

# Traitement des Signaux Aléatoires

## Estimation Spectrale

4 ETI – CPE Lyon

Travaux Pratiques TSA

Noms, Prénoms: BURNOT Jean-Christophe, LANGUILLE Antoine

Groupe: D

Date: 29 Novembre 2022

### 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

### Notes post scriptum

Une erreur s'est glissée dans les graphiques des simulations, les légendes "théoriques" et "biaisées" ont été inversées.

Les tracés théoriques sont alors en couleur en orange / rouge et les tracés biaisés, prenant en compte la limitation temporelle apparaissent en jaune.

La simulation théorique repose sur une analyse d'une infinité de réalisations de signaux aléatoires d'une longueur infinie, son atténuation en bande atténuée tends vers l'infini (tracés rouge). En revanche, la limitation temporelle entraîne la périodisation en fréquence du signal, cette atténuation idéale est alors inatteignable et l'atténuation stagne alors autour d'une valeur en bande atténuée (tracés jaune).

## 1 Préparation

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

réponse

Pour un signal d'énergie finie, la Densité Spectrale d'Énergie est égale à

$$\Gamma_x(\nu) = E(\nu)^* E(\nu)$$

□

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

\_\_\_\_\_ réponse \_\_\_\_\_

La DSPM d'un processus aléatoire est défini grâce au théorème de Wiener-Kitchine

$$\Gamma_s(\nu) = TF(\gamma_s(\tau))$$

□

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

On chiffre la qualité d'une estimation grâce à 4 grandeurs

- Le biais d'estimation
- La variance de l'estimateur
- La précision d'estimation
- La résolution fréquentielle

□

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

\_\_\_\_\_ réponse \_\_\_\_\_

Un bruit blanc est à corrélation microscopique, sa fonction d'autocorrélation  $\gamma_B(\tau)$  se résume donc à un Dirac, ce qui dans le domaine fréquentiel se traduit par une constante.

De plus on a  $\gamma_s(\tau) = \bar{P}_X$ .

Ainsi

$$\Gamma_B(f) = \Gamma_0 = \bar{P}_X$$

□

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

$$\begin{aligned}\Gamma_X(f) &= TF(\gamma_X(\tau)) \\ &= TF((\gamma_B * h)(\tau)) \\ &= TF((\Gamma_0 \cdot \delta * h)(\tau)) \\ &= \Gamma_0 \cdot 1 \cdot H(f)\end{aligned}$$

□

**Question 6** En une phrase (sans formule), décrire le procédé de calcul de la DSPM estimée  $G1(f)$  d'une séquence aléatoire via l'estimateur simple.

\_\_\_\_\_ réponse \_\_\_\_\_

Avec un estimateur simple, la transformée de Fourier de nos échantillons, puis on l'élève au carré avant de la diviser par le nombre de points. Ce calcul revient à diviser la Densité Spectrale d'Énergie par le nombre d'échantillons. ☐

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

\_\_\_\_\_ réponse \_\_\_\_\_

Une TFD-N points est graduée de 0 à  $1 - \frac{1}{N}$  pour un pas égal à  $\frac{1}{N}$  ☐

**Question 8** Décrire (avec une phrase) le procédé de calcul de la DSPM estimée  $G2(f)$  d'une séquence aléatoire via l'estimateur moyenné.

\_\_\_\_\_ réponse \_\_\_\_\_

On découpe le signal en plusieurs tranches, on calcule la Densité Spectrale d'Énergie de chaque échantillon que l'on divise par le nombre d'échantillons dans chaque tranche. On obtiens alors autant d'estimations de DSPM que de tranches. On moyenne ensuite ces estimations pour obtenir notre estimation finale. ☐

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

\_\_\_\_\_ réponse \_\_\_\_\_

Le compromis biais-variance dans le cas de l'estimateur moyenné caractérise l'évolution inversement proportionnelle du biais et de la variance de l'estimateur en fonction des paramètres (ici la largeur M). ☐

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

\_\_\_\_\_ réponse \_\_\_\_\_

L'estimateur de Welch autorise le recouvrement des bandes de découpages de l'estimateur moyenné. On applique également une fenêtre différente de la porte à ces tranches avant d'en calculer la DSE, pour réduire au maximum les parasites dus à la convolution fréquentielle engendrée par la limitation temporelle. ☐

## 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

---

La génération de nombre pseudo-aléatoires se fait à partir d'une "graine" (seed) qui est par défaut la valeur de l'horloge système à l'exécution. Cette valeur permet le caractère suffisamment aléatoire de la génération de nombres entre deux exécutions.

