

Influence d'un échantillonnage irrégulier sur les performances de la reconstruction

Wilfried CHAUVET¹, Marie CHABERT², Bernard LACAZE¹

¹TéSA 14-16 Port Saint Etienne 31000 Toulouse, France

²Université de Toulouse, Institut de Recherche en Informatique de Toulouse (IRIT)

INP-ENSEEIH, 2 rue Charles Camichel, BP 7122, 31071 Toulouse, France

chauvet@tesa.prd.fr, chabert@enseeiht.fr, lacaze@tesa.prd.fr

Résumé – Cet article étudie le problème de la reconstruction d'un processus aléatoire à spectre borné dans le cas d'un échantillonnage périodique et d'un échantillonnage irrégulier. Diverses méthodes (interpolation par splines, méthode dite matricielle, algorithmes itératifs et formule de Lagrange modifiée) sont envisagées lorsque l'échantillonnage est irrégulier. Dans le cas périodique, la reconstruction utilise la formule de Shannon. Les deux schémas d'échantillonnage sont comparés en terme d'erreur de reconstruction. Lorsque la condition de Shannon est vérifiée, la formule de Lagrange modifiée en présence de gigue donne des résultats similaires à la formule de Shannon pour un échantillonnage régulier. Dans le cas contraire, l'échantillonnage irrégulier permet d'étaler les interférences hors bande utile. En particulier, l'influence d'une composante spectrale en dehors de la bande de Shannon est étudiée : le repliement obtenu avec un échantillonnage irrégulier limite le risque de mauvaise interprétation du contenu spectral.

Abstract – This paper studies the problem of band-limited random signal reconstruction in the case of periodic and irregular sampling. Various methods (spline interpolation, matrix method, iterative algorithms, modified Lagrange formula) are considered in the case of irregular sampling. The reconstruction is performed using the Shannon formula in the case of periodic sampling. Periodic and random sampling are compared in terms of reconstruction error. When the Shannon condition holds, the modified Lagrange formula applied to jitter shows similar performance to the Shannon formula applied to regular sampling. Otherwise, an irregular sampling allows to spread the out-of-band interference. The influence of a spectral component outside the Nyquist band is studied: in this case, the periodizing sampling may occur misleading interpretation of the spectral content.

1 Introduction

Considérons un signal analogique modélisé par un processus aléatoire stationnaire à bande limitée $\mathbf{Z} = \{Z(t), t \in \mathbb{R}\}$, de densité spectrale $s(\omega)$ régulière définie par [4] :

$$K(\tau) = E[Z(t)Z^*(t-\tau)] = \int_{-\pi}^{\pi} e^{i\omega\tau} s(\omega) d\omega.$$

Nous travaillons donc sur un ensemble de processus à puissance finie tels que ceux que l'on peut rencontrer en télécommunications [7] et non sur des fonctions déterministes à énergie finie plus couramment utilisées en traitement d'images par exemple. Une conséquence importante est que les échantillons, même lointains, pris en compte dans la reconstruction ne peuvent pas être considérés comme négligeables. Le schéma d'échantillonnage irrégulier étudié correspond à un échantillonnage périodique perturbé par de la gigue. L'annulation de la gigue conduit donc à l'échantillonnage périodique. La séquence des instants d'échantillonnage est définie par $\{t_n = n\theta + G_n, n \in \mathbb{Z}\}$. La suite $\{G_n, n \in \mathbb{Z}\}$ modélise la gigue supposée bornée et observable, de distribution uniforme sur un intervalle centré, la largeur de l'intervalle réglant la puissance de la gigue. Dans nos simulations, cet intervalle est $[-\theta/2, \theta/2]$. Ce choix est cou-

rant dans les applications d'analyse spectrale de signaux échantillonnés puisqu'il permet de supprimer le repliement lorsque Shannon n'est pas vérifié [2]. La gigue intervient dans deux types de situation : elle peut modéliser les imperfections d'un procédé d'échantillonnage idéalement périodique ou être intentionnelle. Un domaine d'application émergent est l'échantillonnage compressif qui permet la conversion analogique numérique de signaux large bande à représentation spectrale parcimonieuse [3], [5].

