

Simulation de fonctionnement  
de filtres numériques récurrents  
pour le choix de la structure  
à bruit d'arrondi minimal

Round of noise simulation for IIR digital structure choice



**Vincent DEVLAMINCK**

Laboratoire de mesures automatiques, Université des Sciences de Lille-Flandres-Artois, Bât. P 3, 3<sup>e</sup> étage, 59655 VILLENEUVE-D'ASCQ CEDEX.

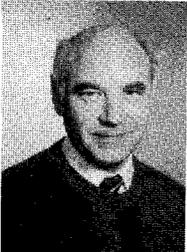
Maître ès Sciences, Université de Lille (1984). DEA Électronique, Lille (1986). Allocataire de recherche, inscrit en thèse de doctorat en électronique à l'Université de Lille-Flandres-Artois au Laboratoire de Mesures Automatiques. Domaines d'intérêt : structures numériques du traitement du signal, applications aux télécommunications.



**François WAUQUIER**

Laboratoire de mesures automatiques. Université des Sciences de Lille-Flandres-Artois, Bât. P 3, 3<sup>e</sup> étage, 59655 VILLENEUVE-D'ASCQ CEDEX.

Maître ès Sciences, Université de Lille (1982). DEA Électronique, Lille (1983). Allocataire d'enseignement à l'IUTA de Lille, inscrit en thèse de doctorat en électronique à l'Université de Lille-Flandres-Artois au Laboratoire de Mesures Automatiques. Domaines d'intérêt : Outils pédagogiques sur le filtrage numérique.



**Jean-Paul DUBUS**

Laboratoire de mesures automatiques, Université des Sciences de Lille-Flandres-Artois, Bât. P 3, 3<sup>e</sup> étage, 59655 VILLENEUVE-D'ASCQ CEDEX.

Ingénieur CNAM (1965), Docteur 3<sup>e</sup> cycle (1968), Docteur ès Sciences Physiques (1974). Professeur d'électronique Traitement du signal à l'UFR IEAA de l'Université de Lille-Flandres-Artois. Domaines d'intérêt : traitement numérique du signal appliqué à l'électronique des Télécommunications et à son enseignement.

**RÉSUMÉ**

Après un rappel de la méthode classique de calcul du bruit d'arrondi dû aux multiplications dans un filtre numérique RII, nous présentons une méthode faisant appel à la représentation des filtres dans l'espace d'états.

Les estimations théoriques du bruit sont vérifiées par comparaison avec des valeurs expérimentales obtenues à l'aide d'un simulateur, pour trois structures de filtres, tous-pôles réalisant une même fonction de transfert  $H(z)$ .

Les résultats sont exprimés en fonction des pôles de  $H(z)$ , ce qui permet une détermination immédiate de la structure la moins bruyante.

**MOTS CLÉS**

Bruit d'arrondi, filtres numériques RII, espace d'état, simulation.

## SUMMARY

After recalling the classical method to calculate roundoff noise into IIR digital filters, we present a method using state-space filter representation. For three different all-poles structures, having the same transfer function  $H(z)$ , we compare the theoretical estimations with results given by simulation. The results depend on poles of  $H(z)$ , that immediately give us a determination of the less noisy structure.

## KEY WORDS

Roundoff noise, IIR digital filters, state space, simulation.

## 1. Introduction

Dans le cadre de la conception des filtres numériques, un des points importants à mettre en valeur est l'introduction du bruit numérique résultant de l'arithmétique des mots de longueur finie.

Parallèlement, dans le domaine de l'enseignement des circuits numériques en Électronique, il est indispensable de mettre l'accent sur ces effets.

Dans ce but, nous avons développé un simulateur de filtres numériques que nous avons voulu d'orientation didactique. La mesure par simulation, des puissances de bruit recueillies en sortie des filtres est une de ces utilisations possibles. Nous la présentons ici en comparaison avec différentes méthodes théoriques d'estimation de ces puissances.

Le développement de deux de ces méthodes permet de mettre en évidence l'intérêt de la représentation des filtres dans l'espace d'état.