Fixer ce nombre empêche la variation inter-exécutions, garantissant ainsi un signal aléatoire lors de l'exécution mais identique entre les différentes exécutions.

On peut alors comparer les performances de plusieurs estimateurs, sachant qu'ils estimeront le même signal. □

### 3.2 Estimateur spectral simple

#### 3.2.1 Script de la fonction Matlab développée

---

code ci-dessous

---

```
function [Gamma, f, N] = simpleDSPM(X, nd, nf, NFFT)

ana = X(nd:nf);
N = nf - nd; %Indices de début et de fin de sélection

% On fait la transformé de fourrier N-points
Xf = fft(ana, NFFT);
f = [0 : 1/NFFT : 1-1/NFFT];

% On calcule la DSPM (DSE divisée par nombre d'échantillons)
Gamma = (1/N).*abs(Xf).^2;
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

---

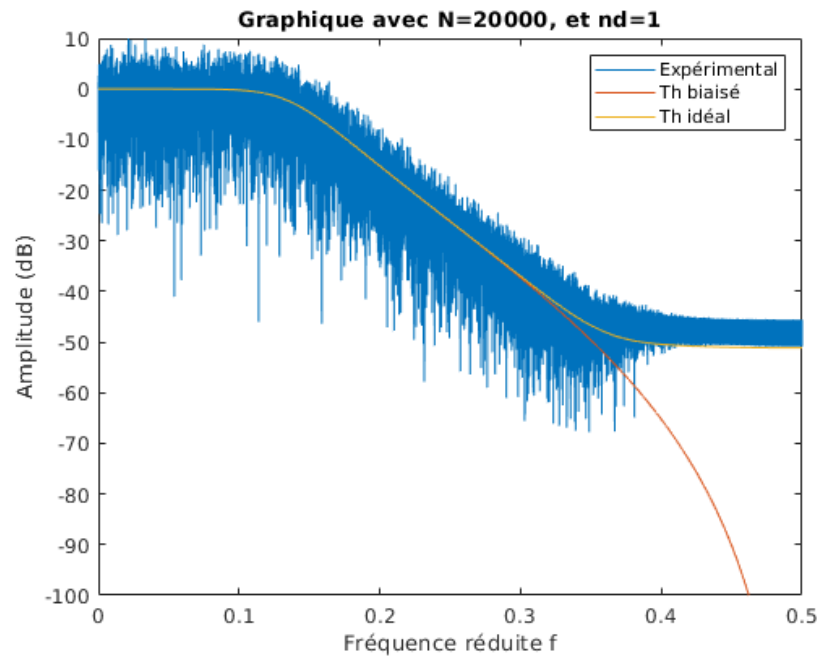


Figure 1:  $N$  faible – indice de début de la séquence à 1

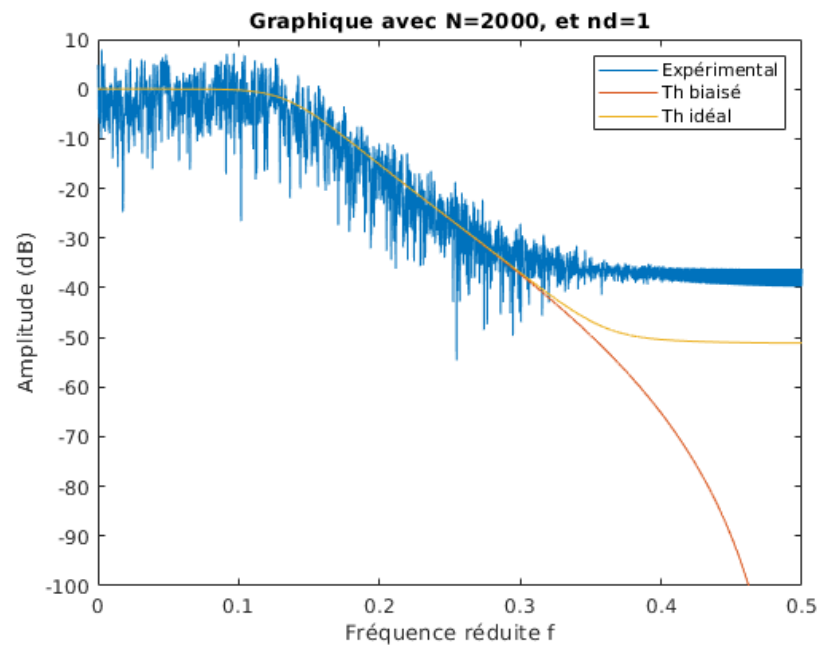


