

# Traitement des Signaux Aléatoires

## Estimation Spectrale

4 ETI – CPE Lyon

Travaux Pratiques TSA

2020-2021

Noms, Prénoms : BRUYERE Axel CHARRAT Philippe

Groupe : D

Date : 20/10/2020

### 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 \_\_\_\_\_

☐

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

\_\_\_\_\_ réponse \_\_\_\_\_

☐

**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 \_\_\_\_\_

☐

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

\_\_\_\_\_ réponse \_\_\_\_\_

☐

**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 \_\_\_\_\_

☐

**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 \_\_\_\_\_

☐

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

\_\_\_\_\_ réponse \_\_\_\_\_

☐

**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 \_\_\_\_\_

☐

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

\_\_\_\_\_ réponse \_\_\_\_\_

☐

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

\_\_\_\_\_ réponse \_\_\_\_\_

☐

## 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 ?

Le problème est qu'il est compliqué d'analyser des processus aléatoire car cela reviendrait à comparer des signaux avec les même caractéristique mais des réalisations différentes. L'option *seed* des fonctions *rnd*, *rand* sur MATLAB nous permet de fixer la suite de nombre aléatoires dans un ordre précis pour que les *n* réalisations soit identiques. Ainsi, on étudie un processus pseudo-aléatoire car prédictible.

### 3.2 Estimateur spectral simple

#### 3.2.1 Script de la fonction Matlab développée

Le code de la fonction *ESS*, il est tronqué de sa partie affichage et de ses commentaires.

```
function ESS(x,nd,nf,NFFT) ;
    % ---Initialisation des variables ---
    x_seq = x(nd : nf); %Sequence Ã analyser
    N = nf - nd +1; %Longueur de la sequence

    % ---Création de l'estimateur 1 ---
    X = fft(x_seq,NFFT); %Transofrmée N points de la séquence
    gamma_x_c = ((abs(X)).^2)/N; %Estimation simple
    log_gamma_x_c = 10*log10(gamma_x_c); %Passage au log (forme quadratique donc log * 10)

    % ---DSP moyenne vraie et (gamma(f) * Wbm(f))(f)
    [Gth,Gbiais,fth]=sptheo(N,'simple');
    f_abs = 0:1/NFFT:1-1/NFFT;

    % ---Partie affichage ---
    figure(2)
    plot(f_abs,log_gamma_x_c,fth,Gth,'k',fth,Gbiais,'r')
    axis([0 0.5 -60 10])
    legend('Estimation de la DSP','DSPMV','Convolution de la DSP et de la fenetre de Barlett')
    title("Etude de l'estimateur simple")
end
```

### 3.2.2 Expérimentation

- A. Étude du biais et de la variance en fonction du nombre  $N$  d'échantillons de bruit  
Pour l'étude du biais, nous allons faire varier le paramètre d'entrée  $N$  de la fonction  $ESS$  de 1000 échantillons à 100000 échantillons.

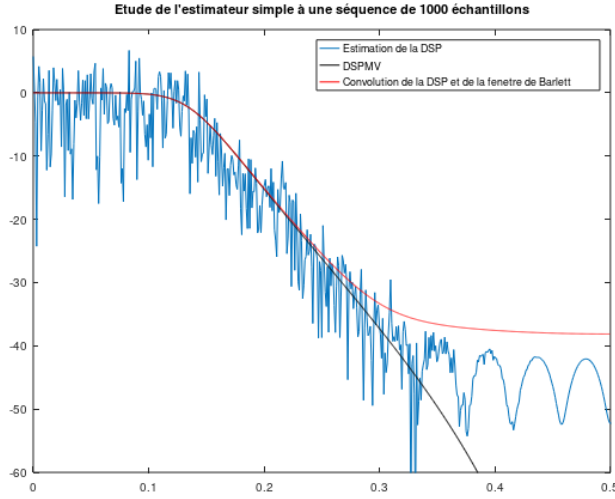


FIGURE 1 –  $N$  faible (1000)– indice de début de la séquence à 1

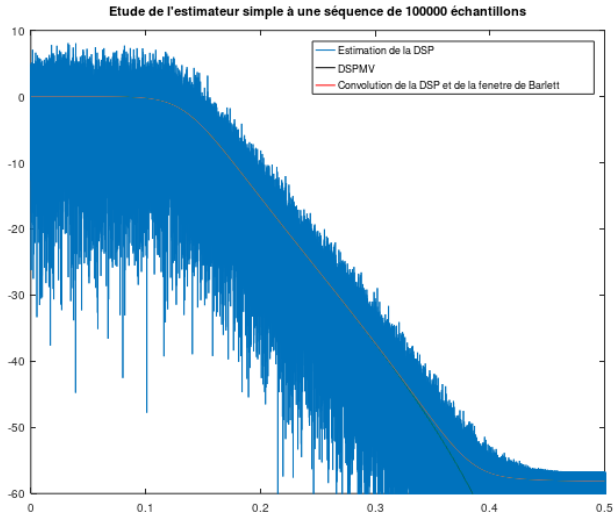


FIGURE 2 –  $N$  élevé (100 000)– indice de début de la séquence à 1

Sur les figures 1 et 2, nous pouvons observer 3 courbes, la bleu correspond à la densité spectrale obtenues à partir d'une séquence empirique à  $N$  points. La courbe noire correspond à la densité spectrale théorique non biaisée et la rouge correspond à la densité spectrale convolué par une fenêtre de BARLETT (donc biaisé). Ainsi, nous pouvons constater que si  $N$  est faible alors la DS empirique correspond à l'estimateur non-biaisé sur les basses-fréquences mais va être biaisé sur les hautes fréquences et une variance qui est "faible" sur toutes la DSP. Inversement, avec un  $N$  important le biais va être "faible" car la courbe bleu suit la DST mais la variance est très importantes.