Après avoir vérifié par comparaison une bonne concordance entre les estimations théoriques et les valeurs expérimentales obtenues à l'aide du simulateur, nous présentons pour une fonction de transfert  $H(z)$  du deuxième ordre donnée, un exemple de choix de la structure la moins bruyante parmi trois types de réalisations possibles.

## 2. Rappel de la méthode classique d'évaluation du bruit de calcul dans un filtre numérique récursif

La valeur à calculer est celle de la variance du bruit global en sortie du filtre. Ce bruit résulte des différents bruits dus aux opérations arithmétiques intervenant au cours du calcul du filtre.

Parmi les opérateurs arithmétiques que l'on considère, on suppose que la réalisation du filtre prévoit un facteur d'échelle à l'entrée pour réduire au minimum un éventuel dépassement de capacité en sortie des additionneurs. L'effet de quantification des coefficients n'est pas pris en considération pour des filtres à coefficients fixes. Dans l'étude que nous présentons, nous ne considérerons que les sources de bruit engendrées en sortie des multiplieurs. Dans la technologie moderne, l'arrondi est effectué après accumulation en double longueur.

Compte tenu des hypothèses communément admises :

H. 1. deux échantillons différents d'une même source de bruit sont supposés non corrélés;

H. 2. deux sources différentes de bruits (c'est-à-dire associées à deux multiplieurs différents) sont supposées non corrélées;

H. 3. chaque source de bruit est supposée non corrélée avec la séquence d'entrée,

et des propriétés des systèmes linéaires, la variance du bruit en sortie d'un filtre possédant  $M$  sources de bruit, est donnée par la formule [1, 2] :

$$(1) \quad \sigma^2 = \sum_{k=1}^M \sigma_{e_k}^2 \int_{\text{Cercle Unité}} \frac{H_k(z) H_k(z-1)}{2\pi j} z^{-1} dz$$

où  $H_k(z)$  représente la transformée en  $z$  de la réponse impulsionnelle du filtre qui aurait pour entrée la  $k$ -ième source de bruit  $e_k$  de variance  $\sigma_{e_k}^2$ . Ces bruits sont classiquement modélisés par des variables aléatoires uniformément réparties, ce qui conduit pour une technique de quantification par arrondi, à des variances de la forme  $q^2/12$ , où  $q=2^{-b}$  si l'on traite des mots binaires de  $b+1$  bits.

Des formules analytiques ont été développées [3], qui permettent le calcul des normes  $\|H_k\|_2$  pour toute transmittance rationnelle jusqu'à l'ordre quatre. Nous avons préféré, pour effectuer ce calcul, utiliser une méthode qui fait appel à la représentation des filtres dans l'espace d'état [4, 5, 6] et qui est applicable quel que soit l'ordre du filtre.

## 3. Calcul de la variance du bruit en sortie d'un filtre au moyen de sa représentation dans l'espace d'état

## 3.1. EXPRESSION DE LA VARIANCE DU BRUIT EN SORTIE D'UN FILTRE

Soit le filtre de fonction de transfert en  $z$

$$(2) \quad H(z) = \sum_{n=0}^{+\infty} h(nT) z^{-n} = \frac{N(z)}{D(z)}$$

et ses équations d'état

$$(3) \quad \begin{cases} x((n+1)T) = A \cdot x(nT) + B \cdot u(nT) \\ y(nT) = C \cdot x(nT) + D \cdot u(nT) \end{cases}$$

où  $x(nT)$  est le vecteur d'état du filtre à l'instant  $nT$ ,  $u(nT)$  est l'échantillon d'entrée à l'instant  $nT$ ,  $y(nT)$  est l'échantillon de sortie à l'instant  $nT$ .

On a alors :

$$(4) \quad h(nT) = \begin{cases} D & \text{pour } n=0 \\ C \cdot A^{n-1} \cdot B & \text{pour } n>0 \end{cases}$$

et

$$(5) \quad H(z) = D + C \cdot (zI - A)^{-1} \cdot B.$$

Pour calculer le bruit à la sortie du filtre, dû aux multiplications, nous devons considérer chacune des réponses  $h_k(n)$  de la sortie du filtre qui a pour entrée la variable d'état correspondante,  $x_k(n)$ .

