

Échantillonnage adapté à l'extraction des signaux

Matched sampling for denoising

par **Bernard LACAZE**

TéSA, 2 rue Camichel, BP 7122 Toulouse cedex 7, France
e-mail : Bernard.Lacaze@tesa.prd.fr

résumé et mots clés

Le préfiltrage est une opération qui est utilisée pour éliminer le bruit dans les bandes de fréquence où le signal est absent ou peu présent. En particulier, c'est une manière d'éviter des repliements de bandes qui peuvent s'avérer très pénalisants, notamment dans le cas d'échantillonnage. Dans cet article, on montre que cette opération doit se faire de manière systématique et adaptée, en tenant compte des valeurs relatives des spectres de puissance du signal et du bruit. D'autre part, on montre que les formules d'interpolation optimales, dans le cas de l'échantillonnage périodique, ont des propriétés spécifiques, et notamment, rendent l'erreur indépendante du temps.

échantillonnage périodique, processus stationnaires, condition de Nyquist, repliement de spectres

abstract and key words

Pre-filtering is used in order to eliminate noise in frequency bands in which there is no signal (or a few signal) of interest. More precisely, this kind of operation is necessary to avoid aliasing in the sampling process. In this paper, it is shown that this pre-filtering operation has to be done taking into account the respective power spectral densities of signal and noise. Moreover, optimal formulae of interpolation are derived taking into account this pre-filtering operation. Specific properties are given in the periodic sampling case. Particularly, the reconstruction error is shown to be time independent.

periodic sampling, stationary processes, Nyquist condition, spectral aliasing.

1. introduction

1) Un signal transmis dans un canal (à travers l'atmosphère, un câble ou tout autre support) arrive à l'utilisateur après avoir subi certaines transformations. Le modèle le plus simple consiste à considérer qu'il a subi une transformation linéaire et qu'il arrive mélangé à un bruit dû au milieu traversé. Le récepteur traite le processus reçu. Dans le cas d'un traitement numérique, il procède

de à un échantillonnage et une quantification, puis utilise un algorithme adapté. On négligera l'erreur due à la quantification pour n'étudier que les problèmes d'échantillonnage, et uniquement dans le cas périodique (ou « uniforme »). On sait que la période d'échantillonnage doit être adaptée à l'occupation spectrale du signal. Plus elle est faible, et plus la reconstruction du signal pourra être améliorée. En même temps, la masse de données à traiter augmente, ce qui fait qu'il y a un compromis à rechercher entre précision d'estimation et complexité de calcul.

Le point de vue adopté dans cet article est celui du « filtrage de Wiener » dans le cadre aléatoire stationnaire. Il s'agit de minimiser la distance quadratique moyenne entre le signal et sa meilleure estimation (linéaire). La solution, dans la situation que l'on vient de décrire, est dans la conscience collective, et peut se trouver par exemple dans [15] (chap. 6) et [14] (chap. 14), sous des formes approchées.

Une manière d'améliorer l'estimation consiste à éliminer le bruit qui est en dehors de la bande de fréquence du signal, avant échantillonnage, à l'aide d'un filtrage sélectif. Cette opération est intuitivement bénéfique, largement pratiquée (et pour d'autres motifs éventuellement, comme la séparation de sources), et l'amélioration est facile à calculer. Mais rien ne dit qu'elle est optimale. Le but de cet article est donc de rechercher le filtre linéaire à utiliser avant un échantillonnage périodique, de façon à minimiser l'erreur d'estimation d'un signal.

2) Supposons que $\mathbf{A} = \{A(t), t \in \mathbb{R}\}$ est un « signal » noyé dans un « bruit » $\mathbf{B} = \{B(t), t \in \mathbb{R}\}$. La somme $\mathbf{Z} = \mathbf{A} + \mathbf{B}$ étant observée à des instants périodiques (par exemple les entiers), on cherche à estimer linéairement $A(n)$, pour chaque $n \in \mathbb{Z}$. Le fait de prendre une période d'échantillonnage égale à 1 ne nuit pas à la généralité du propos, mais simplifie les notations. Dans le contexte stationnaire non corrélé, la meilleure interpolation $\tilde{A}(n)$ de $A(n)$ est le résultat du filtrage (numérique) de la suite $\mathbf{Z}' = \{Z(n), n \in \mathbb{Z}\}$ dans le filtre de gain

$$H(\omega) = \left[\frac{s'_A}{s'_A + s'_B} \right](\omega)$$

$s'_A(\omega)$ et $s'_B(\omega)$ sont les « spectres repliés » définis par

$$s'_A(\omega) = \sum_{k \in \mathbb{Z}} s_A(\omega + 2k\pi) \quad s'_B(\omega) = \sum_{k \in \mathbb{Z}} s_B(\omega + 2k\pi)$$

lorsque l'on suppose l'existence de densités spectrales de puissance $s_A(\omega)$ et $s_B(\omega)$ définies par :

$$\begin{cases} E[A(t)A^*(t-\tau)] = \int_{-\infty}^{\infty} e^{i\omega\tau} s_A(\omega) d\omega \\ E[B(t)B^*(t-\tau)] = \int_{-\infty}^{\infty} e^{i\omega\tau} s_B(\omega) d\omega. \end{cases} \quad (1)$$

L'erreur d'interpolation est définie par

$$\sigma^2 = E[|A(n) - \tilde{A}(n)|^2] = \int_{-\pi}^{\pi} \left[\frac{s'_A s'_B}{s'_A + s'_B} \right](\omega) d\omega.$$

Les expressions de $H(\omega)$ et de σ^2 données ci-dessus sont des cas particuliers de l'appendice 1. Il est évident qu'un préfiltrage (analogique) de \mathbf{Z} avant échantillonnage peut réduire σ , et cette opération est couramment effectuée dans la réalité. Par exemple, si

$$\begin{cases} s_A(\omega) = 1, \omega \in (-\pi, \pi) \\ s_B(\omega) = 2/3, \omega \in (-2\pi, 2\pi) \end{cases}$$

et 0 ailleurs, un filtrage passe-bande de \mathbf{Z} sur $(-\pi, \pi)$, qui élimine une partie du bruit, assure un gain de 30 % (σ^2 passe de $8\pi/7$ à $4\pi/5$). En effet, en l'absence de préfiltrage, le phénomène de repliement fait que la valeur de la densité spectrale du bruit (échantillonné) est égale à $4/3$ sur $(-\pi, \pi)$, alors qu'elle ne vaut plus que $2/3$ après le préfiltrage (qui supprime le repliement). La figure 1 décrit les spectres équivalents du signal et du bruit (échantillonnés) avec et sans le pré-filtrage.

