

Traitement des Signaux Aléatoires

Estimation Spectrale

4 ETI – CPE Lyon

Travaux Pratiques TSA

Noms, Prénoms : Beaucamp Elisa, Deschamps Corto

Groupe : C

Date : Lundi 11 octobre

Objectifs du TP

- Comprendre la notion de densité spectrale d'énergie ou de puissance moyenne
- Manipuler différents estimateurs empiriques (à partir d'une série temporelle de taille finie) de DSE/DSPM
- Etudier l'effet du compromis biais-variance d'un estimateur

1 Préparation

Question 1 Comment peut-on calculer simplement la densité spectrale d'énergie (DSE) d'un signal certain d'énergie finie ?

_____ **réponse** _____
On a la DSE d'un signal à énergie finie :

$$\Gamma_X(\nu) = |X(\nu)|^2$$

□

Question 2 Comment est définie la densité spectrale de puissance moyenne (DSPM) d'un processus aléatoire ?

_____ **réponse** _____
D'après Wiener-Kintchin, la DSPM d'un signal aléatoire vaut :

$$\Gamma_X(\nu) = TF(\gamma_X(\tau)) = \int_{-\infty}^{\infty} \gamma_X(\tau) e^{-2i\pi\nu\tau} d\tau$$

□

Question 3 Quelles sont les grandeurs qui permettent de chiffrer la qualité d'une estimation dans le cas général ? et la qualité de l'estimation spectrale en particulier.

_____ **réponse** _____

Dans le cas général, on caractérise la qualité de l'estimation grâce au biais et à la variance.
 Dans le cas spectral, on peut se servir de la précision d'estimation :

$$\epsilon^2 = \frac{\text{Var}(\Gamma_X(f))}{E^2(\Gamma_X(f))}$$

□

Question 4 Exprimer la densité spectrale de puissance moyenne (DSPM) $G_B(f)$ d'un bruit blanc stationnaire centré.

réponse

La DSPM d'un bruit blanc vaut :

$$\Gamma_B = TF(\gamma_B(\tau)) = TF(\Gamma_0 \delta(\tau)) = \Gamma_0$$

□

Question 5 Exprimer $G_X(f)$, où $X(t)$ est la sortie d'un filtre excité par un bruit blanc centré, en fonction de la DSPM du bruit blanc et des caractéristiques du filtre.

réponse

$$\begin{aligned} \Gamma_X(f) &= \gamma_X(0) = \int_{-\infty}^{\infty} |H\nu|^2 \gamma_X(\tau) d\nu = \Gamma_0 \int_{-\infty}^{\infty} |H\nu|^2 d\nu \\ &= \Gamma_0 \int_{-\infty}^{\infty} \left| \frac{k}{\Gamma_0} TF(X(T)) \right|^2 d\nu = \Gamma_0 \int_{-\infty}^{\infty} \frac{k}{\Gamma_0} X^*(\nu) e^{-2i\pi\nu t_0} d\nu \end{aligned}$$

□

Question 6 En une phrase (sans formule), décrire le procédé de calcul de la DSPM estimée $G_1(f)$ d'une séquence aléatoire via l'estimateur simple.

réponse

Pour calculer $\Gamma_1(f)$ via l'estimateur simple, on réalise la TF N-points du signal, que l'on somme, on met au carré puis on normalise en divisant par N.

□

Question 7 Rappeler le mode de graduation d'une TFD-N points en fréquences réduites.

réponse

On rappelle la graduation d'une TFD N-points en fréquences réduites :

$$0 : \frac{1}{N} : 1 - \frac{1}{N}$$

.

□

Question 8 Décrire (avec une phrase) le procédé de calcul de la DSPM estimée $G_2(f)$ d'une séquence aléatoire via l'estimateur moyenné.

réponse

Pour calculer Γ_2 via l'estimateur moyenné, on divise le signal en L tranches puis on calcule Γ_1 sur chacune de ces tranches. On fait ensuite une moyenne de ces Γ_1 pour trouver Γ_2 .

□

Question 9 Que signifie le terme «compromis biais-variance» dans le cas de l'estimateur moyenné ?

réponse

Le compromis biais-variance fait référence au fait que lorsque l'un diminue, l'autre augmente, et vice-versa. Or, dans l'idéal, nous aimerions que les 2 grandeurs soient petites. Il faut donc doser pour avoir quelque chose de satisfaisant. Dans le cas de Γ_2 , on a pour L tranches fixées :

$$\lim_{M \rightarrow \infty} = \frac{\frac{\Gamma_{X^2}(f)}{L}}{\Gamma_X^2(f)} = \frac{1}{L}$$

□

Question 10 Quelles modifications sont apportées au procédé de calcul de l'estimateur de Welch par rapport à l'estimateur moyenné ?

_____ **réponse** _____

Les modifications apportées sont :

- chaque tranche de signal est pondérée par une fenêtre h,
- les tranches peuvent se recouvrir.

□

3 Estimation de la DSPM d'un bruit blanc gaussien filtré

3.1 Génération du bruit à analyser

A quoi sert l'entier permettant d'initialiser le générateur ?

réponse ci-dessous

Le paramètre qui sert à initier le bruit correspond au nombre de réalisations que nous effectuons. La fonction ne plot que la dernière réalisation, c'est pourquoi nous ne voyons pas de différence à l'exécution du programme, le signal étant ergodique. □

3.2 Estimateur spectral simple

3.2.1 Script de la fonction Matlab développée

code ci-dessous

```
function [gamma1, f, N] = estimateur_simple(x, nd, nf, NFFT)

x_s = x(nd:nf);
N = length(x_s);

delta_f = 1/NFFT;
f = 0:delta_f:1-delta_f;

X = fft(x_s, NFFT);
gamma1 = (abs(X).^2)/N;

plot(f, 10*log10(gamma1), 'r'); hold on;
plot(fth, Gth, 'b');
plot(fth, Gbiais, 'g');
xlim([0 0.5]);
ylim([-50 10]);

end
```