Le but de la reconstruction est d'estimer les $Z(t)$ lorsque $t \neq t_n$, à partir de la suite d'échantillons observés $\{Z(t_n), n \in \mathbb{Z}\}$. On notera $\hat{Z}(t)$ le signal reconstruit et :

$$\varepsilon = \frac{E[(Z(t) - \hat{Z}(t))^2]}{E[(Z(t))^2]} \quad (1)$$

l'erreur quadratique de reconstruction normalisée par rapport à la puissance du signal. La période d'échantillonnage moyenne notée θ est telle que $0 < \theta < 1$ ce qui correspond à un scénario de sur-échantillonnage. Dès lors que la période moyenne d'échantillonnage est inférieure à 1, la suite d'échantillons observés contient une information suffisante pour permettre la reconstruction [10].

Dans un premier temps, nous présentons différentes méthodes

de reconstruction adaptées au cas d'un échantillonnage non périodique : l'interpolation par splines, une méthode dite matricielle qui correspond à la résolution d'un système d'équations linéaires, des méthodes itératives et une formule de Lagrange modifiée. Nous comparons ensuite les performances de ces méthodes dans le cas d'un échantillonnage aléatoire à la formule de Shannon dans le cas d'un échantillonnage périodique. Dans un premier temps, la condition de Shannon est supposée vérifiée. Ceci implique une atténuation suffisante du spectre en dehors de la bande de Nyquist. Nous avons également mesuré l'influence du niveau d'atténuation exigé sur les performances comparées des méthodes. Afin d'illustrer les propriétés d'étalement de l'échantillonnage non périodique, nous considérons ensuite un signal qui présente une raie à l'extérieur de la bande considérée comme utile et qui constitue un exemple de non respect de la condition de Shannon. Les simulations permettent d'étudier l'influence de cette raie sur la reconstruction du signal pour un échantillonnage périodique et non périodique.

2 Reconstruction en présence de gigue d'échantillonnage

Nous avons choisi de considérer quatre types de méthodes de reconstruction dans le cas d'un échantillonnage irrégulier. La reconstruction par splines est parmi les plus utilisées. La méthode dite matricielle est la méthode qui découle naturellement de la formule de Shannon en présence d'échantillonnage irrégulier. Les algorithmes itératifs de reconstruction utilisent des connaissances a priori sur l'occupation spectrale du processus. Enfin, la formule de Lagrange modifiée permet une reconstruction exacte lorsque la fenêtre d'observation n'est pas limitée. Ces méthodes seront bien sûr comparées dans le cas réaliste où la fenêtre d'observation est limitée.

2.1 Interpolation par splines

L'interpolation par splines fournit une approximation polynomiale du signal sur l'intervalle $[t_n, t_{n+1}]$ [1]. Le choix courant d'un polynôme de degré 3 conduit aux splines cubiques. La fonction d'approximation doit être deux fois continument dérivables sur toute la fenêtre d'observation. Le calcul des coefficients des polynômes sur chaque intervalle $[t_n, t_{n+1}]$ s'obtient en imposant, aux instants d'échantillonnage t_n , l'égalité entre le signal observé et le signal reconstruit ainsi que la continuité des dérivées premières. La nullité de la dérivée seconde aux bornes de la fenêtre d'observation est une condition supplémentaire pour l'unicité de la solution.

2.2 Algorithmes itératifs

L'idée générale est de reconstruire le signal sous une contrainte relative à son occupation spectrale. Une méthode de reconstruction est appliquée lors d'une première étape. Le résultat qui ne possède pas nécessairement la bande spécifiée est ensuite fil-

tré par un filtre passe-bas de bande celle de Nyquist. Ceci peut être interprété comme une projection du signal reconstruit. On réitère ensuite cette opération sur l'erreur de projection i.e. la différence entre le signal reconstruit et sa projection. La procédure itérative cesse lors de la validation d'un certain critère de convergence.

2.3 Méthode matricielle

La méthode de reconstruction la plus intuitive est la méthode dite matricielle. Elle consiste à construire, à partir de $2N + 1$ échantillons observés $Z(t_j)$, $-N \leq j \leq N$, un système linéaire à $2N + 1$ inconnues $\tilde{Z}(n\theta)$, $-N \leq n \leq N$, correspondant aux estimations des $Z(n\theta)$, $-N \leq n \leq N$. Ces estimations sont liées aux observations par la formule de Shannon tronquée selon :

$$Z(t_j) = \sum_{n=-N}^N \text{sin}_c \pi \left(\frac{t_j}{\theta} - n \right) \tilde{Z}(n\theta), \quad -N \leq j \leq N$$

où $\text{sin}_c x = \frac{\sin x}{x}$. Une inversion matricielle permet de calculer les $\tilde{Z}(n\theta)$. La reconstruction $\tilde{Z}(t)$ pour une valeur de t quelconque s'obtient en appliquant la même formule :

$$\tilde{Z}(t) = \sum_{n=-N}^N \text{sin}_c \pi \left(\frac{t}{\theta} - n \right) \tilde{Z}(n\theta).$$