Notons que cette opération est inutile si l'on observe \mathbf{Z} sur \mathbb{R} tout entier et non seulement sur \mathbb{Z} . Cet exemple est évidemment caricatural, mais une opération supprimant une part importante du bruit, même si elle détériore dans une faible mesure le signal, peut être avantageuse, comme le montreront les exemples de la section 3.

3) Comme expliqué au-dessus, la construction précédente amène à poser la question suivante. Dans le cas de l'extraction d'un signal \mathbf{A} mélangé à un bruit \mathbf{B} , quel filtrage faut-il effectuer avant un échantillonnage périodique, pour espérer la meilleure interpolation linéaire possible $\tilde{A}(t)$ de $A(t)$, $t \in \mathbb{R}$?

Plus précisément, définissons le processus $\mathbf{U} = \{U(t), t \in \mathbb{R}\}$ par

$$U(t) = \int_{-\infty}^{\infty} f(u) Z(t-u) du, t \in \mathbb{R}. \quad (2)$$

$f(t)$ est supposée suffisamment régulière, avec une transformée de Fourier bornée $F(\omega)$ telle que

$$F(\omega) = \int_{-\infty}^{\infty} f(u) e^{-i\omega u} du. \quad (3)$$

Maintenant, on extrait de \mathbf{U} la suite d'échantillons $\mathbf{U}' = \{U(n), n \in \mathbb{Z}\}$ telle que

$$\begin{cases} U(n) = \int_{-\infty}^{\infty} f(u) Z(n-u) du \\ Z(t) = A(t) + B(t). \end{cases} \quad (4)$$

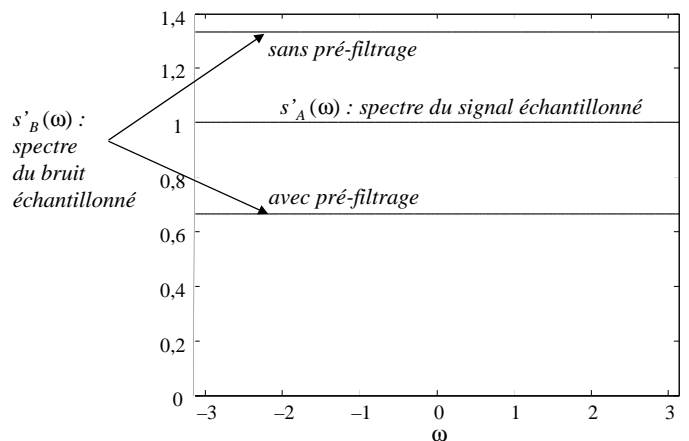


Figure 1. – Spectres du signal échantillonné, du bruit sans et avec préfiltrage.

Le problème étudié dans cet article est la recherche du $f(t)$ (ou du $F(\omega)$) qui permet d'obtenir la meilleure estimation possible $\tilde{A}(t)$ (au sens de la moyenne quadratique) du signal $A(t), t \in \mathbb{R}$, en utilisant la suite des observations U' .

La question posée est sans intérêt lorsque le processus Z est sans repliement de spectre [13], c'est-à-dire lorsque, à ω fixé quelconque, au plus un seul des $[s_A + s_B](\omega + 2k\pi), k \in \mathbb{Z}$, est non nul. Dans cette situation, l'estimation est indépendante de $f(t)$, dans la mesure où $F(\omega)$ ne s'annule pas sur le support spectral de Z . Mais l'on sait d'autre part qu'un processus sans repliement n'est pas réaliste.

Notons que le problème que l'on étudie ici est d'autant plus intéressant que l'on ne dispose jamais dans la réalité des observations de $Z(n), n \in \mathbb{Z}$, mais tout au plus d'un « moyennage » autour des instants théoriques d'échantillonnage. Une telle opération, sous sa forme la plus simple, revient à prendre $f(t)$ constante autour de $t = 0$, et 0 ailleurs [4]. Donc, on peut considérer qu'un préfiltrage de Z existe dans tous les cas. Autant l'adapter au mieux. Le cas où le temps de prise est aléatoire est évoqué dans [8].

4) La suite $U' = \{U(n), n \in \mathbb{Z}\}$ des échantillons, définie par (4), peut être vue comme l'échantillonnage périodique idéal du processus $U = \{U(t), t \in \mathbb{R}\}$, lui-même résultat du filtrage de Z dans le filtre de réponse impulsionnelle $f(t)$ et de réponse (gain) fréquentielle $F(\omega)$. La détermination de la meilleure interpolation linéaire $\tilde{A}(t)$ de $A(t)$ sur U' est calculée dans l'appendice 1, ainsi que l'erreur d'estimation $\sigma_{t,F}$, en fonction de F, s_A et s_B , où

$$\sigma_{t,F}^2 = E \left[|A(t) - \tilde{A}(t)|^2 \right].$$

Dans le cas où $s_Z = s_A + s_B$ est sujet au repliement (le seul cas intéressant), $\sigma_{t,F}$ dépend effectivement de t . Pour comparer valablement deux filtres différents F_1 et F_2 , il est préférable de travailler avec une mesure d'erreur indépendante de t . En effet, la plupart du temps, l'on trouvera des valeurs de t telles que $\sigma_{t,F_1} < \sigma_{t,F_2}$ et d'autres valeurs de t pour lesquelles ce sera l'inégalité opposée.