Si  $\begin{cases} x_k(n) = \delta_k(n) \\ u(n) = 0, \forall n \end{cases}$  on peut alors écrire :

$$(6) \quad y(n) = C \cdot A^n \cdot x(0) = (C \cdot A^n)_k$$

$h_k(n)$  est donc la  $k$ -ième composante du vecteur ligne  $C \cdot A^n$ .

La variance de l'erreur en sortie du filtre, due aux  $V_k$  sources de bruit de multiplication, ramenées au niveau de la variable d'état  $x_k$  est alors donnée par :

$$(7) \quad \sigma_k^2 = V_k \cdot \frac{q^2}{12} \|h_k\|^2.$$

On définit alors une matrice de bruit  $W$ , symétrique, de dimension  $N$ , qui vérifie l'équation :

$$(8) \quad W = A^T W A + C^T C = \sum_{k=0}^{+\infty} (C \cdot A^k)^T \cdot (C \cdot A^k)$$

dont les éléments  $W_{ij}$  sont donnés par les produits  $h_i \cdot h_j$ .

La variance du bruit en sortie du filtre peut donc encore s'écrire sous la forme :

$$(9) \quad \sigma^2 = \frac{q^2}{12} \sum_{k=1}^N V_k W_{kk}$$

où  $N$  est la dimension du vecteur d'état du filtre.

### 3. 2. ALGORITHME DE CALCUL DE $W$

Le calcul des éléments de  $W$  est un problème bien connu [7], [8], [9]. Nous présentons un algorithme pratique donné par [4] qui permet d'obtenir ces éléments aisément.

Soit  $D(z)$  le polynôme d'ordre  $N$  du dénominateur de la fonction de transfert  $H(z)$  d'un filtre :

$$(10) \quad D(z) = z^N + \sum_{i=1}^{N-1} b_i z^{N-i}.$$

On définit  $(r_0, r_1, \dots, r_N)$ ,  $N+1$  inconnues vérifiant la relation :

$$(11) \quad r_i + \sum_{j=1}^N b_j r_{i-j} = \begin{cases} 1 & \text{pour } i=0 \\ 0 & \text{pour } i \geq 1 \end{cases}$$

ainsi qu'une matrice  $R$ , symétrique, de Toeplitz de dimension  $N$  dont le premier rang est donné par :

$$[r_0, r_1, r_2, \dots, r_{N-1}]$$

On introduit d'autre part une matrice  $X$  dont les vecteurs colonnes sont définis par :

$$(12) \quad \begin{cases} X(1) = C^T \\ X(k+1) = A^T \cdot X(k) + b_k C^T, \quad \forall k \text{ dans } [1, N] \end{cases}$$

$W$  est alors donnée par la relation :

$$(13) \quad W = X \cdot R \cdot X^T.$$

## 4. Valeurs expérimentales de la puissance de bruit d'arrondi en sortie d'un filtre

### 4. 1. DESCRIPTION DE L'INSTRUMENTATION DE CALCUL DES VALEURS EXPÉRIMENTALES DE LA PUISSANCE DE BRUIT

L'étude théorique, précédemment décrite, demande une confirmation expérimentale. Nous avons mesuré la puissance de bruit d'arrondi en sortie de différents types de filtres dans le but de vérifier l'exactitude des résultats théoriques et nous permettre d'en apprécier la portée pratique.

Ces mesures ont été faites à l'aide d'un logiciel de simulation qui tient compte des caractéristiques réelles des composants utilisés pour la réalisation des structures de circuits numériques.

Ce logiciel a été conçu et mis au point au laboratoire sur une architecture VME (Microprocesseur MC 68000) munie d'un système d'exploitation type VERSADOS. Il comprend une capture de schéma (logiciel écrit en Assembleur 68000) qui assure une interactivité entre l'utilisateur et la structure de données qui décrit le schéma en mémoire, et un logiciel simulateur (écrit en PASCAL). Ce dernier fournit à chaque instant, les valeurs des sorties de chaque opérateur et à chaque période d'horloge, le vecteur d'état du système (contenus des mémoires internes).

Ce simulateur reçoit les échantillons d'entrée fournis, soit par une fonction, soit par une liste de valeurs, sous forme de mots binaires dont le format est à définir. D'une manière générale, il tient compte des formats finis des registres internes et des bus de transfert de données, ainsi que des temps de propagation de chaque opérateur.