□

3.2.2 Expérimentation

A. Etude du biais et de la variance en fonction du nombre N d'échantillons de bruit

figures ci-dessous

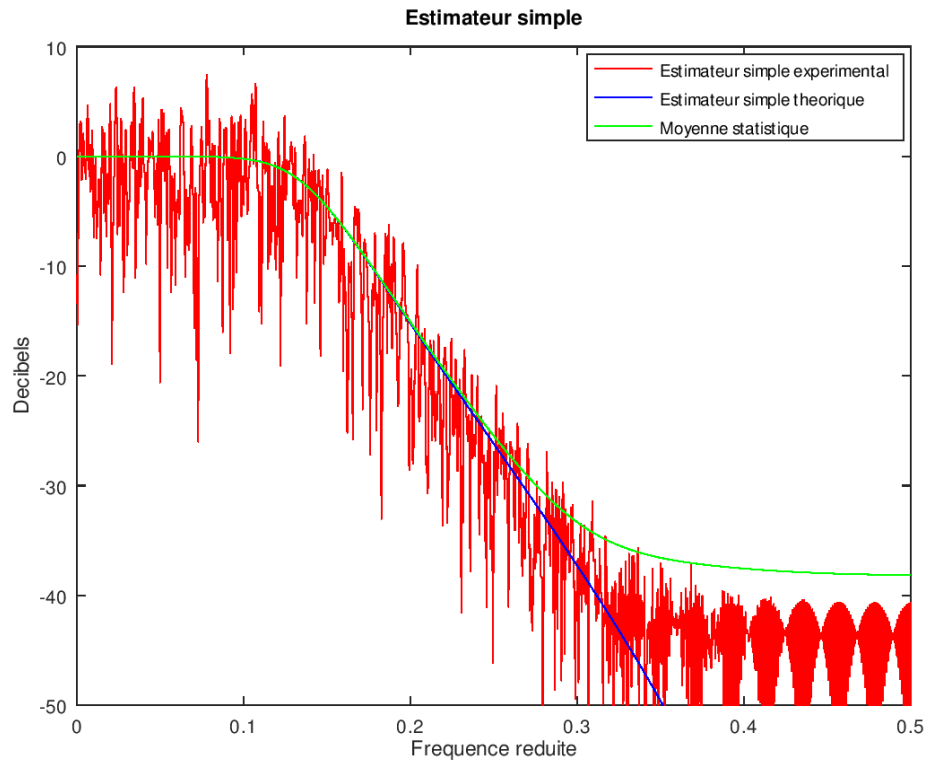


Figure 1 : N faible = 1000 – indice de début de la séquence à 1

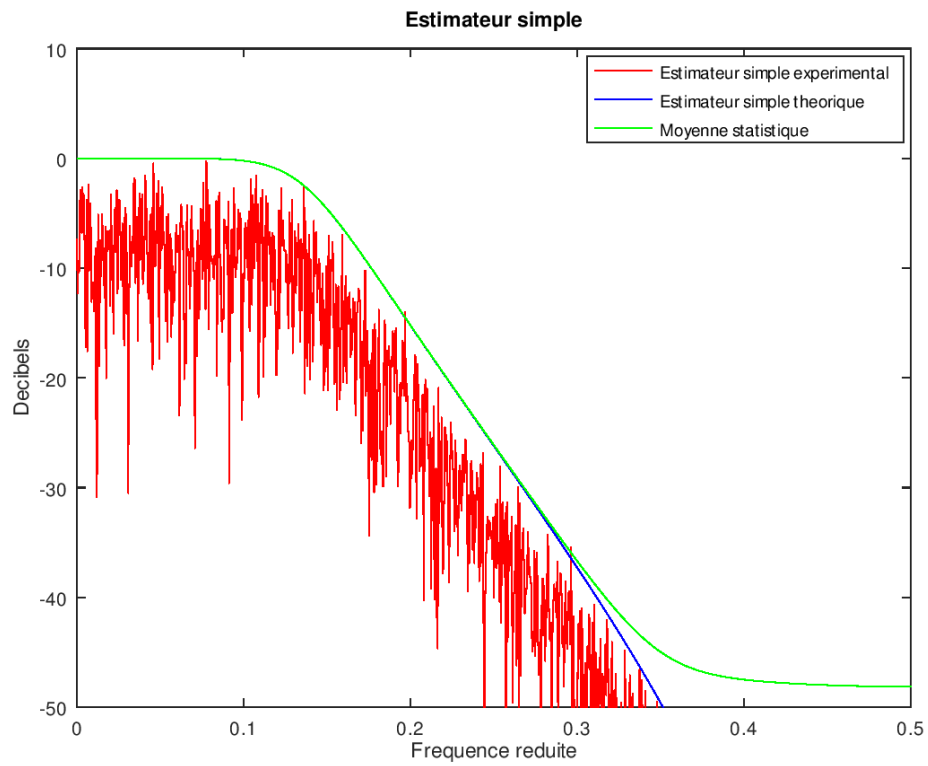


Figure 2 : N élevé = 10000 – indice de début de la séquence à 1

□

Commentaires.

réponse ci-dessous

Nous constatons que lorsque nous augmentons N , soit le nombre d'échantillons choisis, la courbe du biais a tendance à se rapprocher de la courbe de la DSPM théorique. La variance n'est en aucun cas modifiée (sauf cas extrême, par exemple nous avons testé avec $N = 100$). La variance et le biais étant les deux indicateurs de la qualité de notre estimateur, le fait d'augmenter le nombre d'échantillons a donc pour effet d'augmenter la précision, et donc la qualité, de notre estimateur. □