C'est pour cette raison que l'on utilisera la quantité ε_F obtenue en faisant une moyenne sur t de $\sigma_{t,F}$:

$$\varepsilon_F = \sqrt{\int_0^1 \sigma_{t,F}^2 dt}. \quad (5)$$

Il s'agit maintenant de déterminer $F(\omega)$ (ou la quantité équivalente $f(t)$) qui rend ε_F minimum. ε_F s'exprime en fonction de $F(\omega), s_A(\omega), s_B(\omega)$ (voir l'appendice 1) et la solution du problème s'écrit de manière simple. Cela sera exposé dans la section 2 (et démontré dans l'appendice 2). La section 3 sera consacrée à l'étude d'exemples. Dans la section 4, on étudiera le cas voisin d'un signal soumis à un changement d'horloge, ce qui est la règle lorsque le temps de parcours est aléatoire.

2. la minimisation de ε_F

2.1. théorème

1) Soit $l(\omega)$ une fonction mesurable de $(-\pi, \pi)$ dans \mathbb{Z} vérifiant, pour tout $k \in \mathbb{Z}$:

$$\left[\frac{s_A^2}{s_A + s_B} \right](\omega + 2\pi l(\omega)) \geq \left[\frac{s_A^2}{s_A + s_B} \right](\omega + 2\pi k) \quad (6)$$

Alors, une réponse impulsionnelle $f_0(t)$ qui minimise ε_F est définie par

$$\begin{cases} F_0(\omega) = 1 \text{ si } \bar{\omega} = l(\underline{\omega}) \\ F_0(\omega) = 0 \text{ si } \bar{\omega} \neq l(\underline{\omega}) \\ f_0(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_0(u) e^{iut} du, t \in \mathbb{R}, \end{cases} \quad (7)$$

où $\omega = 2\pi\bar{\omega} + \underline{\omega}, \bar{\omega} \in \mathbb{Z}, \underline{\omega} \in (-\pi, \pi)$. $\bar{\omega}$ et $\underline{\omega}$ ont un rapport évident avec les parties entière et fractionnaire de ω (modulo 2π).

2) L'erreur ε_{F_0} et la meilleure interpolation linéaire $\tilde{A}_{F_0}(t)$ vérifient

$$\begin{cases} \varepsilon_{F_0}^2 = \int_{-\infty}^{\infty} s_A(\omega) d\omega - \int_{-\pi}^{\pi} \left[\frac{s_A^2}{s_A + s_B} \right](\omega + 2l(\omega)\pi) d\omega \\ \tilde{A}_{F_0}(t) = \sum_{k \in \mathbb{Z}} a_0(t-k) U(k) \\ a_0(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{s_A}{s_A + s_B} \right](\omega + 2l(\omega)\pi) e^{it(\omega + 2l(\omega)\pi)} d\omega. \end{cases} \quad (8)$$

3) L'erreur quadratique moyenne σ_{t,F_0} est indépendante de t , et donc $\varepsilon_{F_0} = \sigma_{t,F_0}$.

Les démonstrations sont données dans l'appendice 2.

2.2 remarques

1) $l(\omega)$ dépend des valeurs relatives de $s_A(\omega)$ et $s_B(\omega)$, et l'inéquation (6) peut avoir plusieurs solutions, parce que c'est par une inégalité large que $l(\omega)$ est définie.

2) Le processus optimal U , obtenu par filtrage de Z à travers le filtre de gain $F_0(\omega)$, est sans repliement par construction, puisque, quelque soit ω , un seul des $F_0(\omega + 2k\pi), k \in \mathbb{Z}$ est susceptible d'être non nul.

Cette suppression de repliement est adaptée aux spectres de puissance respectifs $s_A(\omega)$ et $s_B(\omega)$. Elle précise une opération classique dans la pratique, qui consiste à supprimer le bruit qui déborde des lobes principaux du signal.

3) La propriété précédente est encore vraie pour le produit de $F_0(\omega)$ par n'importe quelle fonction. En conséquence,

$F(\omega) = [F_0\Phi](\omega)$ est aussi solution du problème, quand $\Phi(\omega)$ est non nul sur le support de $F_0(\omega)$.

4) Le cas $F(\omega) = 1, \omega \in \mathbb{R}$ correspond à l'échantillonnage idéal ($U(n) = Z(n)$). Bien que cette situation ne rentre pas dans le cadre utilisé ici (puisque $f(t)$ n'existe pas au sens des fonctions), il est naturel de comparer les erreurs dans ce cas (notées $\sigma_{t,1}$ et ε_1), l'erreur dans le cas optimum ($\sigma_{t,F_0} = \varepsilon_{F_0}$), et les erreurs dans des cas réalistes approchant l'échantillonnage idéal (par exemple $f(t) = 1/2\delta$ sur $(-\delta, \delta)$ et 0 ailleurs, comme dans [4]).

5) En faisant $s_B \equiv 0$ dans (6) et (8), on se ramène au cas particulier étudié dans [10] de l'échantillonnage sans bruit. Dans cette situation, on retiendra parmi les $s_A(\omega + 2k\pi), k \in \mathbb{Z}$, celui qui est prépondérant. Dans le cas où

$$\inf_{\omega \in (-\pi, \pi)} s_A(\omega) \geq \sup_{\omega \notin (-\pi, \pi)} s_A(\omega)$$

le préfiltrage F_0 sera un filtre passe-bande idéal sur $(-\pi, \pi)$, et la formule d'interpolation optimale se réduit à la formule habituelle

$$\tilde{A}_{F_0}(t) = \sum_{k \in \mathbb{Z}} \frac{\sin \pi(t-k)}{\pi(t-k)} U_k.$$

6) L'erreur $\varepsilon_{F_0}^2$ peut s'écrire sous la forme

$$\begin{cases} \varepsilon_{F_0}^2 = \int_{\Delta_{F_0}} \left[\frac{s_A s_B}{s_A + s_B} \right](\omega) d\omega + \int_{\bar{\Delta}_{F_0}} s_A(\omega) d\omega \\ \Delta_{F_0} = \{\omega; F_0(\omega) = 1\}, \bar{\Delta}_{F_0} = \{\omega; F_0(\omega) = 0\} \end{cases}$$