Figure 2:  $N$  élevé – indice de début de la séquence à 1

□

Commentaires.

réponse ci-dessous

Lorsque  $N$  augmente, on observe une diminution de la variance du signal (trace plus "épaisse") mais également une augmentation du biais sous la forme d'écart entre la valeur biaisée et la valeur expérimentale.  $\square$

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

---

figures ci-dessous

---

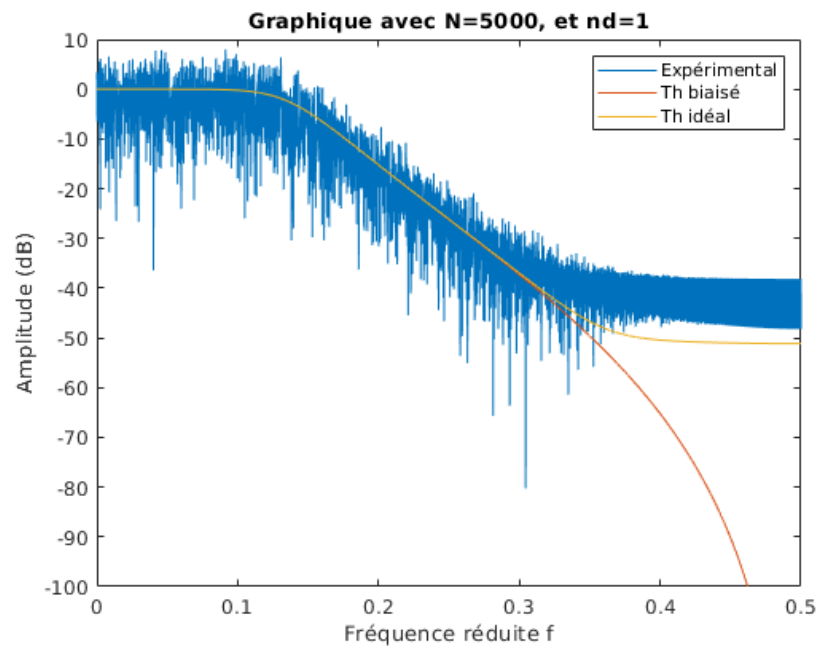


Figure 3:  $N \sim 1000$  – indice de début de la séquence à 1

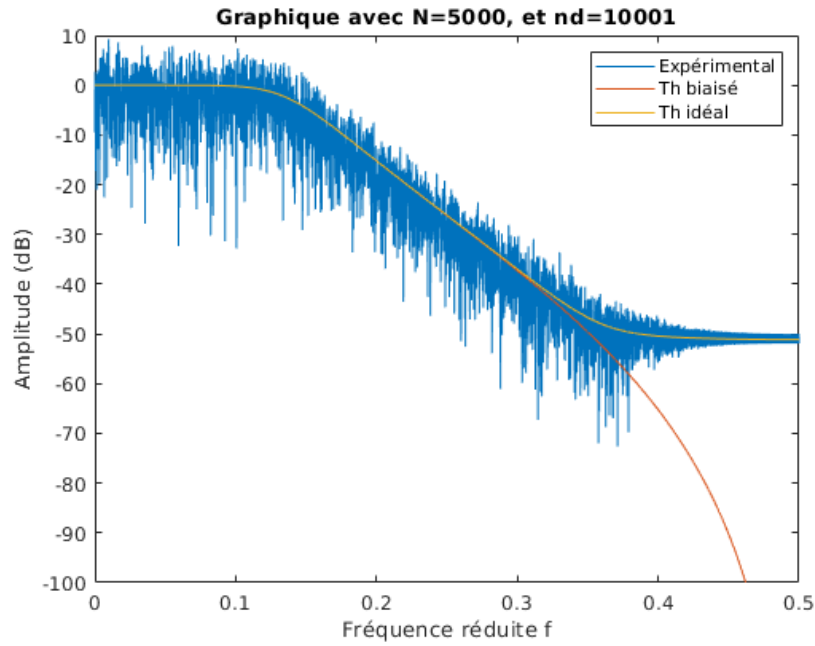


Figure 4:  $N \sim 1000$  – indice de début de la séquence fixé à une autre position ( $\gg 1000$ ) dans la séquence

□

Commentaires.

---

réponse ci-dessous

---

Dans notre deuxième figure, notre réalisation expérimentale est bien plus proche de l'idéal.

On tire comme conclusion que pour avoir une estimation correcte (moins de biais), il est préférable de ne pas démarrer notre séquence à 1.