B. Etude du biais et de la variance en fonction de la réalisation considérée, à N fixé

figures ci-dessous

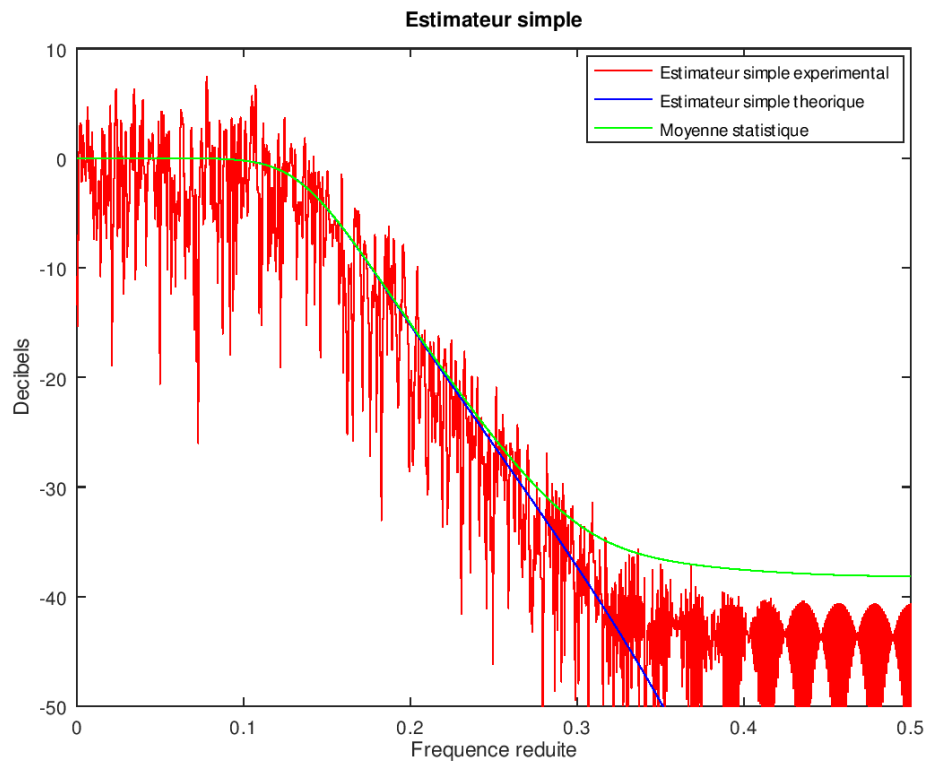


Figure 3 : $N \sim 1000$ – indice de début de la séquence à 1

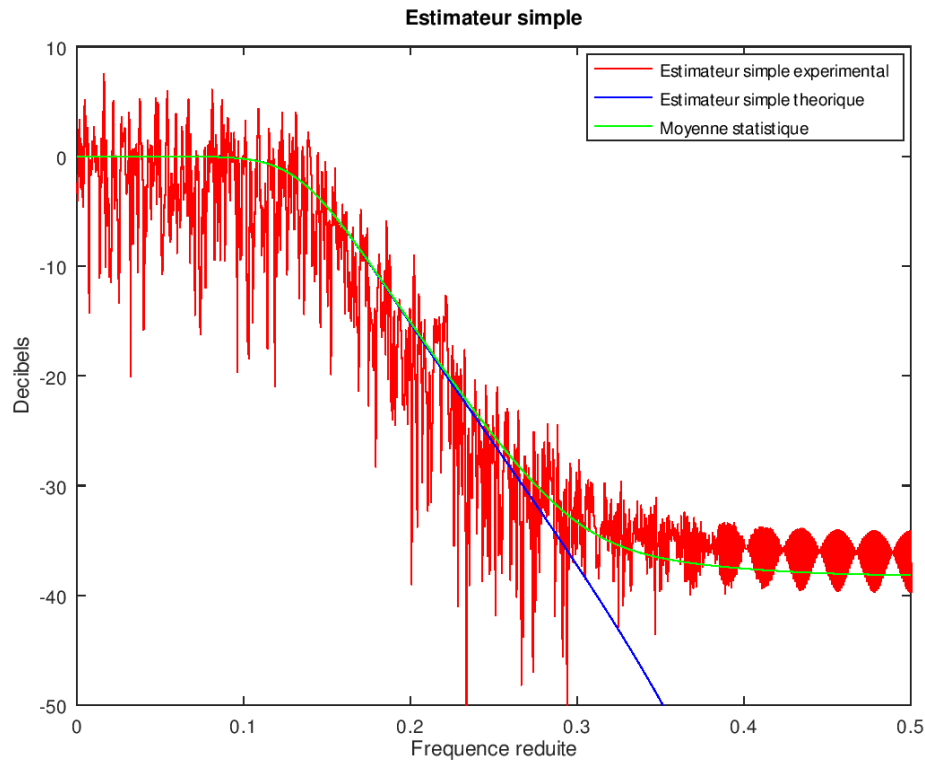


Figure 4 : $N \sim 1000$ – indice de début de la séquence fixé à 50000

□

Commentaires.

réponse ci-dessous

Nous fixons N à 1000. La première séquence est prise de 1 à 1001 échantillons, et la seconde de 50000 à 5100 échantillons. Nous pouvons constater que la DSPM calculée ne change pas en fonction de la séquence choisie. En effet, nous travaillons avec un signal stationnaire, qui voit donc ses propriétés conservées indépendamment du temps. En revanche nous pouvons observer un léger décalage de la DSPM calculée, dans notre cas le milieu de la séquence semble suivre de manière plus fidèle la moyenne statistique. Nous ne pouvons rien conclure sur ces constatations, si ce n'est que l'indice du début de la séquence a de l'importance sur l'allure du signal estimée, car certaines séquences semblent avoir des caractéristiques plus proches d'un signal blanc gaussien que d'autres.

□

C. Etude du biais et de la variance en fonction du nombre $NFFT$ de FFT

figures ci-dessous

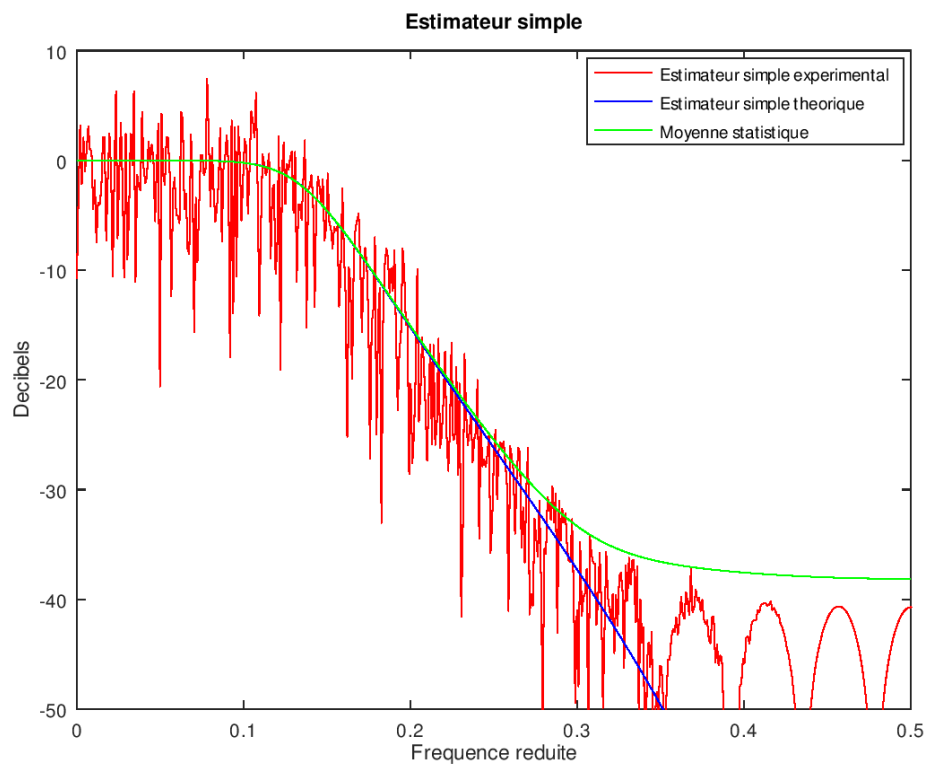


Figure 5 : N constant – indice de début de la séquence à $1 - NFFT : 2^{10} \sim N 1000$