Ces paramètres ainsi que les valeurs des coefficients du filtre sont fournis par l'utilisateur lors de la saisie du schéma. Enfin, il est possible de le faire travailler en utilisation au choix la technique de l'arrondi ou celle de la troncature.

Grâce à ce simulateur, il nous a été très facile d'obtenir les valeurs expérimentales de la puissance de bruit en sortie de différentes structures de filtres du deuxième ordre (structures, directe d'ordre II, premier ordre en cascade, treillis en cascade) (schémas A, B et C).

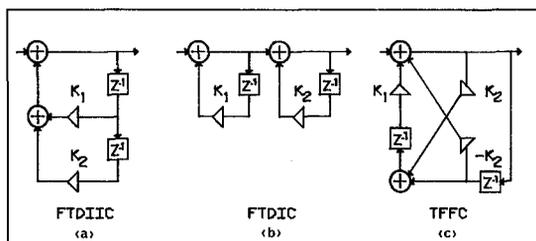


Fig. 1. — Différentes structures de filtres RII du deuxième ordre expérimentées : (a) directe d'ordre II; (b) premier ordre en cascade; (c) treillis en cascade.

Pour chacune de ces structures, nous avons comparé les valeurs des échantillons de sortie du simulateur avec ceux de la sortie d'un filtre (filtre théorique) analogue mais travaillant avec la précision de l'ordinateur (virgule flottante avec 52 bits de mantisse). Les échantillons d'entrée sont fournis par un générateur de bruit gaussien de valeur moyenne nulle et de variance unité. Sachant que pour une puissance donnée, c'est un signal de ce type qui possède l'entropie maximale[11], il permettra donc de caractériser au mieux l'effet, sur la sortie, des différentes sources de bruit considérées.

La simulation ne devant rendre compte que des bruits dus aux multipliers les résultats du simulateur ont été comparés à ceux du filtre théorique pour lequel les échantillons d'entrée et les coefficients ont été préalablement quantifiés de façon à n'introduire aucune source de bruit supplémentaire.

#### 4.2. RÉSULTATS DE LA VÉRIFICATION EXPÉRIMENTALE

L'expression (9) montre que la puissance réduite de bruit en sortie du filtre est de la forme :

$$(14) \quad \frac{\sigma^2}{(q/12)} = \sum_{k=1}^N V_k W_{kk}$$

où  $V_k$  est le nombre de sources de bruit, ramenées à la variable d'état  $x_k$ .  $W_{kk}$  est le  $k$ -ième terme de la diagonale de la matrice  $W$  qui n'est fonction que des seules valeurs des coefficients du filtre.

Il en résulte que sous réserve des hypothèses H.1, H.2 et H.3,  $\sigma^2/(q^2/12)$  doit être indépendant du nombre de bits de quantification.

La figure 2 représente la puissance réduite de bruit en sortie du filtre obtenue à l'aide du simulateur, en fonction du nombre de bits de quantification, et pour différents couples des valeurs  $K_1$  et  $K_2$ . (Toutes les simulations sont faites pour un signal d'entrée comprenant 2000 échantillons.)

On constate sur cette figure que la fluctuation relative de la puissance réduite du bruit en sortie du filtre, reste indépendante du nombre de bits de quantification.

Une seconde utilisation du simulateur a consisté à comparer, pour un nombre de bits donné (24 dans les exemples traités), les valeurs des puissances réduites de bruit obtenues à l'aide de l'expression (14) avec celles obtenues expérimentalement grâce au simulateur. Cette comparaison effectuée pour les trois types de filtres représentés par les schémas a, b et c donne les résultats reproduits en figures 3, 4, 5.

On remarque sur ces figures l'excellente correspondance des résultats issus des expressions théoriques et des simulations du filtre réel.

#### 4.3. ÉTUDE COMPARATIVE DES BRUITS DE MULTIPLICATION ENGENDRÉS PAR TROIS TYPES DE STRUCTURES

Cette étude a été effectuée en considérant trois structures qui réalisent une même fonction de transfert  $H(z)$ . Les filtres sont généralement réalisés à l'aide de structures du premier et deuxième ordre.