Cette variation montre la variation statistique entre les différentes parties du signal, soulignant ainsi que le signal ne respecte pas la propriété de stationnarité.

On préférera alors se limiter à un nombre fixé d'échantillons commençant à un indice fixe pour comparer les estimations. □

C. Etude du biais et de la variance en fonction du nombre  $N_{FFT}$  de FFT

---

figures ci-dessous

---

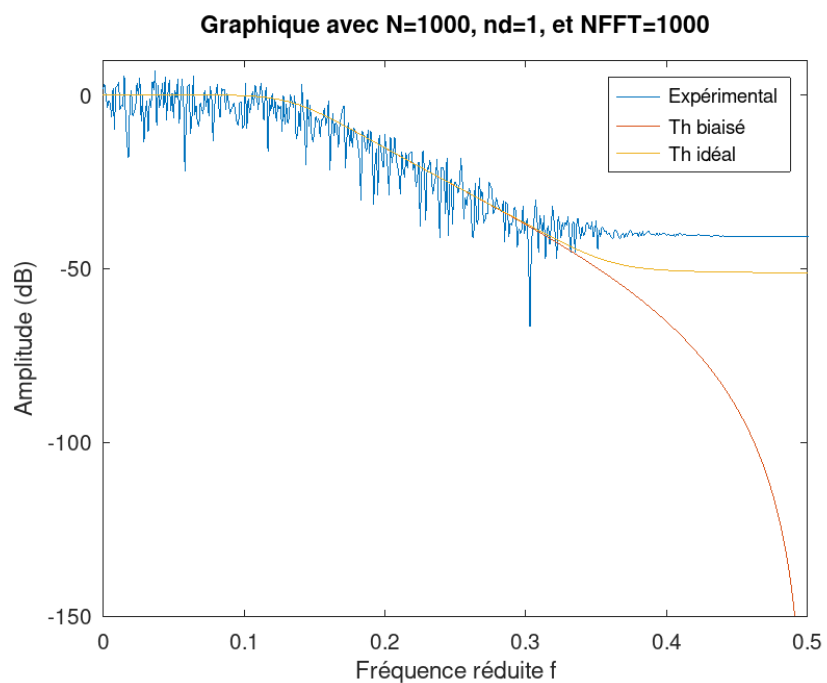


Figure 5:  $N$  constant – indice de début de la séquence à 1 –  $NFFT \sim N$

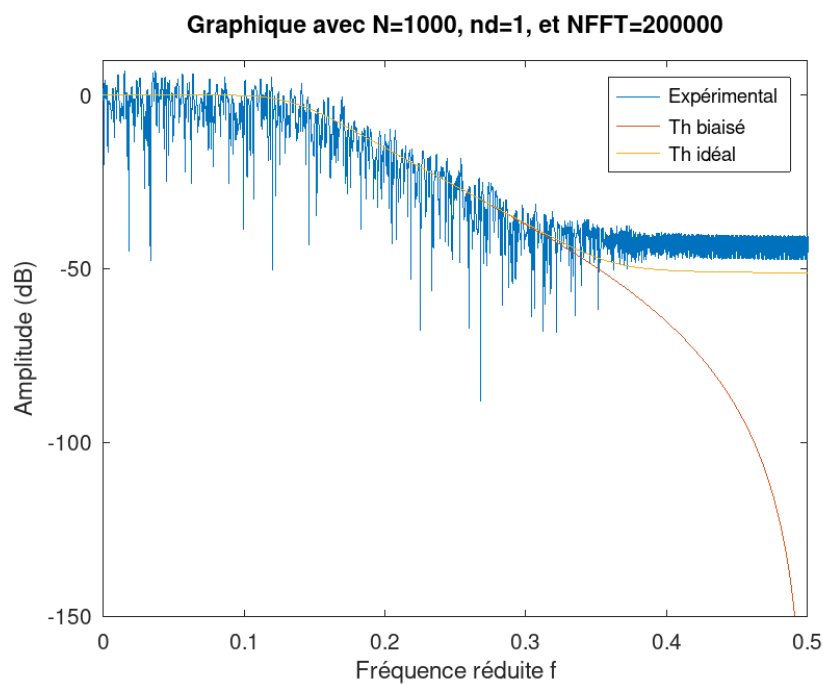


Figure 6:  $N$  constant – indice de début de la séquence à 1 –  $NFFT \gg N$

□

Commentaires.

réponse ci-dessous



On observe une augmentation de la variance du signal lorsque le nombre de points de la FFT devient grand. Il semble qu'avoir  $N$  et  $NFFT$  au même ordre de grandeur nous fournisse la meilleure lisibilité. On ne peut aller en dessous du nombre d'échantillons  $N$ , au risque de perdre de l'information sur la FFT.  $\square$

