $[1, H_1]$, où $H_1 = \frac{a}{a+1}(M+1)$. En effet, notons $p(H)$ le polynôme à gauche de l'égalité dans l'équation 23. On a :

$$p'(H) = [(a+1)\,H - a(M+1)]\,H^{a-1} \tag{24}$$

Il apparaît que $p(H)$ est décroissant entre 1 et H_1 et croissant entre H_1 et M. De plus, on a $p(1) = c - M$ et $p(M) = c - M^a$, et donc $p(M) < p(1)$. Il s'ensuit que s'il existe une solution dans $[H_1, M]$ il en existe aussi une dans $[1, H_1]$. On a dans ce cas tout intérêt à choisir cette dernière solution, car elle donnera une plus faible complexité.

6 Résultats expérimentaux

La figure 13 donne un exemple de gabarit type. La **résolution** détermine la **largeur minimale de la bande passante**. Cette largeur n'est pas vraiment déterminante pour l'obtention du terme LIT (le gabarit). En revanche, plus la largeur est faible plus nous avons de distorsion de repliement (ceci est dû au rapprochement des pics de repliement). Nous devons donc étudier le pire-cas où nous avons une largeur minimale. Pour comparer les différentes méthodes, nous définirons deux gabarits résumés dans le tableau ci-dessous (Fe est la fréquence d'échantillonnage pour les échantillons complexes d'entrée) :

	Gab 1	Gab 2
Ondulation	0.5 dB	0.1 dB
Bande Utile mini	10% Fe	2% Fe
Bande de garde	2.4% Fe	2.4% Fe
réjection adjacente	40 dB	40 dB
réjection générale	50 dB	50 dB

La réjection adjacente est définie par la réjection dans le canal voisin le plus proche, de largeur utile égale à la résolution. Les tableaux ci-dessous résument les résultats obtenus (pour les deux gabarits) avec les différentes méthodes : valeurs optimales de M, L, H, et complexité correspondante (C) :

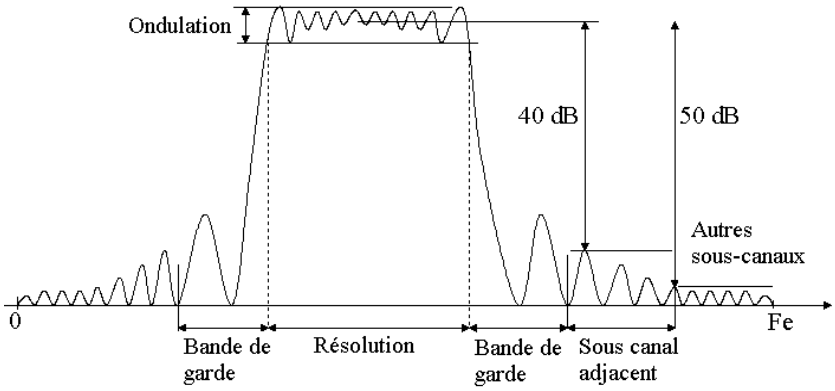

FIG. 13: Gabarit type souhaité

Gabarit 1 :

méthode	M	L	H	C
OLS	512	428	85	59,8
TFD	4096	3865	232	69,9
Rcos	512	450	63	54,6

Gabarit 2 :

méthode	M	L	H	C
OLS	1024	921	104	62,3
TFD	4096	3828	269	70,6
Rcos	512	436	77	56,4

Pour la méthode **OLS**, la formule de Bellanger nous donne un ordre de filtre nécessaire de 85 pour le gabarit 1 et de 104 pour le gabarit 2, ce qui conduit à une complexité de **59,8 orpec** (opérations réelles par échantillon complexe) avec une taille de FFT de 512 pour le gabarit 1, et de **62,3 orpec** avec une taille de FFT de 1024 pour le gabarit 2. L'avantage de cette méthode est qu'il n'y a pas de distorsion de repliement lorsque l'on satisfait la condition d'OLS (car la convolution circulaire est alors égale à la convolution linéaire). La complexité de calcul est faible, mais la commutation de filtre est ici délicate car il est nécessaire de calculer pour tout nouveau filtre la réponse impulsionnelle associée et de charger le nouveau masque de commutation (égal à la TFD de la réponse impulsionnelle).