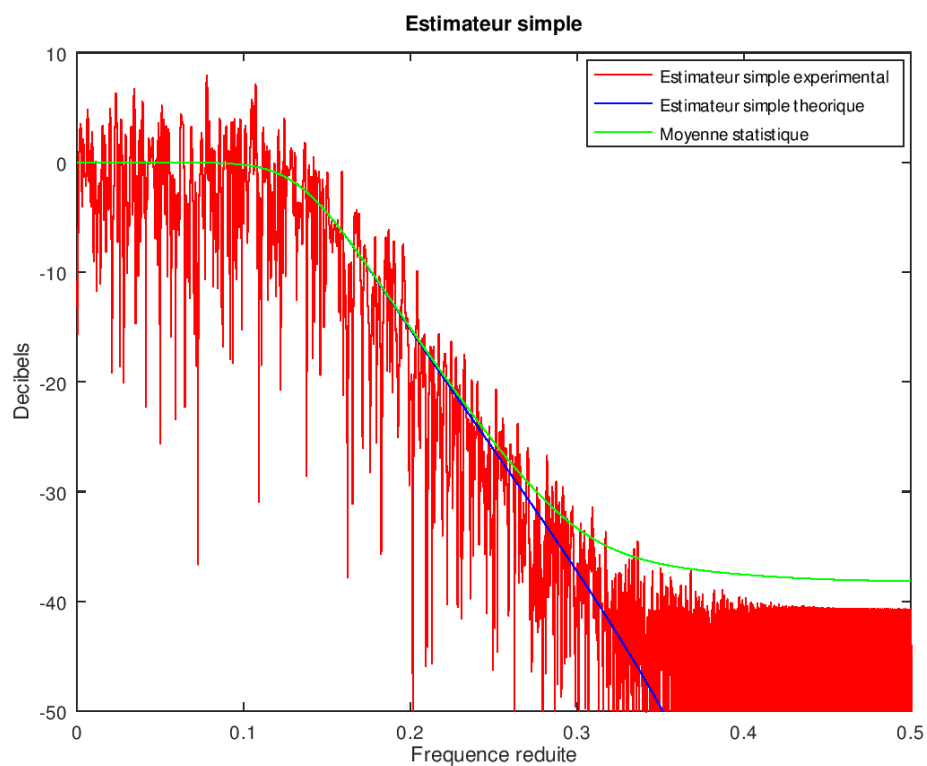


Figure 6 : N constant – indice de début de la séquence à $1 - NFFT : 2^{15} \gg N 1000$

□

Commentaires.

réponse ci-dessous

Nous constatons que si le nombre de points de la FFT est trop grand devant le nombre d'échantillons, le signal perd en lisibilité. Il semble cependant ne pas être modifié, comme nous pouvons l'observer dans les basses fréquences. C'est dans les hautes fréquences que nous perdons de l'information, ici nous ne pouvons plus observer les lobes. □

D. Conclusion Quel est le principal défaut de l'estimateur simple ?

réponse ci-dessous

Nous pouvons constater que N a une influence non négligeable sur la qualité de notre estimateur. Nous devons sélectionner le bon nombre d'échantillons pour que le biais ne soit pas mauvais, et adapter notre nombre de points pour la FFT afin d'obtenir un signal lisible sur lequel on pourra garder toutes les informations. Lorsque nous fixons N , nous observons que le signal ne change pas si on fait varier la position de la séquence. □

3.3 Estimateur spectral moyenné

On fixera $N = 4096$ dans tout ce qui suit.

3.3.1 Script de la fonction Matlab développé

code ci-dessous

```
function [gamma2, f] = estimateur_moy(x, N, M, NFFT)
x_s = x(1:N);
[gamma2, f] = pwelch(x_s, rectwin(M), 0, NFFT, 1, 'twosided');

plot(f2, 10*log10(gamma2), 'r'); hold on;
plot(fth, Gth, 'b');
plot(fth, Gbiais, 'g');
legend("Estimateur moyenne experimental", "Estimateur moyenne theorique", "Moyenne statistique");
title("Estimateur moyenné");
xlabel("Frequence reduite");
ylabel("Decibels");
xlim([0 0.5]);
ylim([-50 10]);

end
```

□

3.3.2 Expérimentation

En précisant bien la valeur des paramètres utilisés pour les essais, affichez les figures correspondantes aux conditions indiquées.

figure ci-dessous

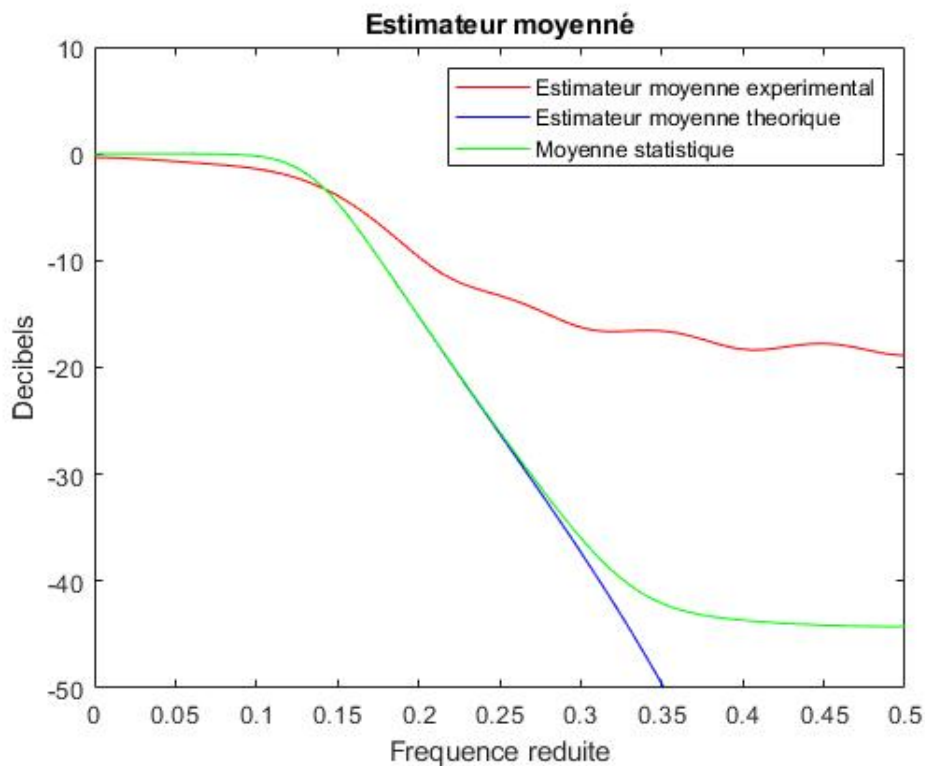


Figure 7 : $N = 4096$ – tranches courtes $M = 10$, $NFFT = 2048$

□

Commentaires

réponse ci-dessous

Ici, nous constatons que la variance est faible. En effet, les tranches sont courtes donc nous en avons beaucoup. Ainsi, nous pouvons faire une moyenne précise. En revanche, le biais est mauvais car nous n'avons pas beaucoup de points par tranches ce qui ne nous permet pas de faire de bonnes estimations. □