Nous pouvons expliquer ce comportement à l'aide du phénomène de convolution par la fenêtre de BARRET car plus on a une fenêtre tend vers l'infinie (plus physiquement le signal est long dans le temps) et plus le replie temporelle va être faible donc les hautes fréquences de la DST ne vont pas être impactées par les basses fréquences et le biais diminues. Le problème est que la variance

est obtenues à l'aide de la relation :

$$VAR(ES) = \Gamma^2(x) + [1 + (\sin(2 * \pi * f * N) / (N * 2 * \pi * f))^2]$$

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

Pour cette étude, nous avons fixé taille du  $N$  à 1000 échantillons et l'on obtient les figures suivantes :

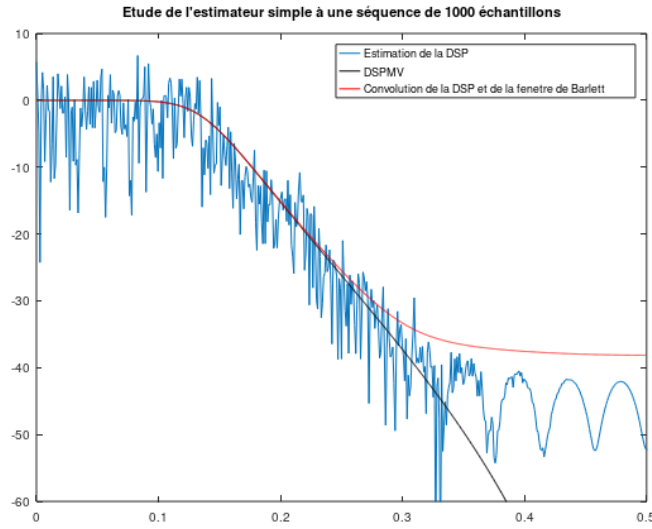


FIGURE 3 –  $N \sim 1000$  – indice de début de la séquence à 1

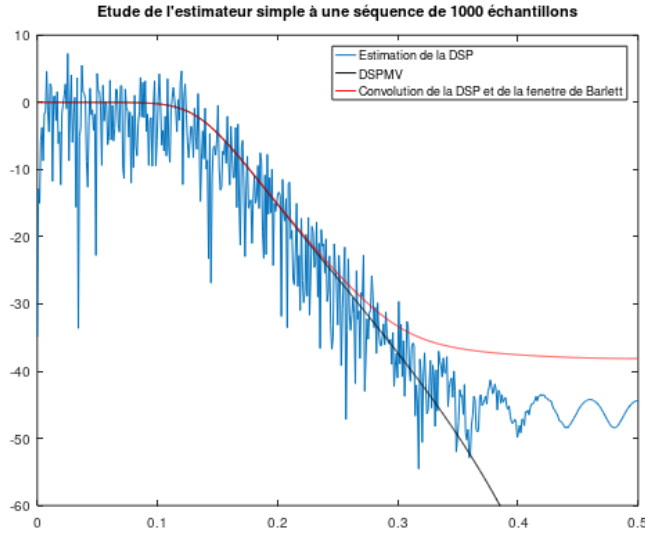


FIGURE 4 –  $N \sim 1000$  – indice de début de la séquence fixé à une autre position ( $\gg 1000$ ) dans la séquence

Nous pouvons observer que sur les figure 3 et 4 (pour une séquence qui débute à 2000), les courbes obtenues sont semblables. Ce résultat est cohérent car le signal observé est un bruit blanc Gaussien qui a comme propriété d'être stationnaire donc l'intervalle d'échantillons peut être quelconque car n'influe pas sur l'étude, tant que l'on conserve une même taille  $N$ .

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

Pour cette étude, nous avons fixé taille du  $N$  à 1000 échantillons et l'on obtient les figures suivantes :