Pour la méthode **TFD/TFD**$^{-1}$, intéressons nous tout d'abord au gabarit 1. Nous pouvons initialiser la procédure décrite dans la section 4.3.2 avec quatre fois la valeur de M trouvée pour l'OLS, c'est-à-dire $M = 2048$. On trouve alors que la valeur de L doit descendre à 1700 pour assurer le respect du gabarit. Cela correspond à un recouvrement de 17%. Les résultats obtenus pour le terme LIT et pour la distorsion de repliement Dr (cf. équation 11) sont respectivement indiqués sur les figures 14 et 15. Sur ces figures et les suivantes, la fréquence est graduée en points FFT (on peut convertir en fréquence normalisée, c'est à dire en proportion de F_e, en divisant par la valeur de M). La complexité est alors de **72,3 orpec.** Cette complexité peut cependant être légèrement réduite. En effet, la méthode décrite dans la section 5 nous permet, à partir de cette solution initiale ($M = 2048$, $L = 1700$) de déterminer la solution optimale : ($M = 4096$, $L = 3865$, recouvrement=5.6%). La complexité est alors de **69,9 orpec**.

En ce qui concerne le gabarit 2, si on choisit la même valeur initiale $M = 2048$, on trouve que la valeur de L doit descendre à 1638. La complexité est alors de **75,0 orpec**. La méthode décrite dans la section 5 nous permet, à partir de cette solution initiale ($M = 2048$, $L = 1638$, recouvrement=20%) de déterminer la solution optimale : ($M = 4096$, $L = 3828$, recouvrement=6.5%). La complexité est alors de **70,6 orpec**.

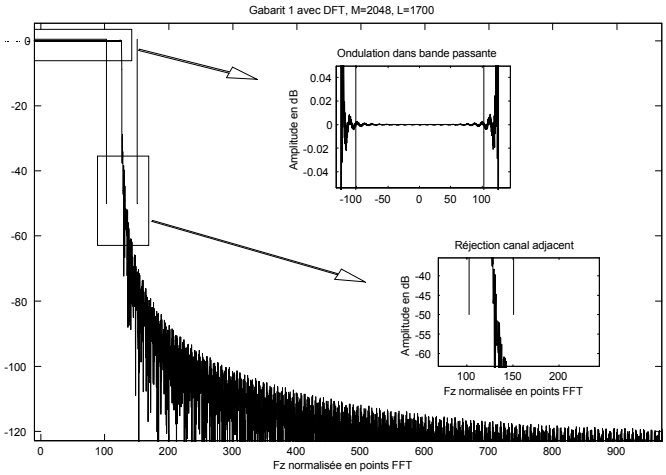

FIG. 14: Terme LIT obtenu pour la méthode TFD/TFD^{-1}(gabarit 1, M=2048, L=1700)

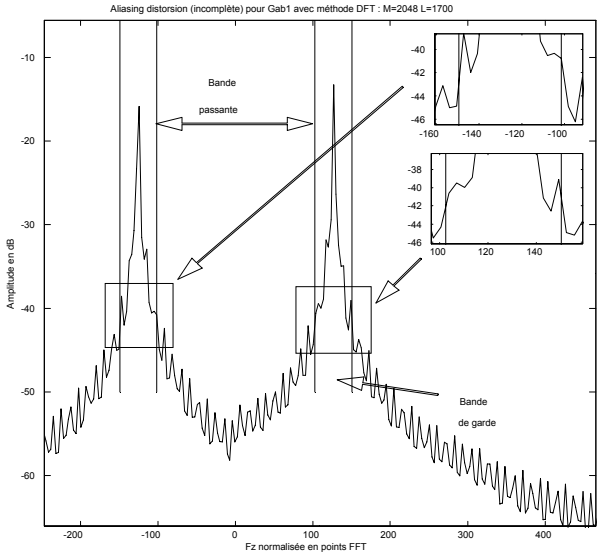

FIG. 15: Distorsion de repliement obtenue pour la méthode TFD/TFD^{-1} (gabarit1, $M = 2048$, $L = 1700$)

Avec la méthode en **cosinus surélevé**, nous obtenons plus vite une distorsion de repliement acceptable mais il est plus difficile d'obtenir nos gabarits. Nous obtenons les spécifications pour les gabarits 1 et 2 pour une taille de FFT de seulement 512. Les figures 16 et 17 indiquent les résultats obtenus pour le gabarit 1. Il nous faut un recouvrement respectif de 12,1% ($L = 450$) et de 14,8% ($L = 436$), soit une complexité de **54,6 orpec** et de **56,4 orpec**. La méthode décrite dans la section 5 nous indique que $M = 512$ est la valeur optimale pour les deux gabarits.