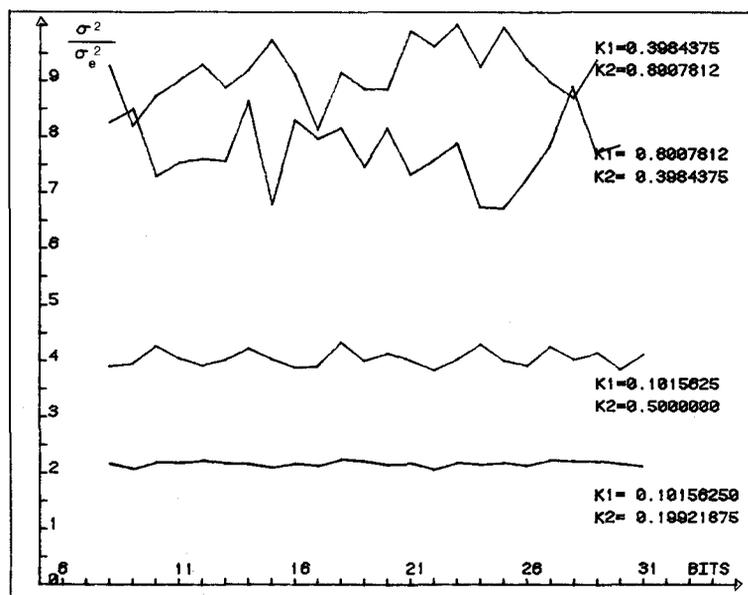


Fig. 2. — Puissance réduite du bruit de sortie du filtre premier ordre en cascade en fonction du nombre de bits de quantification.

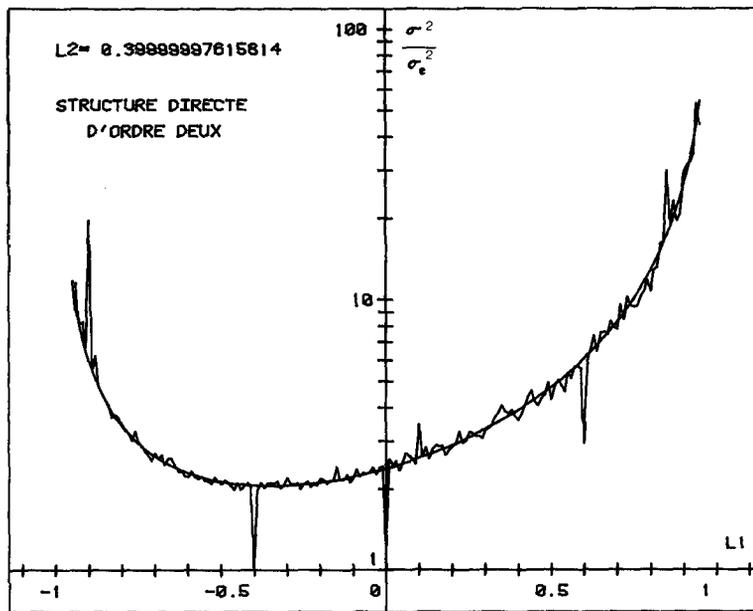


Fig. 3. — Comparaison entre résultats théoriques et expérimentaux de la puissance réduite de bruit en fonction des pôles pour le filtre FTDIC.