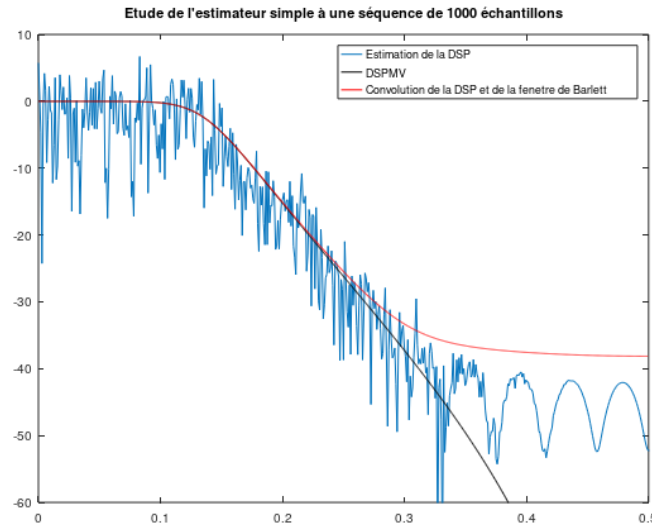


FIGURE 5 –  $N$  constant – indice de début de la séquence à  $1 - NFFT \sim N$

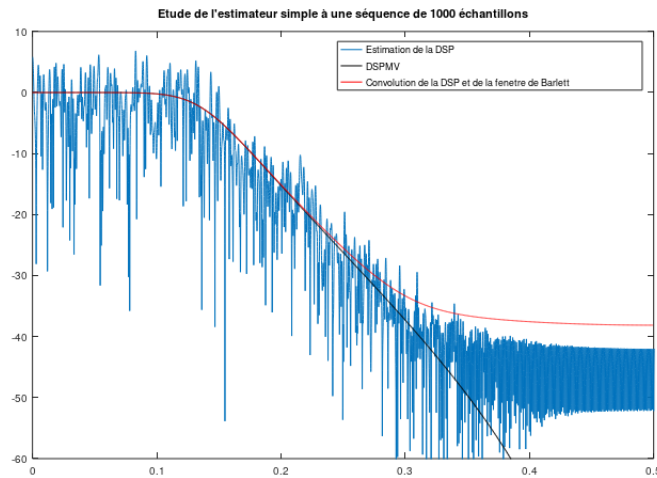


FIGURE 6 –  $N$  constant – indice de début de la séquence à  $1 - NFFT(2 * 14) \gg N$

Nous pouvons observer que sur les figure 5 et 6 (une FFT sur  $2^{14}$  points), les courbes obtenues sont semblables. Ce résultat est cohérent car pour obtenir les densités spectrales le paramètre  $NFFT$  n'influence que le nombre de points pour la Transformée de Fourier à  $N$  points. Et va permettre d'augmenter la précision de la courbe. Ainsi il ne peut pas influer sur l'étude, tant que l'on prend  $NFFT > N$ .

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

A l'aide de l'étude réalisée, nous pouvons conclure que l'estimateur simple a comme principal défaut sa variance car  $N$  est grand est plus le terme carré pour obtenir la variance va "exploser" et cela devient compliqué d'analyser la courbe. Ainsi cet estimateur n'est pas de bonne qualité.

### 3.3 Estimateur spectral moyenné

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

#### 3.3.1 Script de la fonction Matlab développé

Le code de la fonction *ESM* ; il est tronqué de sa partie affichage et de ses commentaires.

```
function ESM(x,N,M,nfft)
    % Inputs :
    % x - séquence brut
    % N - nombre d'échantillons maximum
    % M - nombre d'échantillons d'une sous fenêtre
    % nfft - nombre de points fft à N points
    % Outputs :
    % none

    % Initialisation des variables ---
    x_seq = x(1 : N); %séquence à analyser
    fenetre = rectwin(M); % Fenetre taille M << N
    noverlap = 0; % Chevauchement

    % Création de l'estimateur 2 ---
    [gamma_xd_c,fabs] = pwelch(x_seq,fenetre,noverlap,nfft,1,'twosided');
    log_gamma_xd_c = 10 * log10(gamma_xd_c);

    % DSP moyenne vraie et (gamma(f') * Wbm(f'))(f) ---
    [Gth,Gbiais,fth]=sptheo(M,'moyenne'); % 1er param : M
end
```

#### 3.3.2 Expérimentation

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

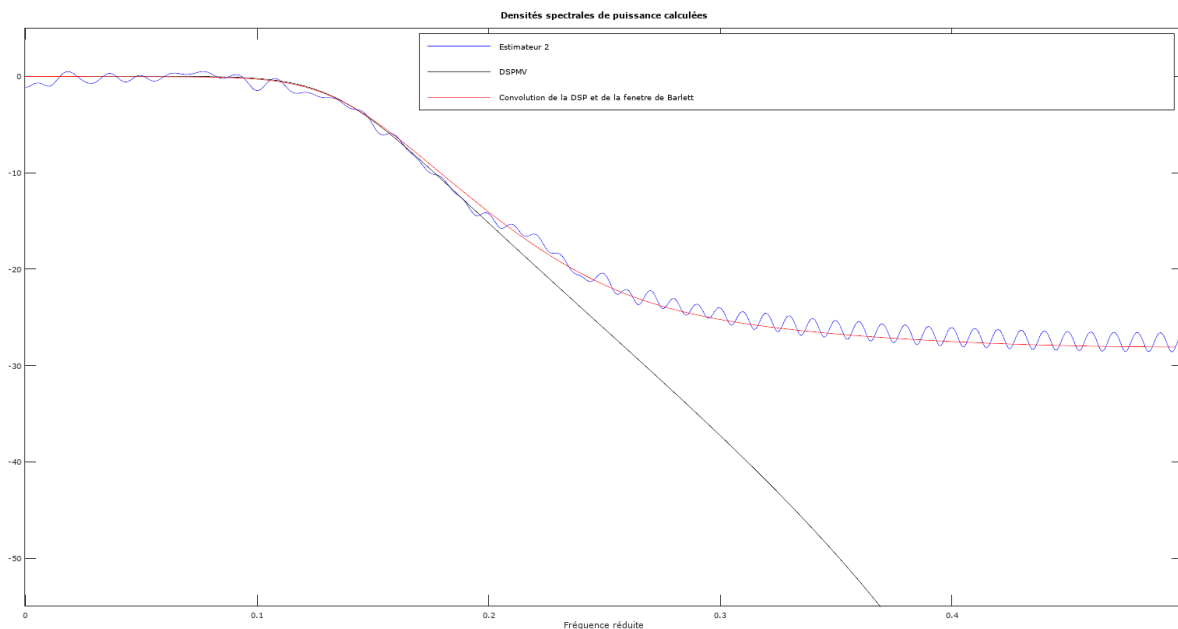


FIGURE 7 –  $N = 10.000$  – tranches courtes  $M = 500$ ,  $NFFT = 1024$

On observe ici une variance plutôt faible mais un biais assez élevé.