Cette méthode possède de bonnes performances en terme d'erreur de reconstruction pour des valeurs suffisantes de N . Toutefois, l'étude théorique de la convergence est inextricable étant donné que deux types d'erreurs se cumulent : l'erreur d'estimation des $Z(n\theta)$ et l'utilisation d'une formule tronquée pour les $\tilde{Z}(t)$. Enfin, l'inversion matricielle implique une complexité calculatoire qui croît exponentiellement avec la taille de la fenêtre d'observation. Nous avons proposé une formule de reconstruction alternative utilisant directement les échantillons observés. L'obtention de cette formule est exposé dans le paragraphe suivant.

2.4 Formule d'interpolation de Lagrange modifiée

La formule considérée utilise des fonctions d'interpolation construites à partir des expressions suivantes [8] :

$$P_N(z) = \pi z \prod_{n=-N, n \neq 0}^N \left(1 - \frac{z}{n} \right) \quad (2)$$

$$Q_N(z) = \prod_{n=-N-1}^{N+1} \left(1 - \frac{z}{t_n} \right) \quad (3)$$

$$R_N(z) = \frac{Q_N(z)}{P_N(z)} \sin \pi z. \quad (4)$$

Quels que soient les instants d'échantillonnage, dès lors qu'ils ne sont pas entiers, la reconstruction exacte de $Z(t)$ s'écrit :

$$Z(t) = R_N(t) [A_N(t) + B_N(t)] \quad (5)$$

avec :

$$A_N(t) = \sum_{|n| \leq N+1} \frac{P_N(t_n)}{(t-t_n)Q'_N(t_n) \sin \pi t_n} Z(t_n) \quad (6)$$

$$B_N(t) = \sum_{|n| \geq N+1} \frac{(-1)^n P_N(n)}{\pi(t-n)Q_N(n)} Z(n). \quad (7)$$

La formule de Lagrange classique est construite à partir des échantillons mesurés. La formule de Lagrange (5) fait intervenir deux termes. Le premier terme utilise les échantillons observés $Z(t_n)$. En revanche, le second terme ne peut être calculé, car fonction des $Z(n)$, $|n| \geq N+1$, qui ne sont pas observés. Néanmoins, on démontre que ce terme converge rapidement vers 0 lorsque le nombre d'échantillons observés tend vers l'infini :

$$\lim_{N \rightarrow \infty} B_N(t) = 0. \quad (8)$$

$$R_\infty(t) = \lim_{N \rightarrow \infty} R_N(t) \quad (9)$$

étant bien définie, on en déduit

$$Z(t) = R_\infty(t) \lim_{N \rightarrow \infty} A_N(t). \quad (10)$$

Après troncature imposée par la fenêtre d'observation, la formule n'utilise que les échantillons observés $Z(t_j)$, $-N \leq j \leq N$. Les échantillons périodiques $Z(n)$ impliqués dans l'expression $B_N(t)$ disparaissent lors du passage à la limite (10). Cette suite fantôme est un artifice calculatoire, qui peut être utilisé de diverses manières [9]. Notons finalement que, sous certaines conditions très particulières et difficilement vérifiables en pratique, il est possible d'intervir la limite et la sommation. La formule de Lagrange modifiée (10) conduit à la formule de Lagrange classique [10], [12]. Dans ce cas, un exemple célèbre est donné par le théorème de Kadec [6].

3 Etude des performances

Dans les simulations qui suivent, la taille de la fenêtre d'observation ($N = 64$ échantillons pris en compte de part et d'autre de la fenêtre de reconstruction, elle-même de taille 1 et centrée sur un échantillon) est choisie suffisamment grande pour garantir une erreur raisonnable pour les cinq méthodes. Un changement du paramètre N conduit globalement à une translation des différentes courbes selon l'axe des ordonnées, le gain relatif des différentes méthodes restant approximativement inchangé. Les performances sont représentées en terme d'erreur quadratique moyenne pour un facteur de suréchantillonnage croissant. L'erreur quadratique relative de reconstruction est estimée par simulations de Monte-Carlo grâce à la moyenne des erreurs quadratiques normalisées calculées sur un échantillonnage dense de la fenêtre de reconstruction.