Δ_{F_0} est donc le « support » de $F_0(\omega)$ et $\bar{\Delta}_{F_0}$ son complémentaire. On sait que la quantité

$$\int_{-\infty}^{\infty} \left[\frac{s_A s_B}{s_A + s_B} \right](\omega) d\omega$$

mesure l'erreur que l'on commet lorsque l'on estime linéairement $A(t)$ à partir de l'observation du processus \mathbf{Z} (donc de toutes les v.a $Z(t), t \in \mathbb{R}$). Il s'ensuit que le surplus d'erreur amené par l'échantillonnage (après le filtrage optimal) est égal à

$$\int_{\bar{\Delta}_{F_0}} \left[\frac{s_A^2}{s_A + s_B} \right](\omega) d\omega.$$

7) Le problème posé dans cet article consiste à déterminer la meilleure suite $\mathbf{U}' = \{U(n), n \in \mathbb{Z}\}$, celle qui permet d'obtenir le meilleur estimateur $\tilde{A}(t)$ de $A(t)$ de la forme (approximative)

$$\tilde{A}(t) = \sum_{k \in \mathbb{Z}} a_k(t) U(k).$$

Comme décrit par exemple dans [6], [7], il est possible de trouver de meilleurs estimateurs dans le cas où l'on admet un traite-

ment différent des parties réelle et imaginaire de $U(n)$ (dans la mesure où \mathbf{A} est un processus complexe ou le devient par filtrage).

8) Le problème de l'estimation de $A(t)$ à partir de l'échantillonnage du processus \mathbf{Z} défini par

$$Z(t) = \mathcal{F}[A](t) + B(t)$$

où \mathcal{F} est un filtre linéaire invariant, n'est pas très différent. Ce cas est étudié sous une forme particulière dans le paragraphe 4.

3. exemples

3.1. exemple 1

Les processus \mathbf{A} et \mathbf{B} sont caractérisés par les spectres

$$\begin{cases} s_A(\omega) = 1 - \alpha, \omega \in (-\pi, \pi), \\ s_A(\omega) = \alpha, \omega \in (-2\pi, -\pi) \cup (\pi, 2\pi) \\ s_B(\omega) = \beta, \omega \in (-2\pi, 2\pi) \\ s_A(\omega) = s_B(\omega) = 0, \omega \notin (-2\pi, 2\pi) \end{cases}$$

La valeur de $s_B(\omega)$ n'est, de toute façon, pas prise en compte en dehors de $(-2\pi, 2\pi)$, dans le cas de l'échantillonnage optimal. En utilisant (6), on voit que le meilleur filtre correspond à une suppression des lobes latéraux (pour $\alpha < 1/2$), et du lobe central (pour $\alpha > 1/2$), et cela, quelle que soit la valeur de β . Par exemple, si $\alpha < 1/2, l(\omega) = 0$ pour tout $\omega \in (-\pi, \pi)$ et :

$$\begin{cases} F_0(\omega) = 1, \omega \in (-\pi, \pi), \\ F_0(\omega) = 0, \omega \notin (-\pi, \pi), f_0(t) = \frac{\sin \pi t}{\pi t} \\ \varepsilon_{F_0}^2 = 2\pi \frac{\alpha(1-\alpha) + \beta}{1-\alpha + \beta} \\ \tilde{A}_{F_0}(t) = \frac{1-\alpha}{1-\alpha + \beta} \sum_{k \in \mathbb{Z}} \frac{\sin \pi(t-k)}{\pi(t-k)} U(k) \end{cases}$$

Le minimum ε_{F_0} doit être comparé avec ε_1 ($F \equiv 1$ correspond à l'échantillonnage idéal $U(n) = Z(n)$). On obtient

$$\begin{cases} \sigma_{t,1}^2 = \frac{4\pi}{1+2\beta} [\beta + 2\alpha(1-\alpha) \sin^2 \pi t] \\ \varepsilon_1^2 = \frac{4\pi}{1+2\beta} [\beta + \alpha(1-\alpha)] \\ \tilde{A}_1(t) = \frac{1}{1+2\beta} \\ \sum_{k \in \mathbb{Z}} \left[(1-2\alpha) \frac{\sin \pi(t-k)}{\pi(t-k)} + 2\alpha \frac{\sin 2\pi(t-k)}{2\pi(t-k)} \right] Z(k) \end{cases}$$

En conséquence, pour $\alpha < 1/2$, le gain est donné par

$$\frac{\varepsilon_1}{\varepsilon_{F_0}} = \sqrt{1 + \frac{1 - \alpha}{1 + 2\beta}}$$

quantité comprise entre 1 et $\sqrt{2}$. La valeur maximale est approchée pour les petites valeurs de α et β .

L'échantillonnage idéal $F \equiv 1$ n'est pas vraiment dans le cadre de cet article car $f(t)$ n'existe pas comme fonction. Ce cas peut être approché par exemple par

$$\begin{cases} g_\delta(t) = \frac{1}{2\delta} \text{ for } t \in (-\delta, \delta) \text{ et } 0 \text{ ailleurs} \\ G_\delta(\omega) = \frac{\sin \delta \omega}{\delta \omega} \end{cases}$$

Cette famille d'échantillonnages donne les résultats suivants (pour $0 < \alpha < \frac{1}{2}$ avec la convention $\text{sinc } x = \frac{\sin x}{x}$):

$$\begin{cases} \varepsilon_{G_\delta} = 2[\alpha(1 - \alpha) + \beta]h(\delta) \\ h(\delta) = \int_0^\pi \frac{\text{sinc}^2(\delta\omega) + \text{sinc}^2(\delta(\omega + 2\pi))}{(1 - \alpha + \beta)\text{sinc}^2(\delta\omega) + (\alpha + \beta)\text{sinc}^2(\delta(\omega + 2\pi))} d\omega. \end{cases}$$

On trouve évidemment $\lim_{\delta \rightarrow 0} \varepsilon_{G_\delta} = \varepsilon_1$ (cette sorte d'échantillonnage est étudiée dans [4] et [8]). La figure 2 représente les diverses erreurs en fonction de β , pour $\alpha = 1/4$.