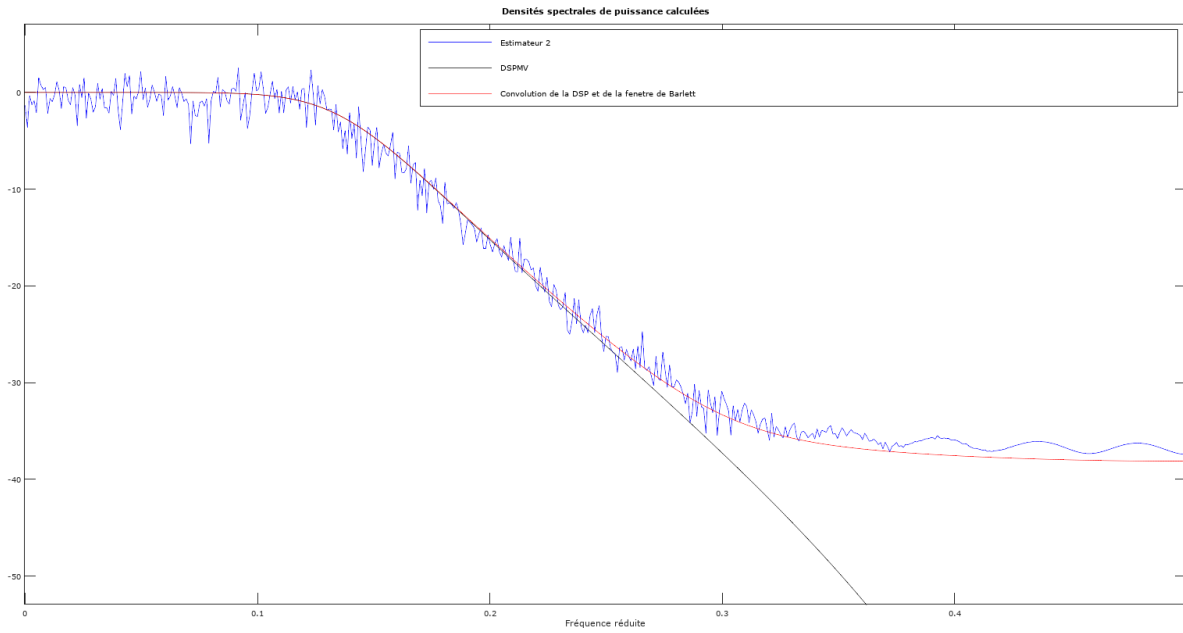


FIGURE 8 –  $N = 10.000$  – tranches courtes  $M = 1000$ ,  $NFFT = 1024$

Ici, on observe un biais plus faible mais une variance qui augmente. En effet, pour réduire le biais, il faut augmenter la largeur des tranches. En revanche, ceci a pour effet d'augmenter la variance. A l'inverse, en réduisant la largeur des tranches, on diminue la variance mais on augmente le biais. Le choix de ces valeurs repose donc sur le cahier des charges, censé spécifier l'importance de chacun de ces paramètres, donc implicitement l'équilibre voulu entre eux.

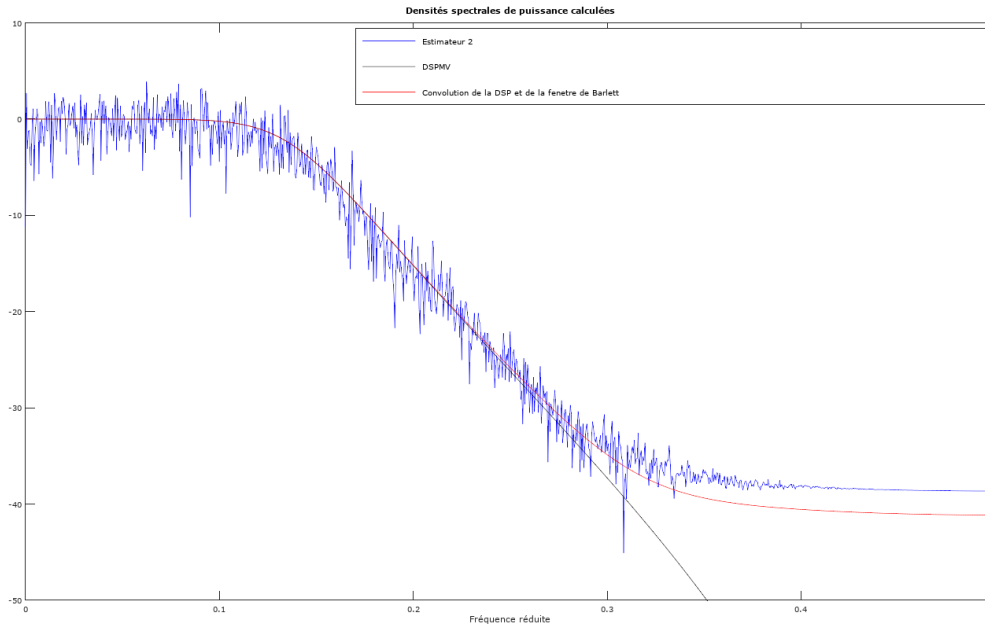


FIGURE 9 –  $N = 10.000$  – « Meilleur » compromis biais variance atteint pour  $M = 2000$ ,  $NFFT = 512$

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

Afin d'obtenir le meilleur compromis biais-variance possible, il nous faudrait définir un degré d'importance pour le biais ainsi que pour la variance, et ainsi savoir lequel prioriser.