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

---

**réponse ci-dessous**

---

L'estimateur simple a pour principal problème son biais élevé. Surtout pour les hautes fréquences. Peu importe la grandeur sur laquelle on influe, on constate un biais élevé pour les hautes fréquences de par le caractère fini (Courbe rouge et courbe jaune écartées).  $\square$

### 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 [Gamma, f] = moyenneurDSPM(X, N, M, NFFT)
    ana = X(1:N);
    window = rectwin(M);
    [Gamma, f] = pwelch(ana,window,0,NFFT, 1, "twosided");
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

---

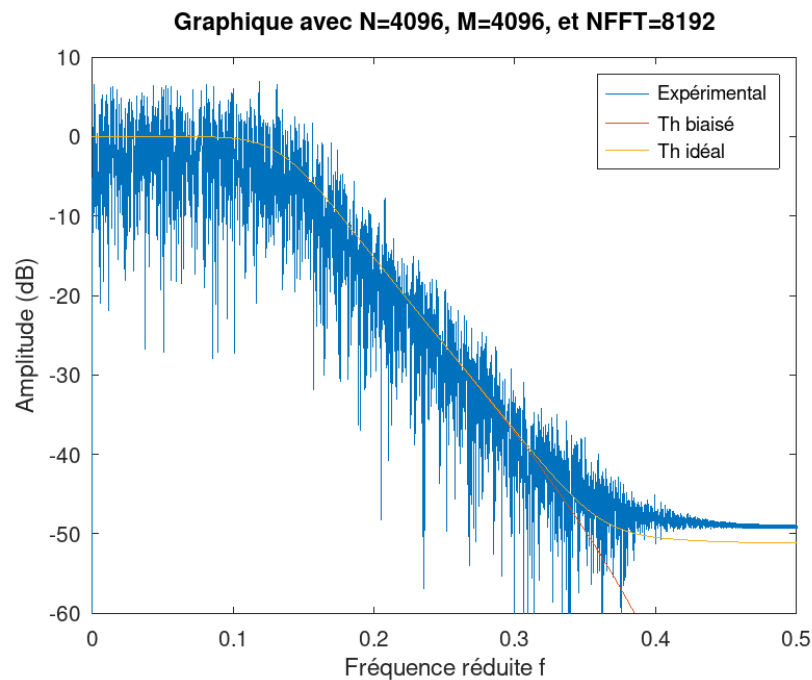


Figure 7:  $N = 4096$  – tranches courtes  $M = 4096$ ,  $NFFT = 8192$

□

Commentaires

---

réponse ci-dessous

---