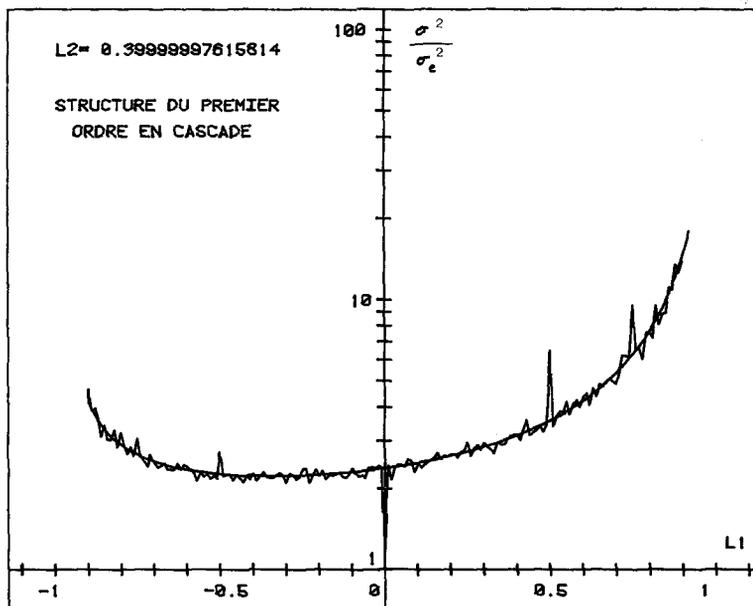


Fig. 4. — Comparaison entre résultats théoriques et expérimentaux de la puissance réduite de bruit en fonction des pôles pour le filtre FTDIC.

Nous avons étudié systématiquement un filtre de fonction de transfert du type :

$$H(z) = \frac{1}{1 + b_1 z^{-1} + b_2 z^{-2}}$$

que nous avons simulé suivant les trois types de réalisations représentées par les schémas *a*, *b* et *c*.

Les coefficients  $K_1$  et  $K_2$  sont déterminés pour chaque structure de façon qu'elle soit une réalisation de la même fonction de transfert.

Nous représentons pour cette raison, la puissance réduite de bruit en sortie, dans les trois cas, en fonction des deux pôles de la fonction de transfert  $H(z)$  qu'ils réalisent.

Ces résultats expérimentaux sont résumés par les figures 6, 7 et 8.

Nous avons choisi de représenter la puissance réduite de bruit en sortie des filtres en fonction de  $L_1$  et  $L_2$ , les deux pôles de la fonction de transfert  $H(z)$  donnée.

Nous ne donnons ici que des courbes pour le pôle  $L_2$  positif car on vérifie expérimentalement que la puissance de bruit réduite  $\sigma^2/\sigma_e^2$  présente une symétrie par rapport à l'axe  $L_1=0$ , suivant la relation :

$$\frac{\sigma^2}{\sigma_e^2} = f(L_1, L_2) = f(-L_1, -L_2)$$

prévisible à partir des estimations théoriques.

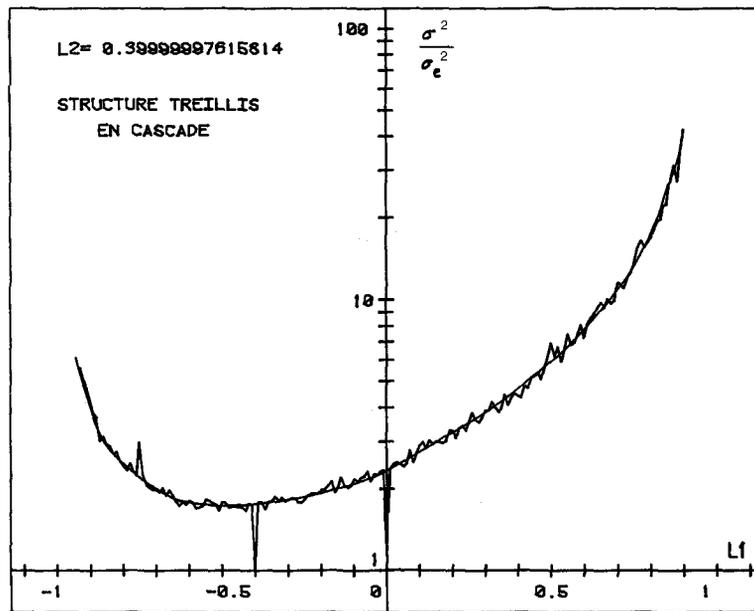


Fig. 5. — Comparaison entre résultats théoriques et expérimentaux de la puissance réduite de bruit en fonction des pôles pour le filtre TFFC.