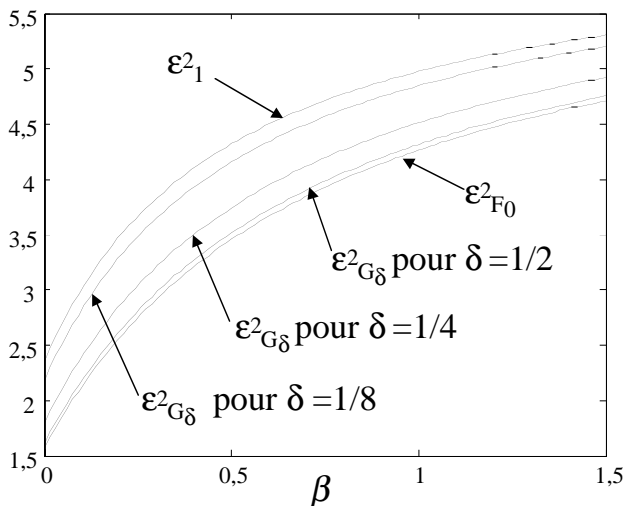


Figure 2. – Différentes erreurs dans le cas de l'exemple 1 du paragraphe 3.1.

3.2. exemple 2

Les processus **A** et **B** sont caractérisés par les spectres

$$\begin{cases} s_A(\omega) = 1, \omega \in (-\pi/2, \pi/2) \cup (\pi, 3\pi/2 + \alpha) \\ \qquad \qquad \qquad \cup (-\alpha - 3\pi/2, -\pi) \\ s_B(\omega) = \beta, \omega \in (-2\pi, 2\pi) \end{cases}$$

$s_A(\omega)$ et $s_B(\omega)$ sont nuls en dehors de ces intervalles et $0 < \alpha < \pi/2$. Dans ce cas, plusieurs solutions pour $l(\omega)$ sont acceptables, par exemple:

$$\begin{cases} l(\omega) = 0, \omega \in (-\pi/2, \pi/2) \\ l(\omega) = 1, \omega \in (-\pi, -\pi/2) \\ l(\omega) = -1, \omega \in (\pi/2, \pi). \end{cases}$$

Dans cette situation, on obtient:

$$\begin{cases} F_0(\omega) = 1, \omega \in (-\pi/2, \pi/2) \cup (\pi, 3\pi/2) \\ \qquad \qquad \qquad \cup (-3\pi/2, -\pi) \text{ et } 0 \text{ ailleurs} \\ f_0(t) = \frac{1}{\pi t} \left(\sin \frac{\pi t}{2} + \sin \frac{3\pi t}{2} - \sin \pi t \right) \\ \varepsilon_{F_0}^2 = 2\alpha + 2\pi \frac{\beta}{1 + \beta} \\ \tilde{A}_{F_0}(t) = \frac{1}{1 + \beta} \sum_{k \in \mathbb{Z}} \frac{1}{\pi(t - k)} \\ \left(\sin \frac{\pi(t - k)}{2} + \sin \frac{3\pi(t - k)}{2} - \sin \pi(t - k) \right) U(k). \end{cases}$$

Le gain est déterminé par la valeur de ε_1 qui se calcule aisément (voir l'appendice 1). Pour α petit, on obtient un rapport $\varepsilon_1/\varepsilon_{F_0}$

près de $\sqrt{1 + \frac{1}{1 + 2\beta}}$.

4. le cas des retards aléatoires

4.1. le modèle

Supposons que le processus **Z** est de la forme

$$Z(t) = A(t - C(t)) + B(t)$$

où $C = \{C(t), t \in \mathbb{R}\}$ représente par exemple un retard de transmission indépendant du processus à reconstruire **A** et du bruit **B**.

\mathbf{C} est caractérisé par les fonctions caractéristiques (ici indépendantes de t)

$$\psi(\omega) = E[e^{i\omega C(t)}], \quad \phi_\tau(\omega) = E[e^{i\omega(C(t)-C(t-\tau))}]$$

\mathbf{C} est donc stationnaire en un sens un peu plus fort que la stationnarité au sens large habituel (si l'on suppose l'existence des deux premiers moments). Dans ce cas, \mathbf{Z} se décompose sous la forme suivante [11]

$$\mathbf{Z}(t) = \psi[\mathbf{A}](t) + D(t)$$

où $\psi[\mathbf{A}](t)$ est le résultat du filtrage de \mathbf{A} par le filtre de réponse fréquentielle $\psi(\omega)$. $\psi[\mathbf{A}]$ et \mathbf{D} sont non corrélés et (lorsque $\phi_\tau(\omega) \xrightarrow{\tau \rightarrow \infty} |\psi(\omega)|^2$ suffisamment rapidement)

$$\begin{cases} s_D(\omega) = s_B(\omega) + \frac{1}{2\pi} \int_{-\infty}^{\infty} \eta(\tau) e^{-i\omega\tau} d\tau \\ \eta(\tau) = \int_{-\infty}^{\infty} [\phi_\tau(\omega) - |\psi(\omega)|^2] s_A(\omega) e^{i\omega\tau} d\omega. \end{cases}$$

Le retard \mathbf{C} induit donc à la fois un filtrage de \mathbf{A} , et un bruit à spectre d'autant plus étalé que les variations de $C(t)$ sont plus rapides. $\eta(\tau)$ caractérise ce bruit. Bien sûr, la puissance totale cumulée des deux composantes est exactement celle de \mathbf{A} .

4.2 le préfiltrage et la reconstruction

Avant l'échantillonnage, \mathbf{A} est déjà filtré, et le retard de propagation \mathbf{C} a ajouté un bruit dont le spectre est lié à celui de \mathbf{A} . Le théorème du § 2.1 se modifie sans difficulté en redéfinissant $l(\omega)$ par le système d'inégalités

$$\begin{aligned} & \left[\frac{|\psi|^2 s_A^2}{|\psi|^2 s_A + s_D} \right] (\omega + 2\pi l(\omega)) \\ & \geq \left[\frac{|\psi|^2 s_A^2}{|\psi|^2 s_A + s_D} \right] (\omega + 2\pi k), k \in \mathbb{Z}, \omega \in (-\pi, \pi). \end{aligned}$$

Le meilleur préfiltrage est encore défini par (7), et les formules (8) deviennent

$$\begin{cases} \varepsilon_{F_0}^2 = \int_{-\infty}^{\infty} s_A(\omega) d\omega - \int_{-\pi}^{\pi} \left[\frac{|\psi|^2 s_A^2}{|\psi|^2 s_A + s_D} \right] (\omega + 2l(\omega)\pi) d\omega \\ \tilde{A}_{F_0}(t) = \sum_{k \in \mathbb{Z}} a_0(t-k) U(k) \\ a_0(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\psi^* s_A}{|\psi|^2 s_A + s_D} \right] (\omega + 2l(\omega)\pi) e^{it(\omega+2l(\omega)\pi)} d\omega. \end{cases}$$