L'utilisation de tranches courtes ( $M$  petit) c'est à dire d'un nombre de segments  $L$  élevés. A pour effet d'augmenter la résolution fréquentielle. Le biais est diminué par rapport à l'estimateur simple mais la variance est plus élevée.  $\square$

---

figure ci-dessous

---

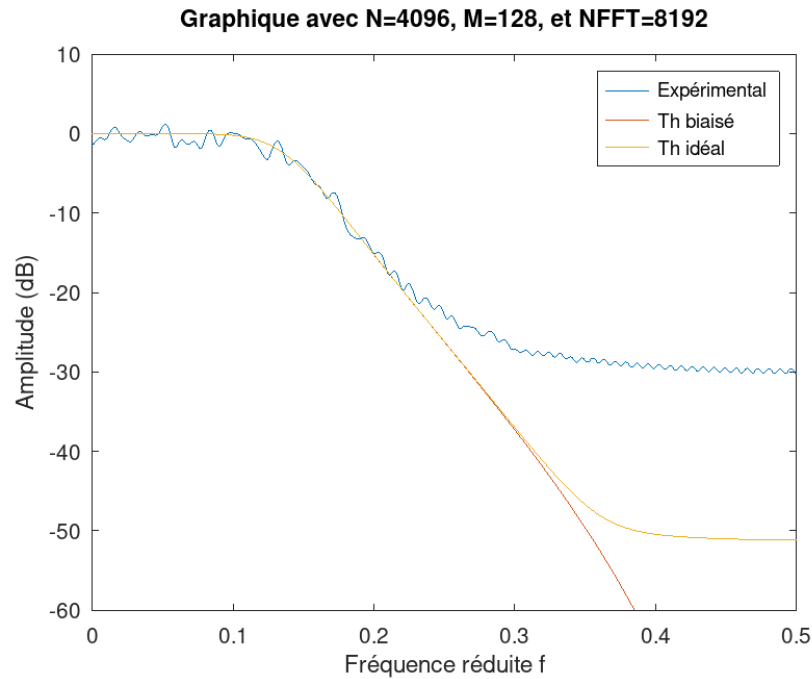


Figure 8:  $N = 4096$  – tranches longues  $M = 64$ ,  $NFFT = 128$

$\square$

Commentaires

---

réponse ci-dessous

---

L'utilisation de tranches longues ( $M$  grand) c'est à dire d'un nombre de segments  $L$  faible. A pour effet de diminuer la résolution fréquentielle. La variance est moins importante mais le biais augmente significativement.  $\square$

---

figure ci-dessous

---

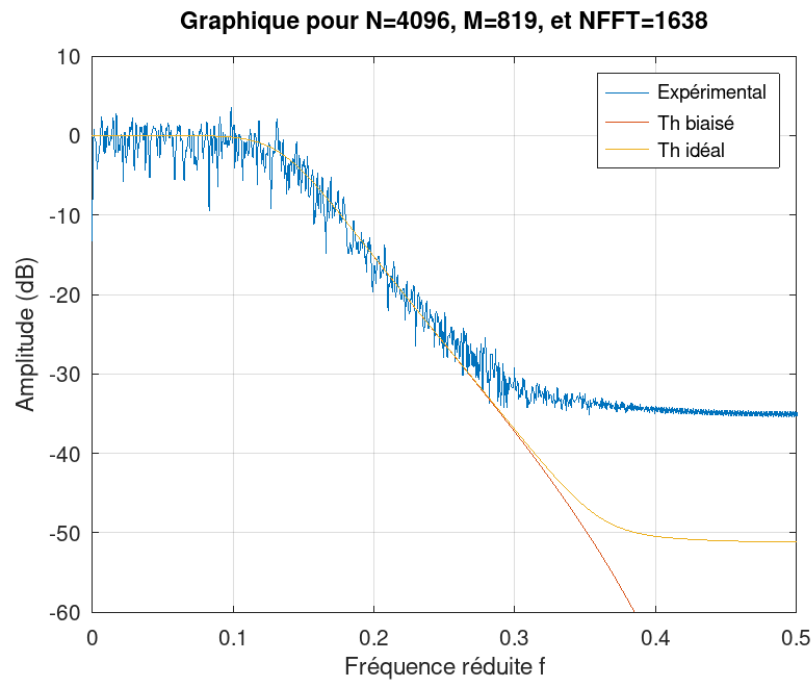


Figure 9:  $N = 4096$  – Meilleur compromis biais variance atteint pour  $M = 819$ ,  $NFFT = 1638$

□

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