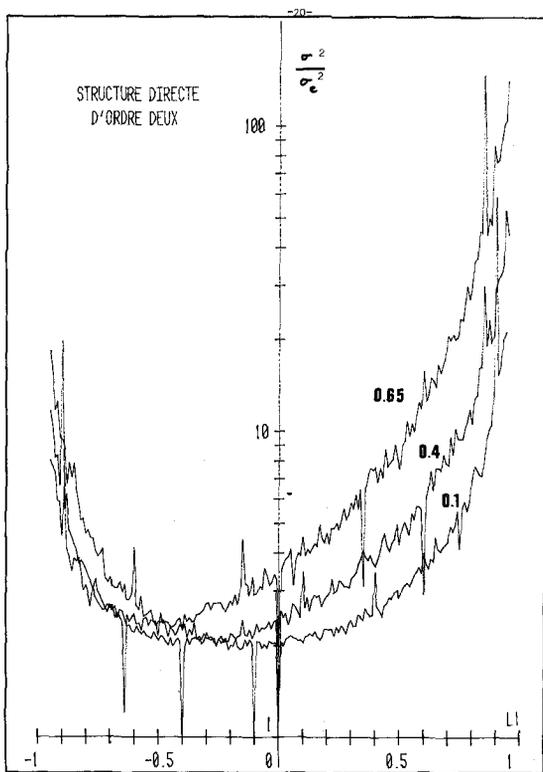


Fig. 6. — Puissance de bruit en sortie du filtre FTDIC en fonction du pôle  $L_1$  pour les valeurs 0,1, 0,4 et 0,65 du pôle  $L_2$ .

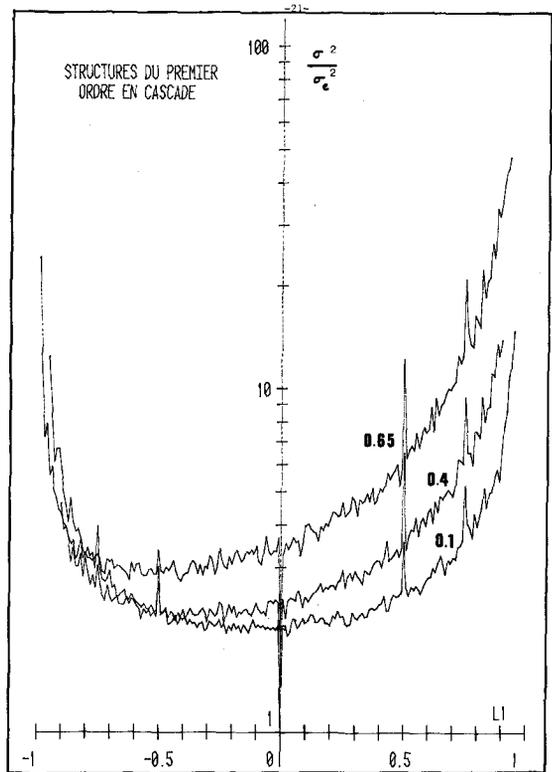


Fig. 7. — Puissance de bruit en sortie du filtre FTDIC en fonction du pôle  $L_1$  pour les valeurs 0,1, 0,4 et 0,65 du pôle  $L_2$ .

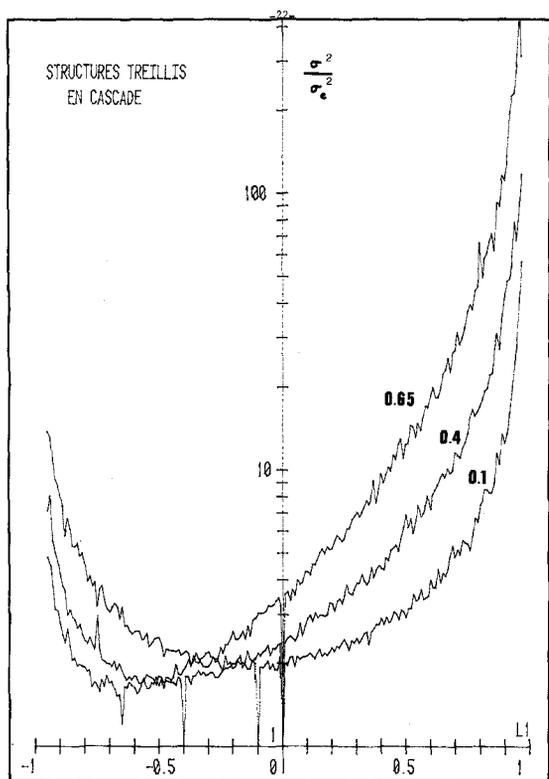


Fig. 8. — Puissance de bruit en sortie du filtre TFFC en fonction du pôle  $L_1$  pour les valeurs 0,1, 0,4 et 0,65 du pôle  $L_2$ .