## 4 Estimateur de Welch

### 4.1 Etude préalable des fenêtres

Quelle différence de comportement fréquentiel peut-on observer sur les 6 fenêtres proposées(lobe principal, lobes latéraux) ?

Plus la fenêtre choisie est large dans le domaine temporel, plus elle comprendra de lobes secondaires, et plus elle sera étroite dans le domaine fréquentiel. De plus, dans le but de laisser passer les maxima du lobe principal, la fenêtre choisie doit être suffisamment haute. La fenêtre rectangulaire possède un lobe principal étroit et atténue de façon égale le reste des fréquences à l'instar de la fenêtre de Hanning. Les fenêtres triangulaire et de Hamming possède un lobe principal large et atténue de façon linéaire les autres fréquences. La fenêtre de Gauss possède, elle, un lobe principal étroit en haut, plus large en bas et atténue très fortement les fréquences qu'elle ne laisse pas passer. Enfin, la fenêtre de Blackman possède un lobe principal large.

#### 4.1.1 Script de la fonction Matlab développée

Le code de la fonction *ESM* ; il est tronqué de sa partie affichage et de ses commentaires.

```
function ESW(x,N,fenetre_char,M,NOVERLAP,NFFT)

% ---Initialisation des variables ---
x_seq = x(1:N);
eval(['WIN=',fenetre_char,'(M)']);
f_abs = 0:1/N:1-1/N;

% ---Création de l'estimateur 3---
[gamma_x_c,fabs] = pwelch(x_seq,WIN,NOVERLAP,NFFT,1,'twosided');
gamma_x_c = 10 * log10(gamma_x_c);