**réponse ci-dessous**

Le meilleurs compromis biais-variance peut être atteint en connaissant la valeur de l'écart-type du signal selon la formule:

$$L = \Delta_{x_{opt}} \approx 3.49 \hat{\sigma}_x N^{-\frac{1}{3}}$$

□

## 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 [Gamma, f] = welchDSPM(X, N, Nom_fenetre, M, NOVERLAP, NFFT)
    ana = X(1:N);
    WIN = eval([Nom_fenetre, "(M);"]);
    [Gamma, f] = pwelch(ana, WIN, NOVERLAP, NFFT, 1, "twosided");
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 \_\_\_\_\_

**$N=4096$ ,  $M=819$ , et  $NFFT=1638$ , Fenetre hamming, recouvrement 0%**

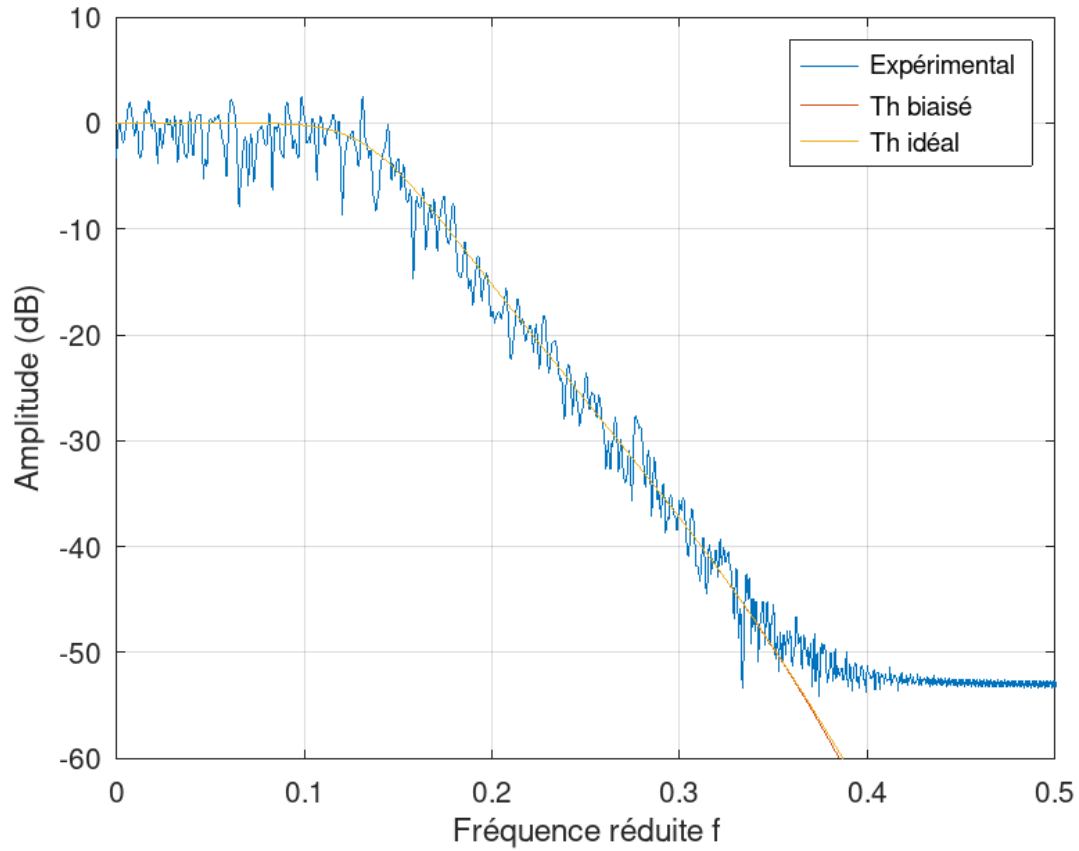


Figure 10:  $N = 4096 - M = 819$ ,  $NFFT = 1638$ . Choix de fenêtrage = Hamming – Recouvrement de 0%

**N=4096, M=819, et NFFT=1638, Fenetre hamming, recouvrement 50%**

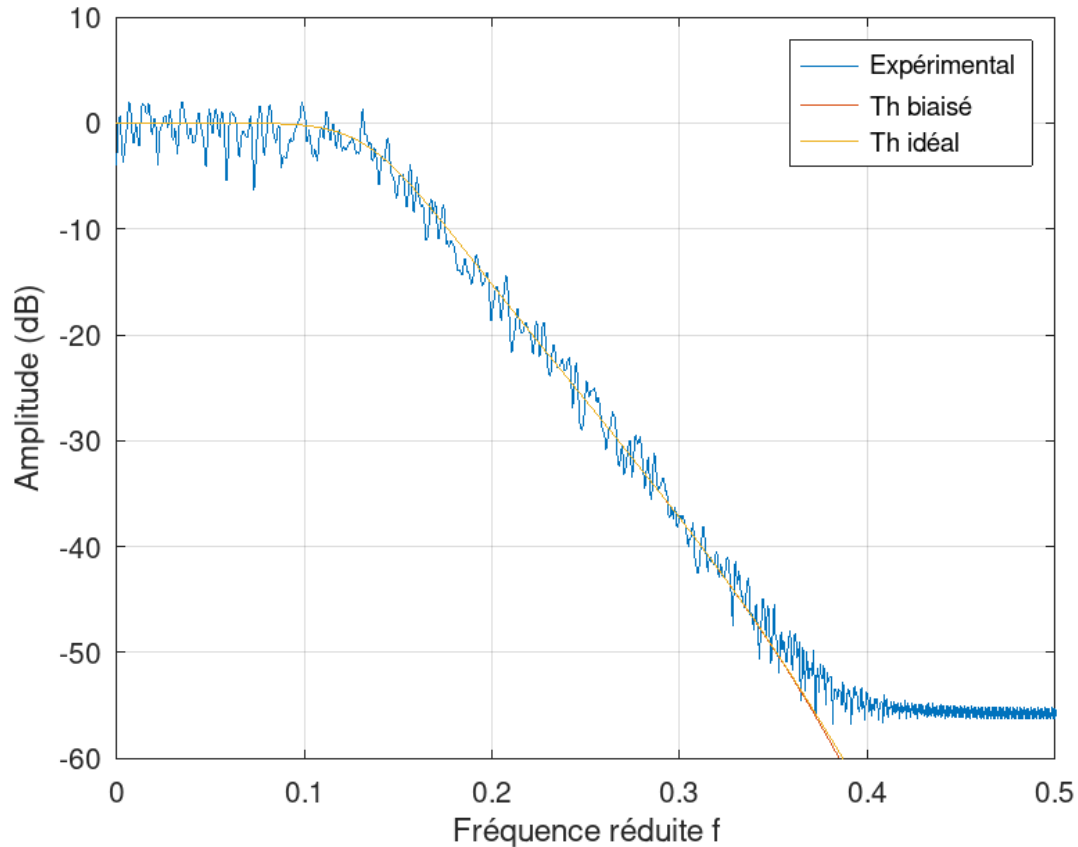


Figure 11:  $N = 4096 - M = 819$ ,  $NFFT = 1638$ . Choix de fenetre = Hamming – Recouvrement de 50%