Nous pouvons à présent résumer et commenter les résultats obtenus. Nous avons mentionné le fait que la méthode OLS présente un inconvénient important en pratique : les coefficients de pondération ne sont pas binaires, et doivent être recalculés dès que l'on modifie un élément tel que la bande passante. La méthode TFD/TFD^{-1} résout ce problème, mais, on vient de le voir, au prix d'une complexité plus importante et d'une taille de FFT requise quatre à huit fois plus importante. L'intérêt de la méthode Rcos

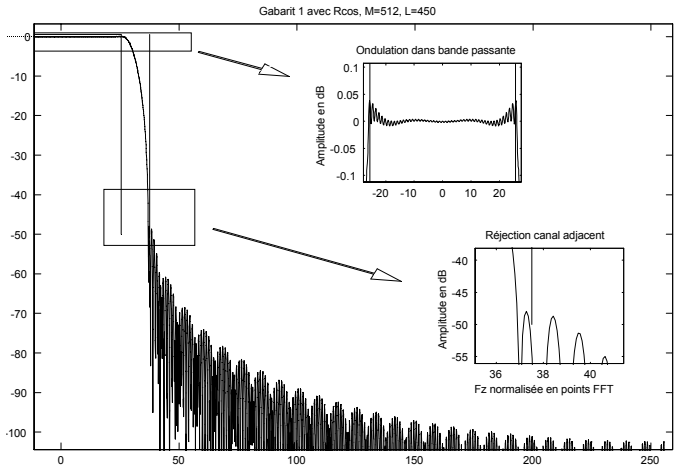

FIG. 16: Terme LIT obtenu pour la méthode Rcos (gabarit 1, $M = 512$, $L = 450$)

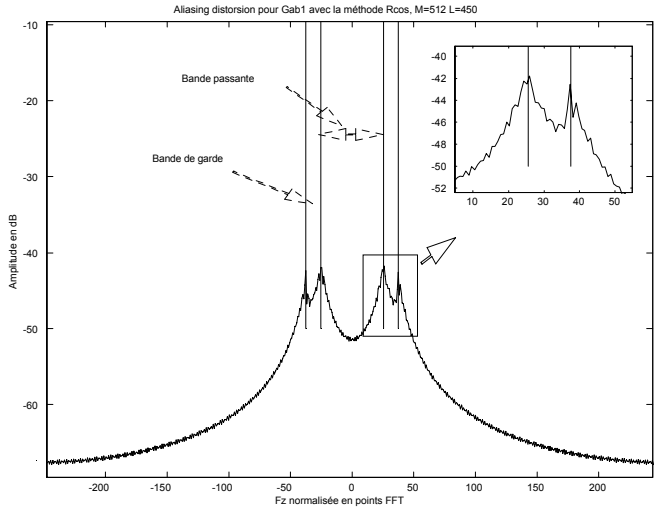

FIG. 17: Distorsion de repliement obtenue pour la méthode Rcos (gabarit 1, $M = 512$, $L = 436$)

est que nous conservons presque en totalité la simplicité de la méthode TFD/TFD^{-1} (simplicité de la sélection des bandes de fréquence désirées puisque la pondération, **unique**, ne concerne que les points de la bande de garde) et que nous arrivons à tailles de FFT et des complexités et qui sont légèrement meilleures qu'avec la méthode OLS.

Le choix de la pondération a été également étudié. Pour cela, nous mettons en place une simulation qui reprend le principe du banc de filtres de la figure 3. Nous mettons en entrée du banc de filtres un bruit blanc gaussien de variance unitaire, nous récupérons le signal filtré en sortie et analysons les spectres du signal en entrée et en sortie avec un périodogramme. Ce processus est itéré pour affiner notre estimation selon le principe de Monte-Carlo. Il reste enfin à faire le rapport des deux spectres estimés pour connaître la fonction de transfert du filtre (valable pour un filtre LIT). Deux exemples de fonction de transfert sont donnés dans la figure 18. Les simulations sont faites pour les valeurs de M et de L satisfaisant le gabarit 1 : la comparaison des pondérations se fait donc pour une complexité de calcul identique. Le tableau suivant résume les performances obtenues avec différentes pondérations (les notations utilisées sont définies dans l'introduction).