figure ci-dessous

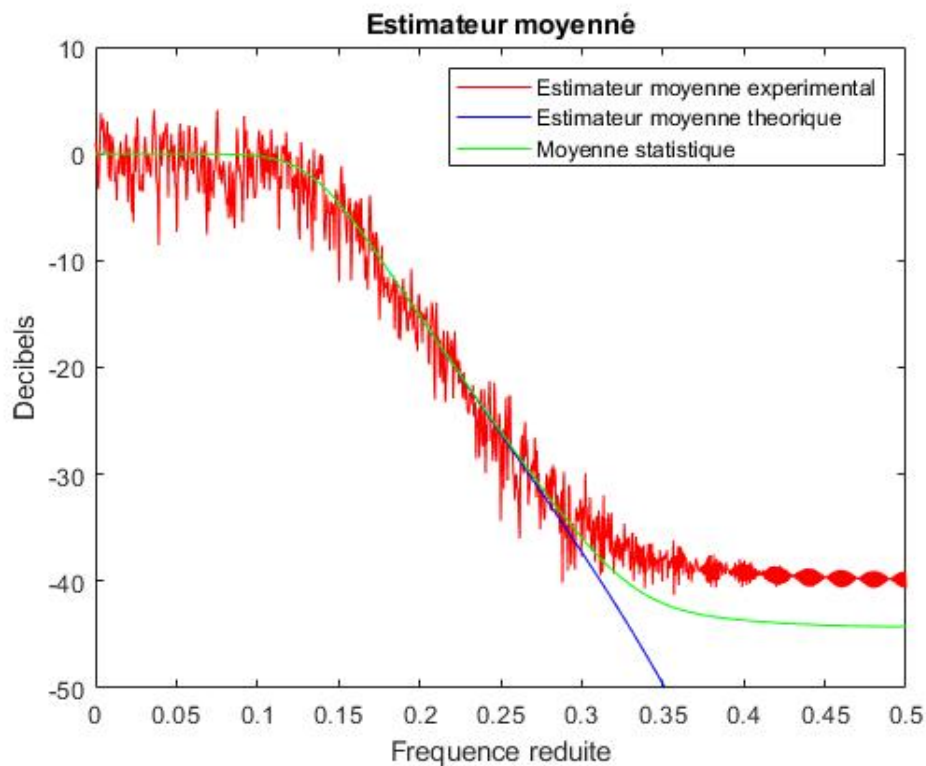


Figure 8 : $N = 4096$ – tranches longues $M = 1000$, $NFFT = 2048$

□

Commentaires

réponse ci-dessous

Ici, nous constatons l'inverse des tranches courtes. Les tranches étant longues, nous avons de nombreux points, donc la somme des espérances est précise. En revanche, nous n'avons pas beaucoup de tranches, donc la variance est mauvaise. □

figure ci-dessous

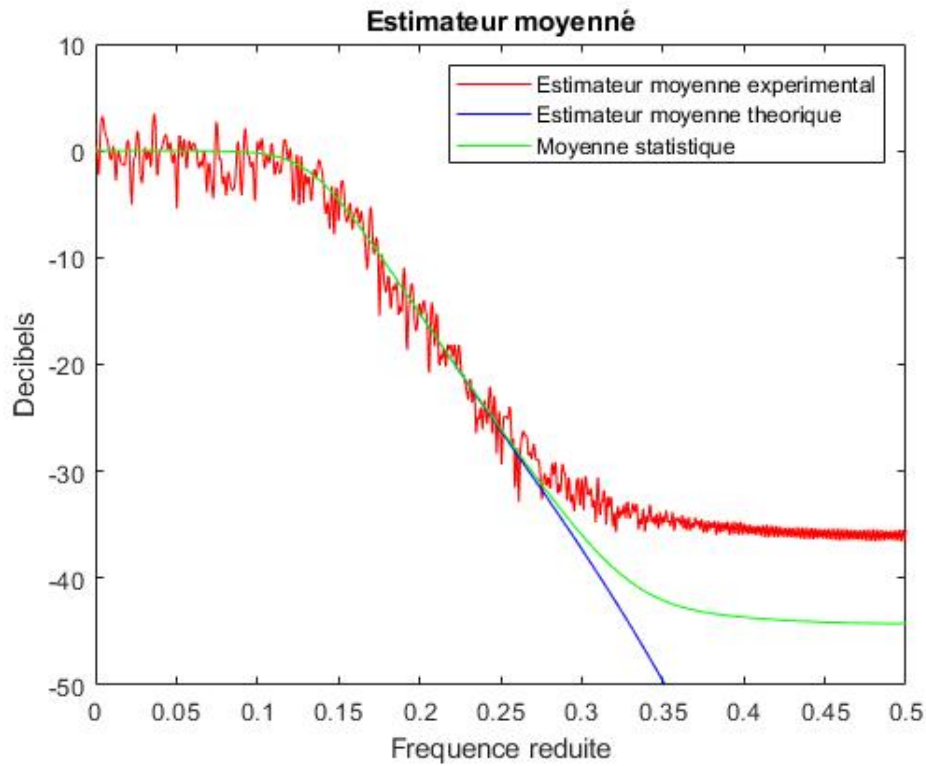


Figure 9 : $N = 4096$ – « Meilleur » compromis biais variance atteint pour $M = 500$, $NFFT = 2048$

□

Quelle information permettrait d'obtenir le meilleur compromis biais-variance ?

réponse ci-dessous

Nous avons fait varier M et constaté qu'il avait une influence directe sur la qualité des grandeurs d'estimations, le biais et la variance. Quand M est petit, la variance est précise mais pas le biais, et vice-versa. Il faut donc que nous calculions le M optimal pour que la qualité de la variance ne dégrade par celle du biais, et l'inverse également. Nous trouvons pour $N = 4096$ et $NFFT = 2^{11}$ un M optimal de 500 échantillons.

□

4 Estimateur de Welch

4.1 Etude préalable des fenêtres

Quelles différences de comportement fréquentiel peut-on observer pour les 6 fenêtres proposées (lobe principal, lobes latéraux...).

4.1.1 Script de la fonction Matlab développée

code ci-dessous

```
function [gamma3, f] = estimateur_welch(x, N, Nom_fenetre, M, NOVERLAP, NFFT)

x_s = x(1:N);
eval(['win=',Nom_fenetre,'(M);']);
[gamma3, f] = pwelch(x_s, win, NOVERLAP, NFFT, 1, 'twosided');

plot(f3, 10*log10(gamma3),'r'); hold on;
plot(fth, Gth, 'b');
plot(fth, Gbiais, 'g');
legend("Estimateur welch experimental","Estimateur welch theorique","Moyenne statistique");
title("Estimateur de Welch");
xlabel("Frequence reduite");
xlim([0 0.5]);
ylim([-50 10]);

end
```

□

4.1.2 Expérimentation

- A. Etude du biais et de la variance en fonction du taux de recouvrement entre tranches
Pour N , M et $NFFT$ fixés et pour une fenêtre choisie, tracez les figures correspondantes aux conditions indiquées ci-dessous.