3.1 Condition de Shannon vérifiée

Les performances en reconstruction sont comparées dans le cas avec et sans gigue lorsque le signal est échantillonné à une

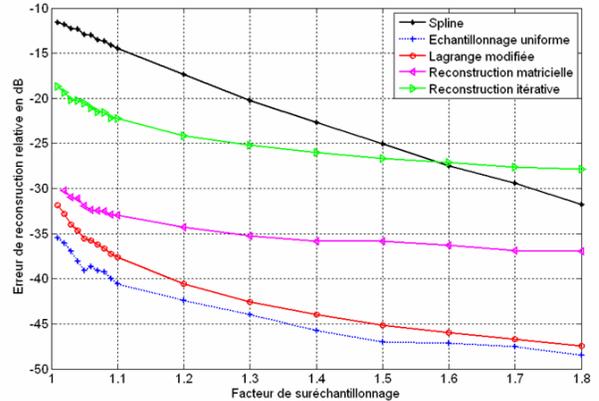


FIG. 1 – Performances en fonction du facteur de suréchantillonnage.

fréquence moyenne supérieure à la fréquence de Shannon. La formule de Lagrange modifiée en présence de gigue donne des résultats similaires à la formule de Lagrange dans le cas régulier (figure 1). Supposer la condition de Shannon vérifiée revient en pratique à supposer que l'atténuation du spectre hors bande de Nyquist n'excède pas un certain seuil. Nous avons observé que, plus ce seuil est contraignant, meilleures sont les performances de la formule de Lagrange modifiée relativement aux autres méthodes (figure 2). L'atténuation a été fixée à 25dB pour la figure 1 et à 50dB pour la figure 2.

3.2 Condition de Shannon non vérifiée

Le cas où la condition de Shannon n'est pas vérifiée peut être simulé en considérant la présence d'une raie en dehors de la bande utile du signal. Pour un échantillonnage périodique, cette raie, par repliement, introduit des raies parasites à l'intérieur de la bande utile (figure 3). Ce phénomène est dommageable dans certaines applications pratiques et peut conduire à de mauvaises interprétations. En revanche, un échantillonnage non périodique conduira également à un repliement de cette raie mais sous une forme étalée (figure 4). L'interférence intervenant comme un bruit large bande conduit à une erreur quadratique de reconstruction similaire au cas précédent mais sa forme est moins préjudiciable au diagnostic (figure 5). Ces résultats ont été obtenus pour un facteur de suréchantillonnage de 1.01 et une atténuation de 25 dB à la fréquence de Shannon.

4 Conclusion

Nous avons étudié l'influence de la gigue sur la qualité de la reconstruction, en considérant la présence éventuelle d'interférences hors bande utile. Dans le cas d'un échantillonnage non périodique respectant la condition de Shannon, la comparaison entre l'interpolation par splines, la méthode matricielle et la formule de Lagrange modifiée montre que les deux der-

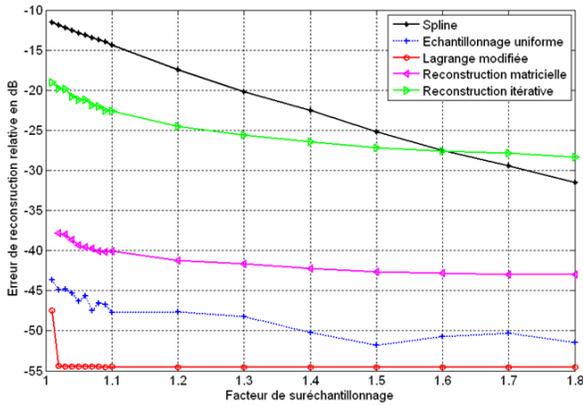


FIG. 2 – Performances en fonction du facteur de suréchantillonnage - Shannon plus strict.

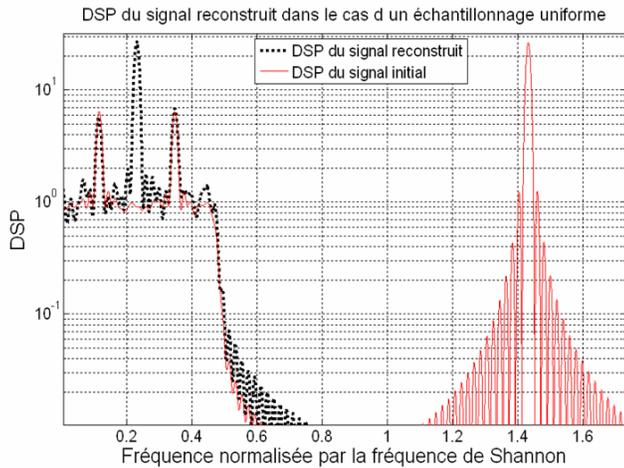


FIG. 3 – Influence d'une raie - Echantillonnage périodique.