% ---DSP moyenne vraie et (gamma(f) * Wbm(f))(f)
[Gth,Gbiais,fth]=sptheo(M,'welch','rectwin');
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.

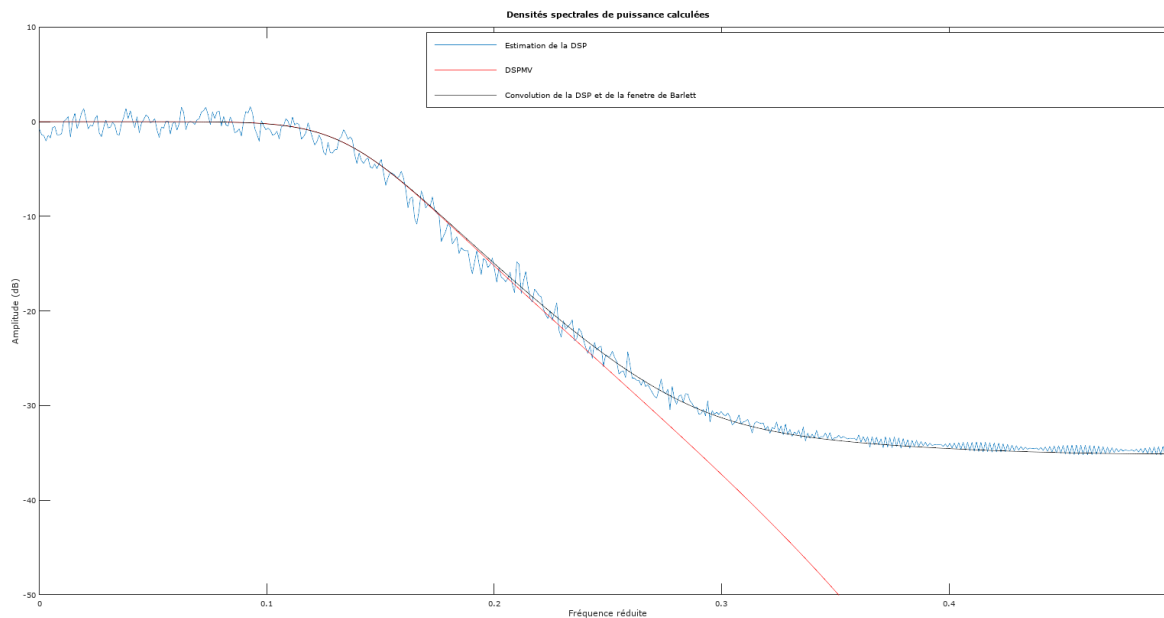


FIGURE 10 –  $N = 10.000$  –  $M = 2000$ ,  $NFFT = 1024$ . Choix de fenêtre = RECTWIN – Recouvrement de 25%

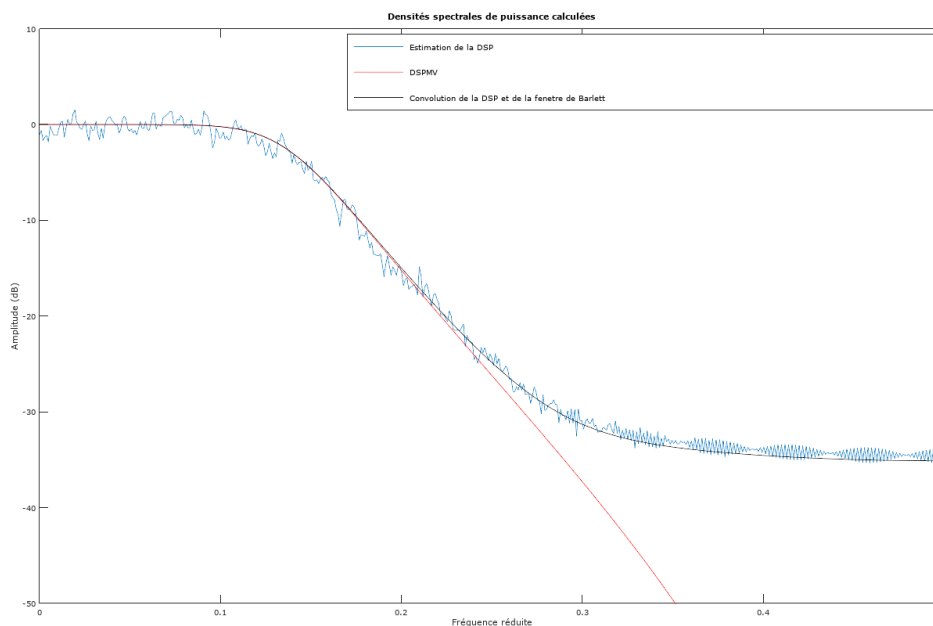


FIGURE 11 –  $N = 10.000$  –  $M = 2000$ ,  $NFFT = 1024$ . Choix de fenêtre = RECTWIN – Recouvrement de 50%

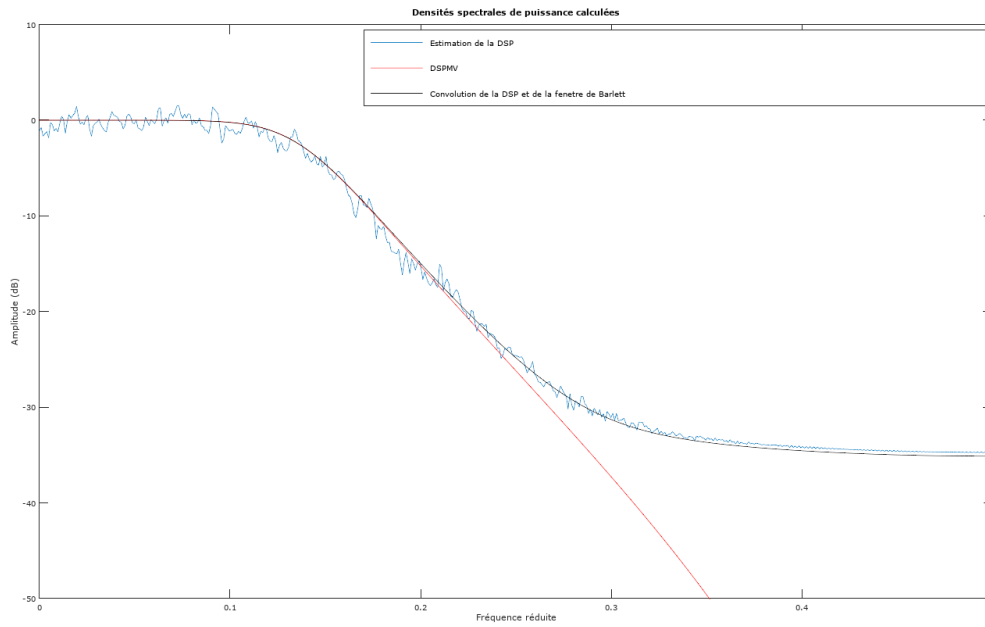


FIGURE 12 –  $N = 10.000$  –  $M = 2000$ ,  $NFFT = 1024$ . Choix de fenêtre = RECTWIN – Recouvrement de 75%

Que permet le recouvrement entre tranches ?

Le recouvrement entre tranche permet de jouer sur deux paramètres à la fois : on augmente le nombre de tranches, sans devoir diminuer leur longueur. Cela permet de trouver un bien meilleur compromis biais-variance que l'estimateur moyenné, et de très largement diminuer la variance par rapport à l'estimateur simple.

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.

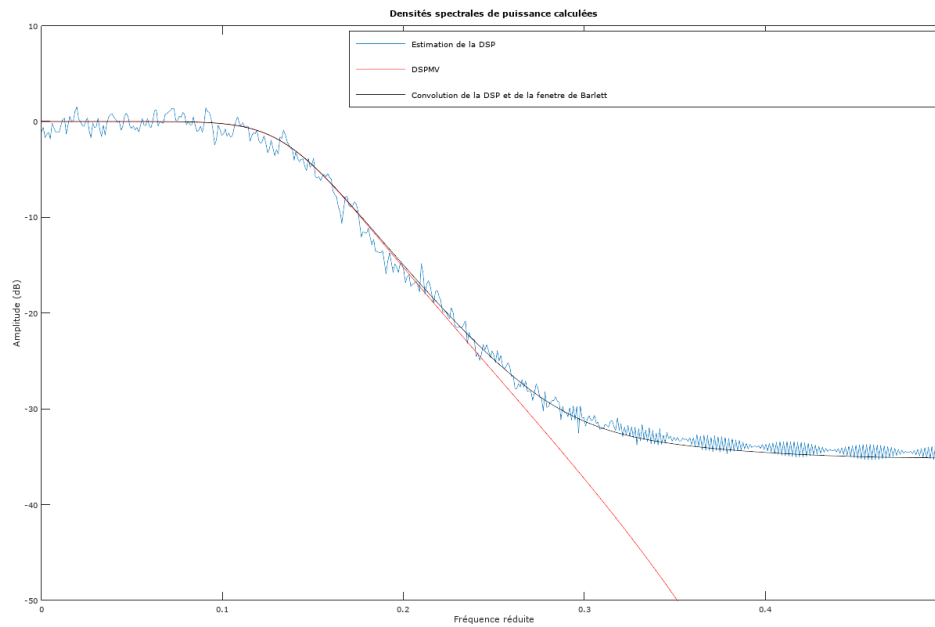
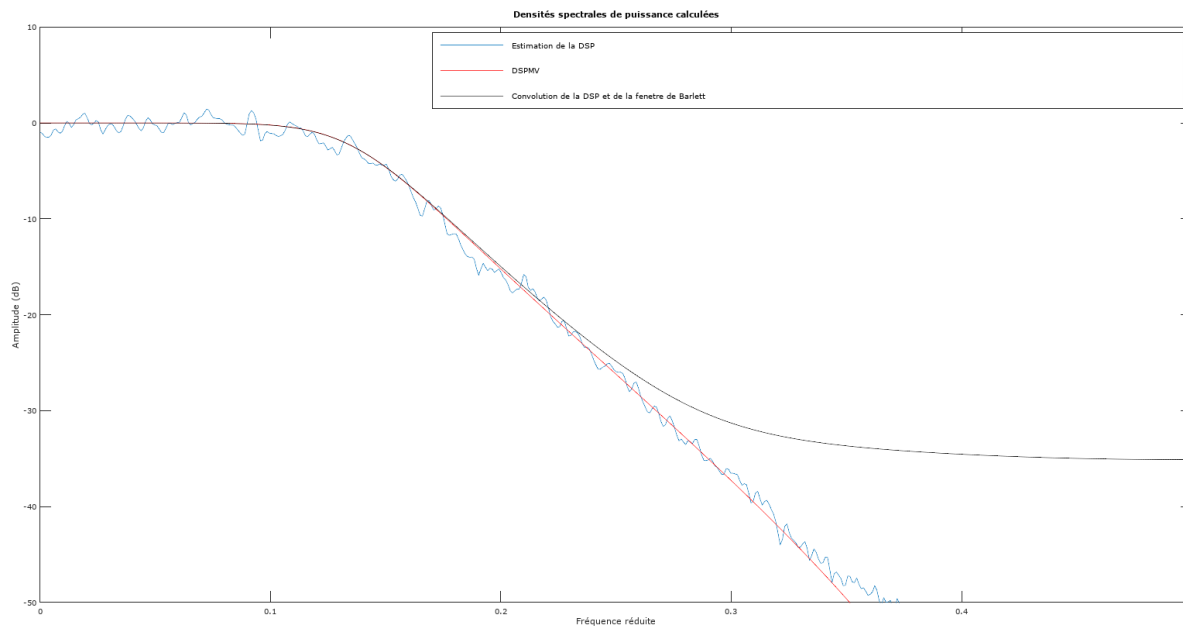


FIGURE 13 –  $N = 10.000$  –  $M = 2000$ ,  $NFFT = 1024$ . Choix de fenêtre = RECTWIN – Recouvrement de 50%



$N = 10.000$  –  $M = 2000$ ,  $NFFT = 1024$ . Choix de fenêtre = HAMMING – Recouvrement de 50%

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