figure ci-dessous

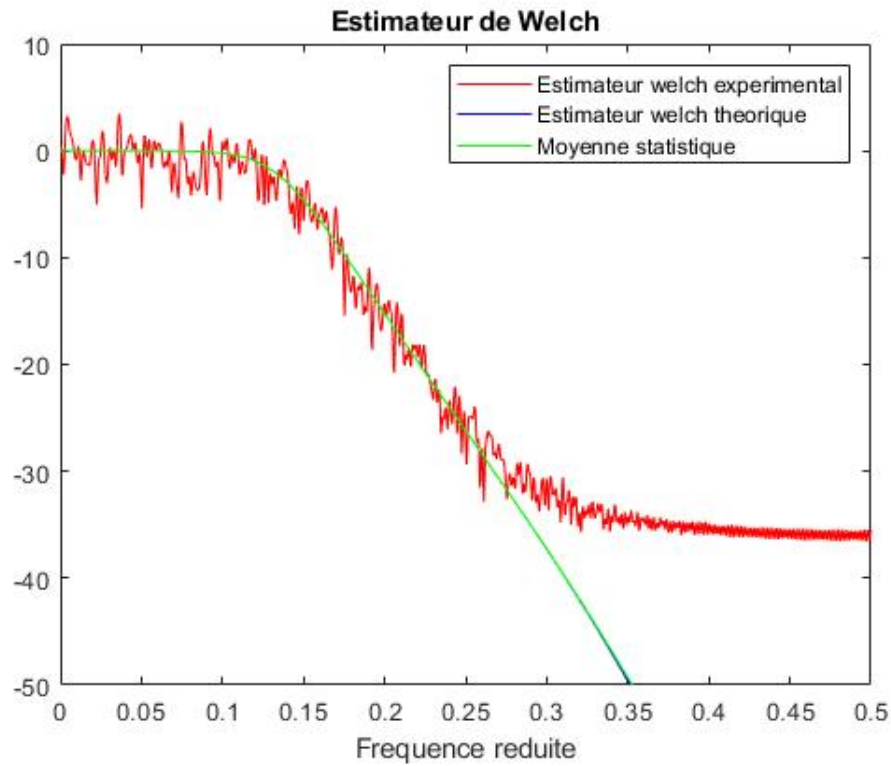


Figure 10 : $N = 4096$ – $M = 500$, $NFFT = 2048$. Choix de fenêtre = Rectangulaire – Recouvrement de 0%

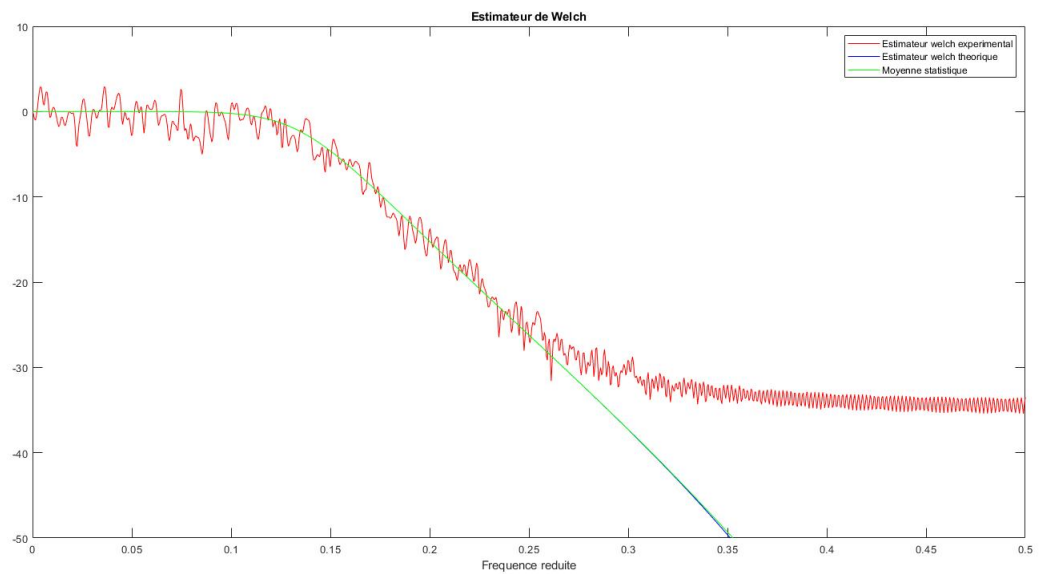


Figure 11 : $N = 4096$ – $M = 500$, $NFFT = 2048$. Choix de fenêtre = Rectangulaire – Recouvrement de 50%

□

Que permet le recouvrement entre tranches ?

réponse ci-dessous

Nous avons vu avec l'estimateur moyenné que nous devons jouer sur le paramètre M pour obtenir une variance et un biais satisfaisants. Implicitement, trouver un compromis entre le biais et la variance revient à devoir trouver un compromis entre le nombre de tranches et leur taille. Le recouvrement permet d'augmenter le nombre de tranches en gardant une taille satisfaisante. Or dans une gaussienne le principal de l'information se situe au milieu de la tranche, donc le recouvrement ne modifiera pas les résultats. \square

B. Etude du biais et de la variance en fonction de la fenêtre utilisée

Pour N , M et $NFFT$ fixés et pour différents choix de fenêtre, tracez les figures correspondantes aux conditions indiquées ci-dessous.

figure ci-dessous

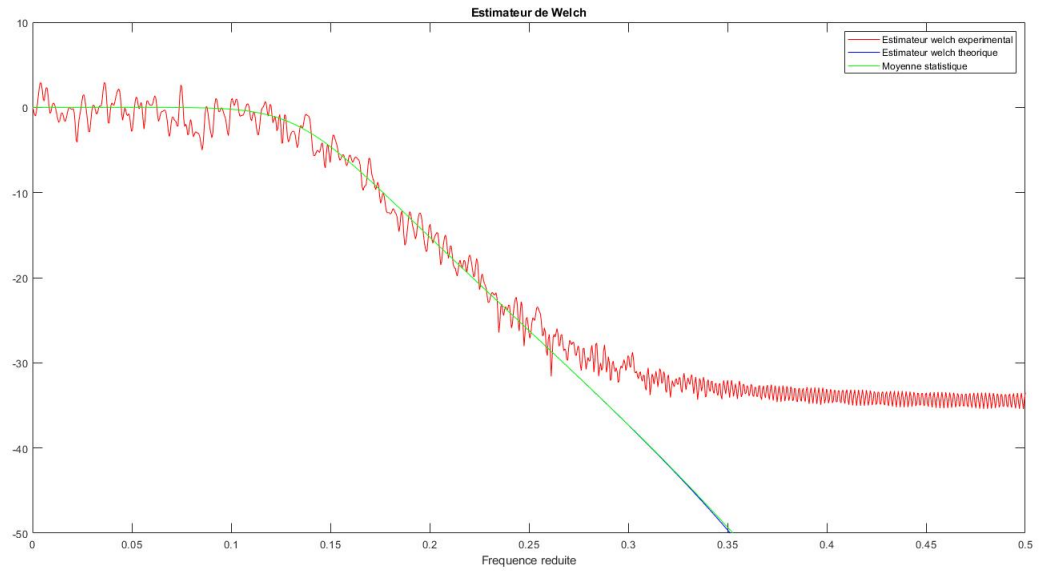


Figure 12 : $N = 4096$ – $M = 500$, $NFFT = 2048$. Fenêtre Rectangle – Recouvrement de 50%