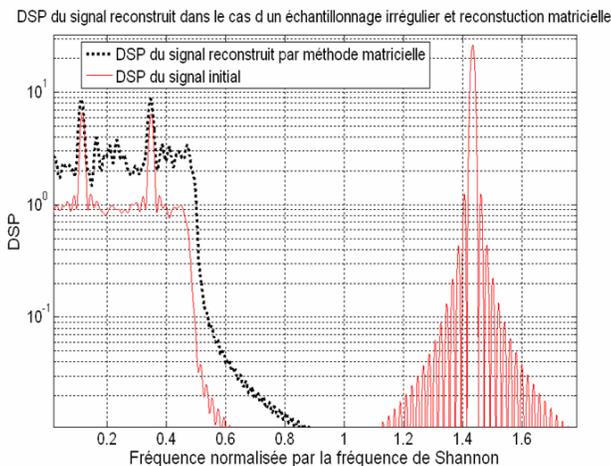


FIG. 4 – Influence d'une raie - Echantillonnage aléatoire.

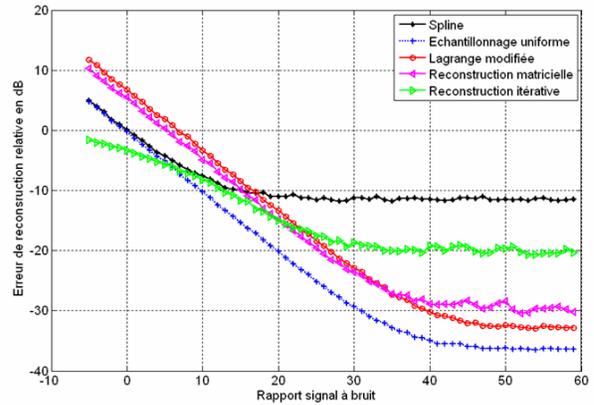


FIG. 5 – Influence d'une raie en fonction de sa puissance.

nières méthodes sont indéniablement supérieures aux splines en terme d'erreur de reconstruction. D'autre part, nous avons vérifié qu'un échantillonnage non périodique permettait d'éta-ler les interférences et ainsi de fournir une reconstruction mieux adaptée à d'éventuelles tâches de décision ultérieures. Ceci peut avoir l'avantage d'assouplir les contraintes sur le gabarit des filtres antirepliement.

Références

- [1] A. Aldroubi, H. Feichtinger, *Complete iterative reconstruction algorithms for irregularly sampled data in spline like spaces*, ICASSP-97, pp. 1857-1860.
- [2] I. Bilinskis, *Digital alias free signal processing*, Wiley, avril 2007.
- [3] E. J. Candès, M. B. Wakin, "An introduction to compressive sampling", *IEEE signal processing Magazine*, March 2008, p 21-30.
- [4] H. Cramér, M. R. Leadbetter, *Stationary and related stochastic processes*, Wiley, New-York (1966).
- [5] D. Healy, D. J. Brady, "Compression at the physical interface", *IEEE signal processing Magazine*, March 2008, p67-71.
- [6] J. R. Higgins, *Sampling Theory in Fourier and Signal Analysis*, Oxford Sc. Pub. (1996).
- [7] B. Lacaze, *Processus aléatoires pour communications numériques*, collection Traitement du Signal, Hermès, 2000.
- [8] B. Lacaze, *Reconstruction formula for irregular sampling*, *Sampling Theory in Signal and Image Processing STSIP 4 (1) (2005) 33-43*.
- [9] B. Lacaze, *The Ghost Sampling Sequence Method*, *Sampling Theory in Signal and Image Processing STSIP 8 (1) (2009)*.
- [10] B. Ja. Levin, *Zeros of Entire Functions*, Am. Math. Soc., 1964.
- [11] A. Papoulis, "A new algorithm in spectral analysis and band limited extrapolation", *IEEE Trans. Circuits Systems*, CAS 22, p735-742, 1975.
- [12] A. I. Zayed, P. L. Butzer, *Lagrange Interpolation and Sampling Theorems*, in *Nonuniform Sampling. Theory and Practice*, ed. F. Marvasti, Kluwer Ac., 2001.