L'utilisation d'une autre fenêtre que la fenêtre rectangulaire permet de réduire le biais et la variance (surtout le biais). En effet, la largeur des lobes principaux et secondaires permet de jouer sur les fréquences plus ou moins coupées de façon à pouvoir se recentrer autour du problème souhaité.

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

Longueur de la séquence analysée  $N = 25.000$  La qualité du résultat augmente avec la longueur de la séquence.

Longueur des tranches  $M = 200$  De petites tranches améliorent la variance

Type de fenêtre : Hamming

Taux de recouvrement = 50%

Nombre de points de transformée de Fourier  $NFFT = 2048$  La qualité augmente avec le nombre de points.

## 5 Utilisation des estimateurs précédents pour analyser un signal inconnu

### 5.1 Modification des programmes

Les fonctions ont été modifiées en retirant la partie création de la partie théorique car le signal étudié n'est plus un bruit blanc mais un signal inconnu fourni sur CPe-campus.

L'autre modification est que nous avons ajouté des paramètres de retours et retiré la partie affichage interne à chaque fonction. Ainsi, nous avons centralisé l'affichage dans le code main et utilisé la fonction *semilogy* pour obtenir une étude plus ergonomique. Ci-dessous le code de la fonction *ESS.c*, l'estimateur de Welch.

Voici le script modifié de l'estimateur de Welch. Là encore, il est tronqué de la partie affichage.