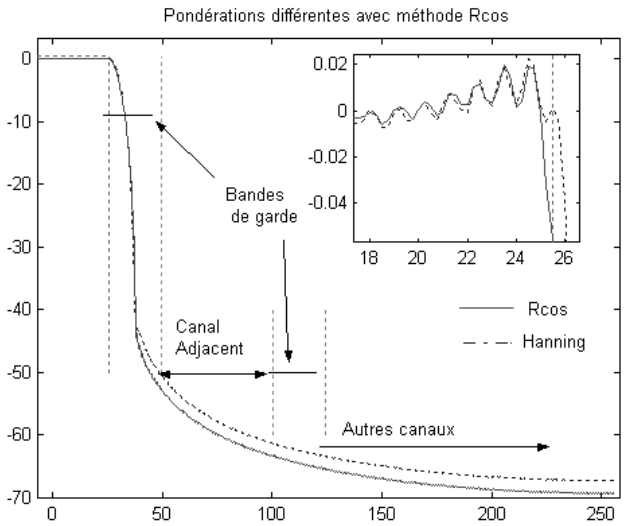

FIG. 18: Simulation de pondérations différentes avec méthode Rcos

	O_{dB}	$\tilde{R}_{dB}$	R_{dB}
Spécifications	0.1	40	50
Cosinus surélevé	0.07	52.7	65.3
Hanning	0.03	50.8	63.5
Hamming	0.03	51.9	64.2
Kaiser(5)	0.03	52.2	64.4

Cette méthode nous permet de vérifier rapidement la "fonction de transfert" du filtre ainsi synthétisé. Elle prend en compte à la fois l'influence du terme LIT et celle des termes de repliement. Cependant, ces derniers termes dépendent du signal en entrée, c'est pourquoi la fonction ainsi réalisée n'est pas tout à fait une fonction de transfert. Seul un coûteux calcul explicite des différents termes de repliement (comme nous l'avons fait pour obtenir les autres figures) permet d'étudier le comportement du filtre dans le cas le plus général.

Nous pouvons dire que les différentes pondérations conduisent à des résultats assez proches, pour un bruit blanc gaussien en entrée. L'ondulation plus élevée pour la pondération en cosinus surélevé s'explique par une décroissance avancée de la fonction de transfert. Nous remarquons également que les spécifications de filtrage sont bien obtenues quelle que soit la pondération. Nous avons même un gain de 10dB à 15dB pour la réjection par rapport aux spécifications. Cela vient du fait que la méthode employée pour calculer la taille de la FFT M et la longueur de recouvrement L est basée sur un calcul analytique pire-cas (tous les termes de repliement sont sommés de manière constructive). Or, la dernière simulation présentée se limite au cas particulier d'un bruit blanc gaussien en entrée.

7 Conclusion

Nous avons introduit trois types d'implémentation de filtre variable à faible complexité. Nous avons rappelé la méthode OLS qui peut être considérée comme une référence, et la méthode TFD/TFD^{-1} qui offre l'avantage d'être très simple d'utilisation mais entraîne une augmentation de la complexité de calcul du fait de l'introduction de termes de repliement importants. Enfin, une troisième méthode est ici proposée pour faire un compromis entre ces deux méthodes. La pondération de la bande de garde offre

l'avantage de réduire considérablement la distorsion de repliement et permet de conserver en grande partie la simplicité de synthèse d'un filtre. De plus, **la complexité de calcul est inférieure à celle de l'OLS**.

Il faut rajouter également que pour les deux dernières méthodes, il est possible de synthétiser simplement des filtres à plusieurs bandes passantes. Il suffit pour cela de sélectionner les points correspondants aux fréquences désirées et de rajouter les points et la pondération nécessaire suivant la méthode employée. Enfin, une permutation circulaire des points en entrée de la deuxième FFT (FFT inverse) permet en outre de faire un décalage en fréquence (switch).

Références

[OS89] Oppenheim and Shaffer, *Discrete-time signal processing*, Prentice-Hall (1989).

[PMC73] J.H. MacClellan and T.W. Parks, "A uniform approach to the design of optimum FIR linear-phase digital filters", *IEEE Trans. on Circuit Theory*, Vol. 20, pp. 697-701, Nov. 1973.

[VAI93] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall Signal Processing series (1993).

[LM96] I.S. Lin & S.K. Mitra, "Overlapped Block Digital Filtering", *IEEE Trans. on Circuits and Systems : Analog and Digital Signal Processing*, Vol. 43, No. 8, August 1996.