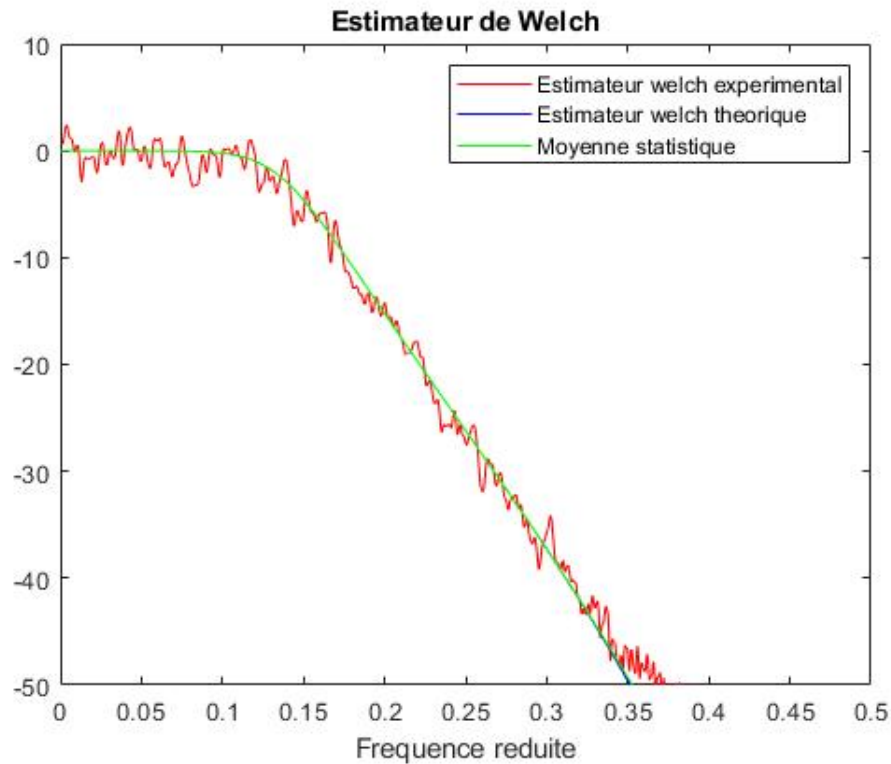


Figure 13 : $N = 4096$ – $M = 500$, $NFFT = 2048$. Fenêtre Hamming – Recouvrement de 50%

□

Que permet l'utilisation d'une fenêtre autre que rectangulaire ? Expliquer.

réponse ci-dessous

Nous avons vu dans le cours que la précision et la résolution dépendent de la fenêtre choisie. En utilisant une fenêtre autre que rectangulaire, nous pourrions améliorer la précision d'un autre paramètre. La fenêtre rectangulaire accorde autant d'importance au signal situé au bord de la fenêtre que n'importe où ailleurs dans le signal. Or avec la gaussienne, l'essentiel est situé au centre de la fenêtre, et les bords viennent en fait fausser nos paramètres de qualité, dans le cas présent le biais. Ainsi, utiliser une autre fenêtre améliorera la qualité de notre biais.

□

Pour quelles valeurs des paramètres d'analyse obtenez vous le « meilleur » résultat (celui qui vous paraît le plus satisfaisant) ?

réponse ci-dessous

Longueur de la séquence analysée $N = 4096$
 Longueur des tranches $M = 500$
 Type de fenêtre Hamming
 Taux de recouvrement = 50
 Nombre de points de transformée de Fourier $NFFT = 2048$

□

5 Utilisation des estimateurs précédents pour analyser un signal inconnu

5.1 Modification des programmes

Script d'une des fonctions modifiée

code ci-dessous

```
function [gamma3, f] = estimateur_welch(x, N, Nom_fenetre, M, NOVERLAP, NFFT)

x_s = x(1:N);
eval(['win=',Nom_fenetre,'(M);']);
[gamma3, f] = pwelch(x_s, win, NOVERLAP, NFFT, 1, 'twosided');

hold on;
semilogy(f, gamma);
legend("DSPM estimée");
title("DSPM estimée avec l'estimateur de welch");
```

□

5.2 Expérimentation

Afficher les spectres estimés obtenus avec chacune des 3 méthodes étudiées.

figure ci-dessous

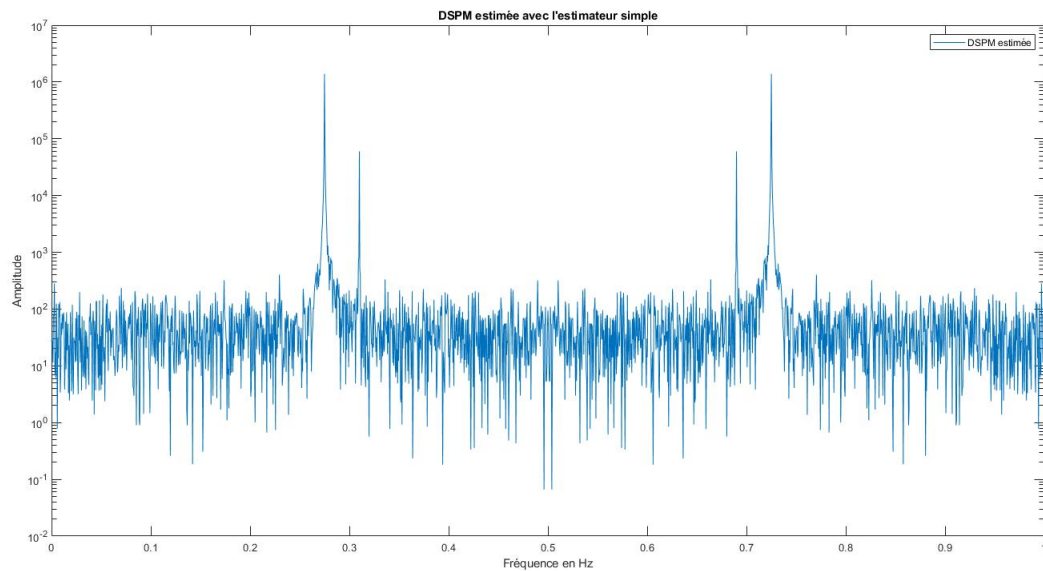


Figure 14 : Estimateur spectral simple. $nd = 1$ $nf = 4097$ $NFFT = 2048$.