```
function [gamma_x_c,f_abs] = ESW2(x,N,fenetre_char,M,NOVERLAP,NFFT)

% ---Initialisation des variables ---
    x_seq = x(1:N);
    eval(['WIN=',fenetre_char,'(M)']);
    f_abs = 0:1/NFFT:1-1/NFFT;

% ---Création de l'estimateur 3---
    [gamma_x_c,fabs] = pwelch(x_seq,WIN,NOVERLAP,NFFT,1,'twosided');
end
```

## 5.2 Expérimentation

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

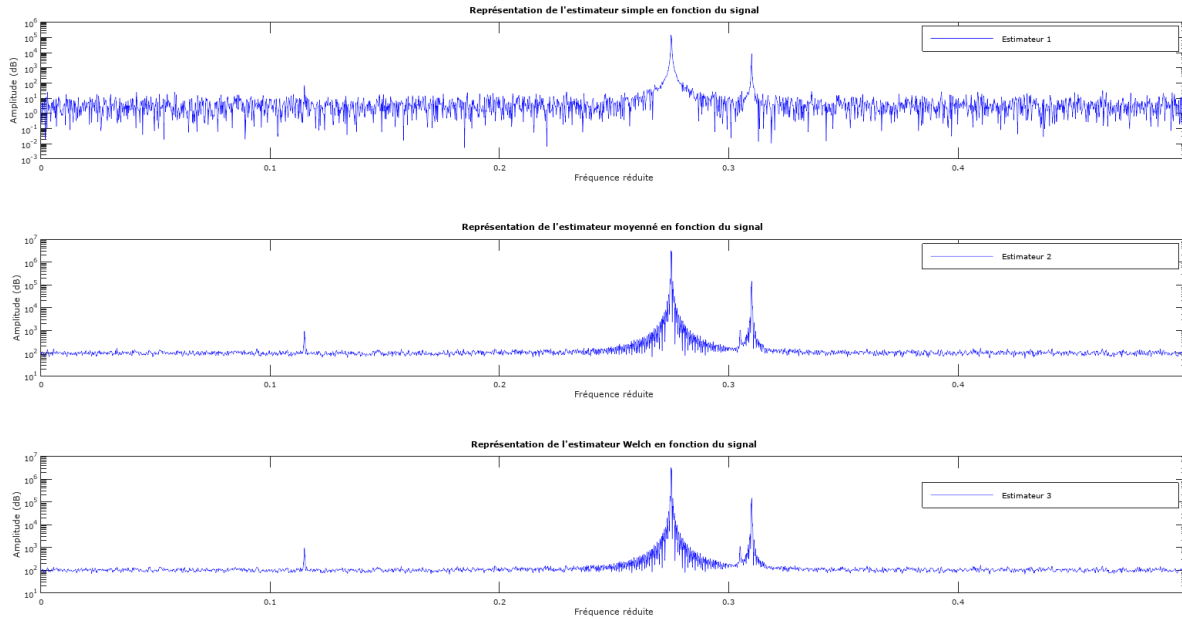


FIGURE 14 – Estimateurs spectraux simple, moyenné et de Welch avec  $N = 100.000 - M = 2500$ ,  $NFFT = 1024$ . Choix de fenêtre = RECTWIN – Recouvrement de 50%

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

Dans un premier temps, nous utilisons l'estimateur simple car bien qu'en partie 1, nous avons conclu qu'il est peu fiable au niveau du compromis biais-variance, l'étude ne prenait pas en compte les propriétés de résolution fréquentielle. Dans son cas, la formule est :  $R_{frequ} = 1.28/2N$  or, pour un  $N$  de 100.000 nous obtenons une résolution précise et nous pouvons étudier facilement les éléments qui compose le spectre.

Cette première étude va permettre d'ajuster les paramètres des Estimateurs Moyenné et de Welch. Nous savons que la résolution fréquentielle est obtenue à l'aide de :  $R_{frequ} = 1.28/2M$ . Or, avec la première étude, il est apparu que deux composantes sont proches donc la taille de la fenêtre devra être grande pour obtenir une résolution fréquentielle suffisante.

Dans notre cas, nous avons choisi empiriquement une fenêtre d'une taille de 1000 et une transformée de Fourier de 1024 points.

Le soucis est que cette configuration ne fonctionne pas dans le cadre de l'estimateur Welch car sinon la Densité Spectrale "perd" une composante car elle est "absorbée" par une autre. Ainsi, nous sommes contraints de doubler la taille de la fenêtre  $M$ .

### 5.2.1 Interprétations

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

L'inconvénient majeur de l'utilisation d'une fenêtre est l'ajout de nouveaux termes dans le domaine fréquentiel. En effet, pour prendre la fenêtre rectangulaire pour exemple, cela revient à convoluer la transformée de Fourier de notre signal par un sinus cardinal. C'est pour cette raison que nous essayons de choisir une fenêtre se rapprochant au mieux d'un dirac.

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

Nous voyons 4 pics : le premier à environ 0.11, le deuxième à environ 0.27, le troisième à environ 0.30 et le dernier très proche du troisième à environ 0.31.

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

Elle nous apprend que notre signal est une somme de 4 sinusoïdes d'amplitude et de phase différentes car nous observons 4 pics (et nous ne visualisons pas les fréquences réduites négatives).

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

La DSPM représente la répartition de la puissance d'un signal suivant les fréquences. Il donne donc comme information les fréquences (et les amplitudes du signal pour celles-ci) par lesquelles il transmet le plus de puissance, c'est-à-dire les fréquences à couper en priorité dans le cas d'un bruit que l'on veut supprimer.