4.3 exemple

Supposons que le retard \mathbf{C} est caractérisé par

$$\begin{cases} \psi(\omega) = \cos a\omega \\ \phi_\tau(\omega) = \cos^2 a\omega + e^{-\lambda|\tau|} \sin^2 a\omega. \end{cases}$$

Un tel modèle pourrait correspondre à un transit à deux voies prises alternativement avec des probabilités égales, suivant un rythme poissonien représenté par le paramètre λ . Plaçons-nous dans le cas où (cf. figure 3)

$$s_B = 0, a = 0.1, s_A(\omega) = \frac{6 \sin^4(2\omega/3)}{\pi \omega^2}.$$

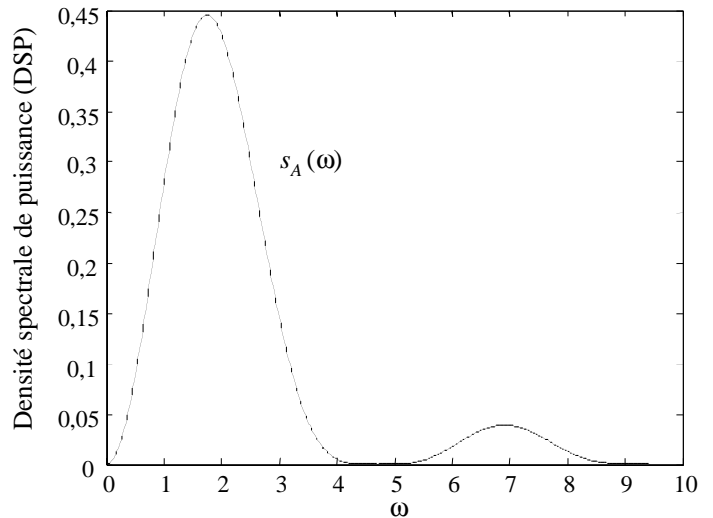


Figure 3. – Spectre du signal biphase.

Le retard relatif entre les deux chemins est donc égal à 0.2 (unités de temps). \mathbf{A} est (par exemple) un « biphase » (ou code Manchester) de puissance unité [9], et $(-3\pi/2, 3\pi/2)$ est le support du lobe principal où est concentrée 85 % de la puissance de \mathbf{A} , tandis que les lobes ± 1 en contiennent au total 7 %. L'intervalle $(-\pi, \pi)$ participe pour 82 %. La figure 3 représente $s_A(\omega)$. Comme $s_A(0) = 0$ et $s_A(\pm 2\pi) \neq 0$, il est clair que $l(0) = \pm 1$, et que ce résultat persiste au voisinage de 0. Ailleurs, $l(\omega) = 0$, lorsque λ n'est pas trop petit. Dans cette situation, on aura un (petit) avantage à utiliser les lobes ± 1 en plus du lobe principal. Pour donner un ordre d'idées, on peut dresser le tableau suivant, qui donne ε_{F_0} (préfiltrage optimal), ε_1 (échantillonnage « idéal »), ε_{F_1} (préfiltrage passe-bande $(-\pi, \pi)$), comme fonction de λ (paramètre qui règle le rythme des changements de voie de transmission)

λ	0.03	0.1	0.3	1	3	10
ε_{F_0}	0.223	0.214	0.212	0.207	0.203	0.199
ε_{F_1}	0.227	0.218	0.216	0.212	0.208	0.204
ε_1	0.302	0.293	0.292	0.291	0.287	0.280

On a bien l'inégalité $\varepsilon_{F_0} < \inf[\varepsilon_1, \varepsilon_{F_1}]$ comme prévu. L'amélioration par rapport à ε_{F_1} est faible (de l'ordre de 2 %) et quasiment indépendante de λ ; rappelons qu'elle consiste à récupérer de l'information en dehors de $(-\pi, \pi)$ où la puissance utile est inférieure à 20 % (7 % dans les lobes ± 1), et qu'il n'est possible d'en utiliser qu'une petite partie. Par contre, l'amélioration par rapport à l'échantillonnage idéal est importante (autour de 40 %) dans les deux cas (suppression des bandes latérales et filtrage optimal). On pourra augmenter le gain en sous-échantillonnant le processus. La figure 4 représente le spectre du signal résultant $s_{\psi[A]}(\omega)$ et celui du bruit $s_D(\omega)$ pour diverses valeurs de λ .

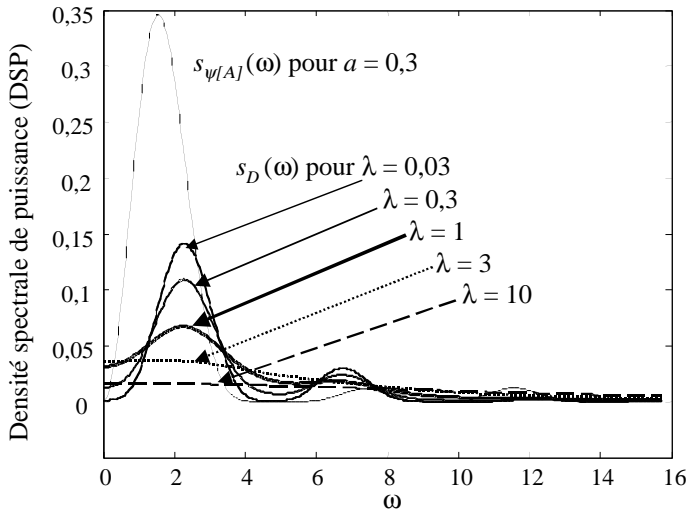


Figure 4. – Spectres du signal et du bruit après retard de propagation, pour différentes valeurs de λ .