De plus, si nous définissons une relation d'ordre notée : « > » par :

$a > b$  : si la puissance réduite de bruit en sortie de la structure  $a$  est plus faible que celle de  $b$  [pour une même valeur du couple  $(L_1, L_2)$ ].

Nous constatons les relations suivantes :

si  $L_2 > 0$  :

si  $L_1 < 0$ ,  $FTTC > FTDIIC > FTDIC$ ,

si  $L_1 > 0$ ,  $FTDIC > FTDIIC > FTTC$ ;

si  $L_2 < 0$  les résultats sont inversés (d'après les remarques de symétrie) :

si  $L_1 < 0$ ,  $FTDIC > FTDIIC > FTTC$ ,

si  $L_1 > 0$ ,  $FTTC > FTDIIC > FTDIC$ .

## 5. Conclusion

Les représentations en fonction non plus des coefficients des filtres mais des pôles de la fonction de transfert  $H(z)$  permettent de comparer immédiatement les puissances réduites de bruit en sortie des trois structures proposées. Nous définissons ainsi un

critère de choix entre ces différentes structures qui réalisent toutes la même fonction de transfert  $H(z)$ , mettant par là même en évidence la non-équivalence entre les différentes réalisations d'une fonction  $H(z)$  donnée.

Cette étude a été grandement facilitée par l'utilisation du simulateur que nous avons mis au point. Outre l'avantage de sa rapidité, il permet en plus de s'affranchir de toutes les hypothèses restrictives (H. 1, H. 2, H. 3) nécessaires aux calculs théoriques des puissances réduites, confirmant encore ainsi, l'indispensable rôle de la simulation en matière de réalisations des circuits numériques.

La généralisation de cette étude à des structures d'ordres plus élevés ne se conçoit bien entendu que par la simulation, compte tenu de la relative lourdeur des calculs théoriques.

Manuscrit reçu le 15 juin 1987.

## BIBLIOGRAPHIE

- [1] R. RABINER et B. GOLD, *Theory and applications of digital filters*, Prentice Hall Inc., Englewood Cliffs, New Jersey.
- [2] M. BELLANGER, *Traitement numérique du signal*, Masson.
- [3] E. J. JURY, *Theory and Application of the Z transform Method*, John Wiley, 1964.
- [4] C. T. MULLIS et R. A. ROBERTS, Synthesis of minimum Roundoff Noise fixed point digital filters, *IEEE Transactions on circuits and systems*, CAS-23, n° 9, septembre 1976.
- [5] C. T. MULLIS et R. A. ROBERTS, Roundoff noise in digital filters: Frequency transformations and invariants, *IEEE Transactions on acoustics, Speech and Signal processing*, ASSP24, n° 6, Dec. 1976.
- [6] F. J. TAYLOR et J. W. MARSCHALL, Computer Aided Design and analysis of standard IIR Architectures, *IEEE Circuits and Systems magazine*, 3, n° 4, Dec. 1981.
- [7] C. S. BERGER, A numerical solution of the matrix equation  $P = \phi P \phi^T + S'$ , *IEEE Transaction on automat. Contr.*, AC-16, August 1971.
- [8] J. A. HEINEN, A technique for solving the extended discrete Liapunov Matrix Equation, *IEEE Transaction on automat. Contr.*, AC-17, Feb. 1972.
- [9] S. BARNETT, Simplification of the Liapunov Matrix equation:  $A^T P A - P = -Q$ , *IEEE Transaction on Automat. Cont.*, AC-19, August 1974.
- [10] M. LABARRERE, J. P. KRIEF et B. GIMONET, *Le filtrage et ses applications*, Collection Sup. Aero, Capadues Editions.
- [11] J.-P. DUBUS, Cours d'électronique, traitement du signal, *Maîtrise EEA*, Université de Lille-Flandres-Artois.