[MAL92] H.S. Malvar, *Signal Processing with Lapped Transforms*, 1992, Artech House.

[COR96] Robert M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W Function", *Advances in Computational Mathematics*, Vol. 5, 1996, pp. 329-359 (voir aussi http ://pineapple.apmaths.uwo.ca/~rmc/papers/LambertW/index.html)

[PRO95] J.G. Proakis, *Digital Communications*, Electrical and Computer Engineering, Mac Graw Hill, 3rd Edition, 1995

ANNEXE

Relation non-linéaire entre M et H dans le cas des méthodes TFD et Rcos

Dans la suite, on prendra comme unité de temps la période d'échantillonnage. Les spécifications déterminent la diagonale de la matrice G, comme cela a été expliqué dans l'article. La réponse impulsionnelle du filtre est alors voisine de celle d'un cosinus surélevé de facteur de retombée (roll-off) α ([PRO95], p. 546) :

$$h_{\alpha,T}(t) = h_\alpha(t/T) \tag{25}$$

avec :

$$h_\alpha(t) = \frac{\sin(\pi t)}{\pi t}\frac{\cos(\pi \alpha t)}{1-4\alpha^2 t^2} \tag{26}$$

Les valeurs de α et de T sont déterminées par les spécifications :

$$T = \frac{1}{B_u + B_g} \tag{27}$$

Pour la méthode TFD/TFD^{-1}, on a $\alpha = 0$. Pour la méthode Rcos, on a :

$$\alpha = \frac{B_g}{B_u + B_g} \tag{28}$$

De par le principe du filtrage par blocs, cette réponse impulsionnelle est en fait tronquée, de manière plus ou moins importante, selon la position du point considéré dans le bloc. La qualité du filtre obtenu va donc être directement liée à la troncature maximale, c'est-à-dire à la valeur de H. La proportion de l'énergie du filtre qui est perdue lors de la troncature est :

$$\Psi_{\alpha,T}(H) = \frac{\int_H^\infty \left|h_{\alpha,T}(t)\right|^2 dt}{L\int_{-\infty}^\infty \left|h_{\alpha,T}(t)\right|^2 dt} \tag{29}$$

On notera que cette proportion a été rapportée à un échantillon. Un changement de variable élémentaire permet de montrer que l'on a :

$$\Psi_{\alpha,T}(H) = \frac{1}{L}\Psi_\alpha(H/T) \tag{30}$$

En posant :

$$\Psi_\alpha(x) = \frac{\int_x^\infty |h_\alpha(t)|^2\, dt}{\int_{-\infty}^\infty |h_\alpha(t)|^2\, dt} \tag{31}$$

Les spécifications imposent un niveau de qualité, qui se traduit par une valeur numérique y de $\Psi_{\alpha,T}(H)$. On a donc :

$$\Psi_\alpha(H/T) = L.y \tag{32}$$

Ce qui conduit à :

$$H = T\,\Psi_\alpha^{-1}(L.y) \tag{33}$$

Numériquement, on peut vérifier que la représentation de $\ln\{\Psi_\alpha(x)\}$ en fonction de $\ln(x)$ peut être bien approximée par une droite[2]. Une approximation de la fonction Ψ_α est donc :

$$\Psi_\alpha(x) = \frac{d}{x^a} \tag{34}$$

où les coefficients a et d dépendent uniquement du coefficient de retombée α (et peuvent donc être précalculés). On pourra prendre $a = 2$ pour $\alpha = 0$ et $a = 6$ pour $\alpha > 0.15$ (ces valeurs étant assez proches des valeurs trouvées par une approximation aux moindres carrés). On a alors :

$$\Psi_\alpha^{-1}(z) = \left(\frac{d}{z}\right)^{1/a} \tag{35}$$

D'où :

$$H = T\left(\frac{d}{L.y}\right)^{1/a} \tag{36}$$

Soit :

$$H^a = \frac{c}{L} \tag{37}$$

où la constante c dépend des spécifications et de la valeur de α (mais est indépendante de M, H et L). En remplaçant L par $M - H + 1$, cette relation nous donne une équation non-linéaire liant M et H :

$$H^{a+1} - (M+1)H^a + c = 0 \tag{38}$$

[2]L'approximation est moins bonne pour les faibles valeurs de x, mais on notera que ces valeurs sont de toute façon de peu d'intérêt pratique puisqu'elles correspondent à des fortes valeurs de $\Psi_\alpha(x)$, c'est-à-dire à des gabarits très peu contraignants.