5. conclusion

Les dispositifs physiques de prise d'échantillons sur des processus à temps continu $\mathbf{Z} = \{Z(t), t \in \mathbb{R}\}$ ne sont pas idéaux. La plupart du temps, ils prennent en compte un ensemble de valeurs du processus étudié autour d'instantanés périodiques. Par exemple, les échantillons $U(n)$ sont le résultat d'un moyennage autour des entiers [4], [8]. Plus généralement, ce peut être une opération linéaire invariante, autour de ces mêmes instants. Dans ce cas, on peut considérer que les échantillons $U(n)$ proviennent de l'échantillonnage idéal d'un processus $\mathbf{U} = \{U(t), t \in \mathbb{R}\}$ qui est la sortie d'un filtre linéaire invariant attaqué par le processus \mathbf{Z} .

Quand le processus étudié \mathbf{Z} est sujet au phénomène de repliement, l'erreur de reconstruction $\sigma_{t,F}$ est fonction du filtre linéaire qui définit les échantillons. Dans cet article, on cherche le filtre qui minimise l'erreur moyenne ε_F (c'est-à-dire la racine carrée de l'intégrale de $\sigma_{t,F}^2$ sur $(0, 1)$). Il apparaît que la solution est caractérisée par les entiers $l(\omega)$ tels que le rapport $[s_A^2 / (s_A + s_B)](\omega + 2\pi l(\omega))$ est maximum, et le procédé qui en résulte est une forme adaptée de suppression de repliement. Enfin, comme le procédé est une suppression de repliement, l'erreur σ_{t,F_0} est indépendante de t (et donc égale à sa moyenne ε_{F_0}).

6. appendices

6.1. appendice 1 : la meilleure estimation linéaire

La technique utilisée ci-dessous, décrite dans [1], est rappelée notamment dans [9] et [12]. Au sens de l'erreur en moyenne quadratique, le meilleur estimateur linéaire $\tilde{A}(t)$ de $A(t)$ à partir de l'observation de la suite $\mathbf{U}' = \{U(n), n \in \mathbb{Z}\}$ est défini par

$$\tilde{A}(t) = pr_{\mathbf{H}(\mathbf{U}')} A(t)$$

où $\mathbf{H}(\mathbf{U}')$ est l'espace de Hilbert engendré par \mathbf{U}' et $pr_{\mathbf{H}(\mathbf{U}'')}$ la projection orthogonale. $\tilde{A}(t)$ est parfaitement caractérisé par le système d'équations

$$\begin{cases} E[(A(t) - \tilde{A}(t)) U^*(n)] = 0, n \in \mathbb{Z} \\ \tilde{A}(t) \in \mathbf{H}(\mathbf{U}') \end{cases} \quad (9)$$

où $U(t)$ est défini par (2) :

$$U(t) = \int_{-\infty}^{\infty} f(u) Z(t-u) du, t \in \mathbb{R}.$$

De manière équivalente, en utilisant l'isométrie fondamentale $I_{\mathbf{U}'}$ qui relie $\mathbf{H}(\mathbf{U}')$ et $\mathbf{L}^2(s'_{\mathbf{U}'})$ [1], [9], [12], (9) s'écrit :

$$\begin{aligned} \int_{-\infty}^{\infty} f^*(u) E[A(t) A^*(n-u)] du \\ - \int_{-\pi}^{\pi} [\phi_t s'_{\mathbf{U}'}](\omega) e^{-i\omega n} d\omega = 0, n \in \mathbb{Z} \end{aligned} \quad (10)$$

où

– $s'_{\mathbf{U}'}(\omega)$ est la densité spectrale de la suite \mathbf{U}' définie ici sur $(-\pi, \pi)$ [2] :

– $\phi_t(\omega)$ est la fonction qui correspond à $\tilde{A}(t)$ dans l'isométrie $I_{\mathbf{U}'}$.

Le fait que les processus **A** et **B** sont non corrélés est pris en compte dans la première intégrale.

En utilisant (1) et (3), on obtient

$$\int_{-\infty}^{\infty} f^*(u) E[A(t) A^*(n-u)] du = \int_{-\infty}^{\infty} [F^* s_A](\omega) e^{i\omega(t-n)} d\omega. \quad (11)$$

s'_U se calcule à partir de l'utilisation des formules de Wiener-Lee [9], [14]. Et l'unicité des séries de Fourier permet d'écrire à partir de (10) et (11)

$$\begin{cases} \phi_t(\omega) = \left[\frac{\mu_t}{s'_U} \right](\omega) e^{i\omega t}, \omega \in (-\pi, \pi) \\ \mu_t(\omega) = \sum_{k \in \mathbb{Z}} [F^* s_A](\omega + 2k\pi) e^{2i\pi k t} \\ s'_U(\omega) = \sum_{k \in \mathbb{Z}} [|F|^2 (s_A + s_B)](\omega + 2k\pi). \end{cases} \quad (12)$$

La paire de transformées de Fourier $f(t)$ et $F(\omega)$ est définie par (3). $f(t)$ est la réponse impulsionnelle et $F(\omega)$ le gain en fréquence du filtre qui transforme **Z** en **U**. Un développement en série de Fourier de $\phi_t(\omega)$ sur $(-\pi, \pi)$ donne $\tilde{A}(t)$, en remplaçant $e^{in\omega}$ par $U(n)$. L'erreur quadratique moyenne $\sigma_{t,F}^2$ correspondante est calculée à partir du théorème de Pythagore :

$$\begin{aligned} \sigma_{t,F}^2 &= E[|A(t) - \tilde{A}(t)|^2] \\ &= \int_{-\infty}^{\infty} s_A(\omega) d\omega - \int_{-\pi}^{\pi} \left[\frac{|\mu_t|^2}{s'_U} \right](\omega) d\omega. \end{aligned} \quad (13)$$

D'autre part, le théorème de Plancherel des séries de Fourier permet d'écrire l'égalité :

$$\int_0^1 [|\mu_t|^2](\omega) dt = \sum_{k \in \mathbb{Z}} [|F|^2 s_A^2](\omega + 2k\pi).$$