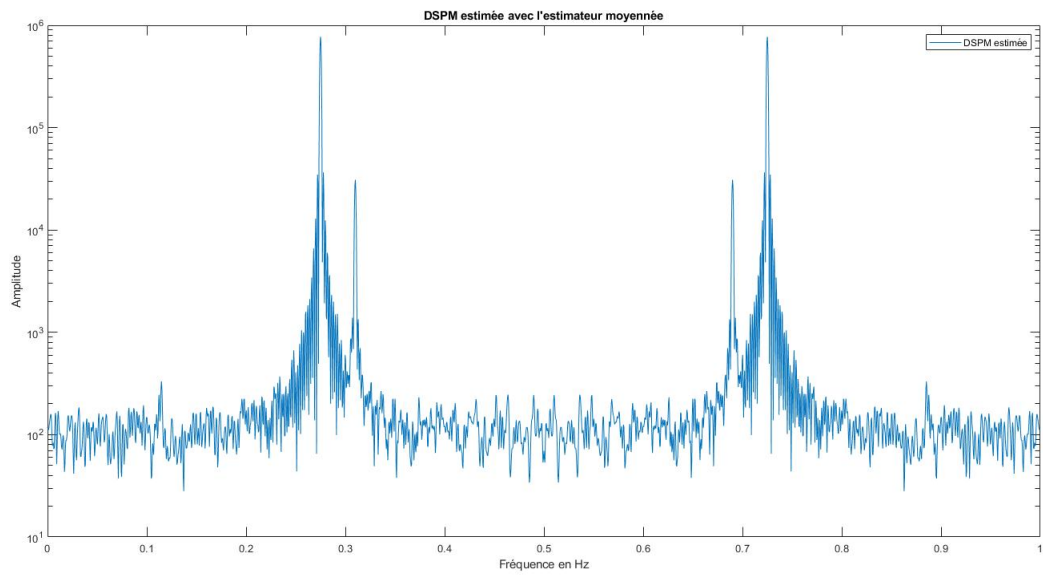


Figure 15 : Estimateur spectral moyenné. $N = 4096$ $NFFT = 2048$ $M = 500$.

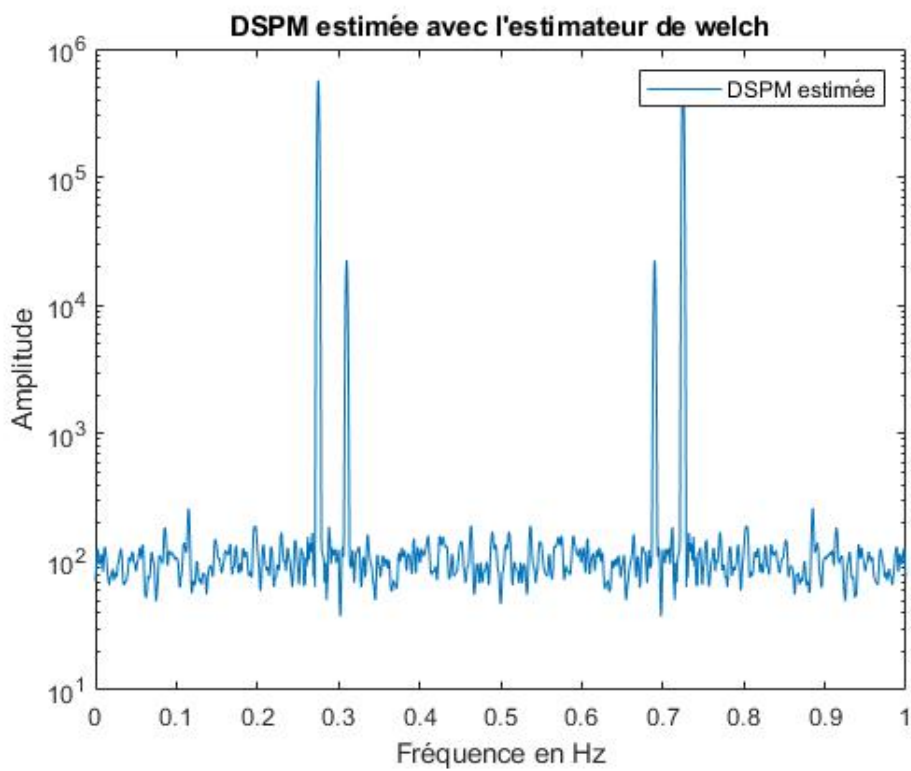


Figure 16 : Estimateur spectral de Welch. $N = 4096$ $NFFT = 2048$ $M = 500$ $NOVERLAP = 50\%$
fenêtre utilisée = hamming.

□

Décrivez précisément la démarche expérimentale suivie. Avec quelle méthode êtes vous capable avec certitude de décrire le contenu fréquentiel de ce signal ?

réponse ci-dessous

Afin de déterminer la DSPM de ce signal inconnu avec le plus de certitude possible, nous avons appliqué les différents estimateurs sur ce dernier avec les paramètres les plus appropriés.

Nous avons donc pris pour l'estimateur simple : $nd = 1$, $nf = 4097$ afin d'avoir comme demandé 4096 échantillons et $NFFT = 2048$ afin de ne pas avoir une variance mauvaise dans les hautes fréquences.

Pour l'estimateur moyenné, nous avons pris les mêmes valeurs de N et $NFFT$ pour les mêmes raisons citées précédemment. Pour le nombre d'échantillons par tranches M , nous nous sommes logiquement placés dans le cas optimal déterminé dans les parties précédentes, c'est-à-dire pour $M = 500$ échantillons.

Pour l'estimateur de Welch, nous nous sommes replacé dans les mêmes conditions que précédemment pour les paramètres N , $NFFT$ et M .

Pour ce qui est du taux de recouvrement, nous l'avons pris égal à 50% afin de pouvoir avoir à la fois une largeur et un nombre de fenêtres satisfaisants, et donc d'améliorer le compromis biais-variance comme nous le recherchions dans la partie précédente.

En ce qui concerne le choix de la fenêtre, comme expliqué dans les questions précédentes, la fenêtre rectangulaire n'est pas un bon choix, nous nous sommes donc tournés vers la fenêtre "hamming" avec laquelle nous avons eu le meilleur résultat.

Nous sommes capable de décrire avec certitude le contenu fréquentiel de ce signal grâce à l'estimateur de Welch. En effet, cet estimateur nous permet d'affiner le résultat grâce à des paramètres supplémentaires tel que le choix de la fenêtre et le taux de recouvrement du signal.

□

5.2.1 Interprétations

- A. Quel inconvénient majeur l'utilisation d'une fenêtre (d'apodisation en temps) engendre-t-elle ?

_____ **réponse ci-dessous** _____

La fenetre rectangulaire est en fait une porte donc nous n'étudions pas le signal, mais la convolution entre le signal et la TF de la porte.

□

- B. Décrire (sans dessin) la forme de la DSPM obtenue.

_____ **réponse ci-dessous** _____

On observe 4 pics situés autour de 0.3 Hz et 0.7 Hz, avec une symétrie du signal située à 0.5 Hz.

□

- C. Quelles informations la forme de cette DSPM apporte-t-elle sur le contenu (la nature) du signal ?

_____ **réponse ci-dessous** _____

On peut donc conclure que la quasi-totalité de la puissance du signal est concentrée autour de 2 fréquences qui sont 0.3 Hz et 0.7 Hz environ. Il semble aussi que la courbe de la DSPM ressemble à des diracs. Nous pouvons conclure quand à l'allure sinusoïdale du signal temporel.

□

- D. Quelles mesures concernant les caractéristiques du signal peut-on effectuer sur la DSPM ?

_____ **réponse ci-dessous** _____

La DSPM représente la répartition fréquentielle de la puissance d'un signal suivant les fréquences qui le composent. D'après le théorème de Wiener-Khintchine, elle est égale à la TF de l'auto-corrélation du signal. Nous pouvons donc en tirer la puissance du signal et son autocorrélation.

□