L'erreur moyenne ε_F^2 définie par (5) est obtenue en intégrant (13) sur $(0, 1)$. En conséquence :

$$\varepsilon_F^2 = \int_{-\pi}^{\pi} \left[a - \frac{c}{d} \right](\omega) d\omega \quad (14)$$

où les fonctions a, c, d sont définies par

$$\begin{cases} a(\omega) = \sum_{k \in \mathbb{Z}} s_A(\omega + 2k\pi) \\ c(\omega) = \sum_{k \in \mathbb{Z}} [|F|^2 s_A^2](\omega + 2k\pi) \\ d(\omega) = \sum_{k \in \mathbb{Z}} [|F|^2 (s_A + s_B)](\omega + 2k\pi). \end{cases} \quad (15)$$

6.2. appendice 2 : le filtrage optimal

1) Il est facile de construire une fonction mesurable $l(\omega)$ qui vérifie (6) [10]. Alors, on a évidemment l'égalité

$$\int_{-\infty}^{\infty} F_0(\omega) d\omega = 2\pi.$$

2) Pour démontrer (7), on utilise l'expression (14) :

$$\varepsilon_F^2 = \int_{-\pi}^{\pi} \left[a - \frac{c}{d} \right](\omega) d\omega$$

où $a(\omega), c(\omega), d(\omega)$, sont définis en (15). L'inégalité de Schwarz implique que $\left[a - \frac{c}{d} \right](\omega) \geq 0$ pour tout ω . Le minimum de ε_F sera donc obtenu quand, pour chaque $\omega \in (-\pi, \pi)$, la quantité $\left[\frac{c}{d} \right](\omega)$ est maximale. On utilise pour ce faire l'inégalité élémentaire

$$\frac{g_1 + g_2 + \dots + g_n}{h_1 + h_2 + \dots + h_n} \leq \sup_{k=1,2,\dots,n} \frac{g_k}{h_k}$$

valable lorsque tous les éléments pris en compte sont positifs ou nuls (en éliminant les éventuelles indéterminations). Alors :

$$\begin{cases} \left[\frac{c}{d} \right](\omega) \leq \sup_n \frac{\lambda_n}{\mu_n} \\ \lambda_k = [s_A^2](\omega + 2\pi k), \mu_k = [s_A + s_B](\omega + 2\pi k) \end{cases} \quad (16)$$

puisque

- les paramètres λ_k, μ_k sont positifs (ou nuls)
- $F(\omega)$ est bornée (par hypothèse)
- les séries de termes généraux λ_k et μ_k sont convergentes (presque partout).

Le maximum de $\left[\frac{c}{d} \right](\omega)$ sera donc obtenu lorsque l'on supprime dans $c(\omega)$ et $d(\omega)$ les termes d'indices autres que $l(\omega)$.

3) En utilisant (7), on déduit que les fonctions $\mu_t(\omega)$ et $s'_U(\omega)$ définies par (12) sont réduites à un seul terme (pour $F = F_0$). Plus précisément, on obtient $\phi_t(\omega)$ qui caractérise $\tilde{A}_{F_0}(t)$ sous la forme

$$\phi_t(\omega) = \left[\frac{s_A}{s_A + s_B} \right](\omega + 2\pi l(\omega)) e^{2i\pi l(\omega)t + i\omega t}, \omega \in (-\pi, \pi).$$

Puisque $\phi_t(\omega) = \phi_{t+n}(\omega) e^{-i\omega n}, n \in \mathbb{Z}$, la série de Fourier de $\phi_t(\omega)$ sur $(-\pi, \pi)$ prend la forme simple

$$\phi_t(\omega) \sim \sum_{k \in \mathbb{Z}} \beta_0(t-k) e^{ik\omega}, \beta_0(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_t(\omega) d\omega$$

ce qui explique la deuxième expression dans (8).

4) Finalement, puisque $\mu_t(\omega)$ est réduit à un seul terme, $|\mu_t(\omega)|^2$ est une quantité indépendante de t , et la même pro-

priété est vraie pour σ_{t, F_0} . Bien entendu, il existera des valeurs de t pour lesquelles $\sigma_{t, F_0} > \sigma_{t, 1}$ ($F = 1$ correspond à l'échantillonnage idéal $U_n = Z(n)$). Évidemment, $\sigma_{n, 1} = 0, n \in \mathbb{Z}$.

BIBLIOGRAPHIE

- [1] H. Cramer, M. R. Leadbetter, *Stationary and related stochastic processes*, Wiley, New-York (1966).
- [2] J. L. Doob, *Stochastic Processes*, Wiley, New-York (1953).
- [3] J. R. Higgins, *Sampling Theory in Fourier and Signal Analysis*, Oxford Sc. Pub. (1996).
- [4] T. Hurt, *Spectral effects of sampling a continuous-scan Fourier transform spectrometer with a capacitive transimpedance amplifier*, Opt. Eng. 40 (7) (2001) 1415-1417.
- [5] A. Jerri, *The Shannon sampling theorem. Its various extensions and applications. A tutorial review*, Proc. IEEE 65 (11) (1977) 1565-1596.
- [6] B. Lacaze, *Periodic bi-sampling of stationary processes*, Signal Processing, 68 (1998) 283-293.
- [7] B. Lacaze, *Improving the Nyquist rate for complex stationary process sampling*, Samp. Th. in Sig. and Im. Proc. Vol.1, No.1 (2001) 33-51.
- [8] B. Lacaze, C. Mailhes, *Influence of a random integration width on periodic sampling*, ICASSP, Orlando, (2002) II-1189-1192.
- [9] B. Lacaze, *Processus aléatoires pour communications numériques*, Hermes, (2000).
- [10] B. Lacaze, *Matched shapes for uniform sampling (soumis à Stat. and Prob. Lett.)*
- [11] B. Lacaze, *Stationary clock changes on stationary processes*, Signal Processing, 55 (1996) 191-205.
- [12] B. Lacaze, *La formule d'échantillonnage et A. L. Cauchy*, Traitement du Signal, Vol. 15, n° 4, (1998) 289-295.
- [13] S. P. Lloyd, *A sampling theorem for stationary (wide sense) stochastic processes*, Trans. Am. Math. Soc. Vol. 92 (1959) 1-12.
- [14] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, Mac-Graw Hill (1991).
- [15] A. M. Yaglom, *An Introduction to the Theory of Stationary Random Functions*, Dover (1962).

Manuscrit reçu le 13 mars 2003