□

Que permet le recouvrement entre tranches ?

réponse ci-dessous

Le recouvrement permet d'augmenter le nombre d'échantillons de chaque tranche (et donc la précision de la FFT), sans pour autant réduire le nombre de tranches. Pour une fenetre de Hamming, on observe une légère réduction de la variance en bande passante ainsi qu'une légère diminution du biais en bande atténuée.

On observe cependant une grande amélioration de variance et biais par rapport à la fenetre rectangulaire de l'estimateur moyenné. □

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

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

figure ci-dessous

**N=4096, M=819, et NFFT=1638, Fenetre bartlett, recouvrement 50%**

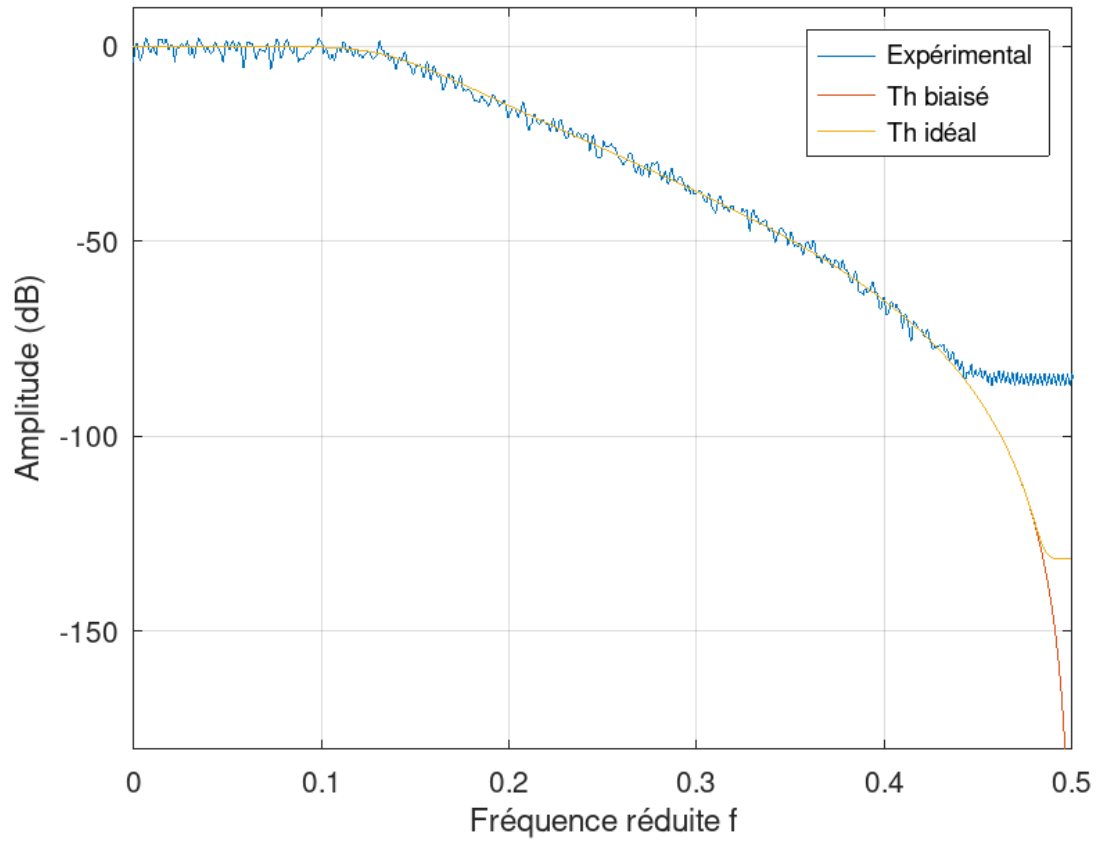


Figure 12:  $N = 4096 - M = 819$ ,  $NFFT = 1638$ . Fenêtre trianngle – Recouvrement de 50%



**N=4096, M=819, et NFFT=1638, Fenetre gausswin, recouvrement 50%**

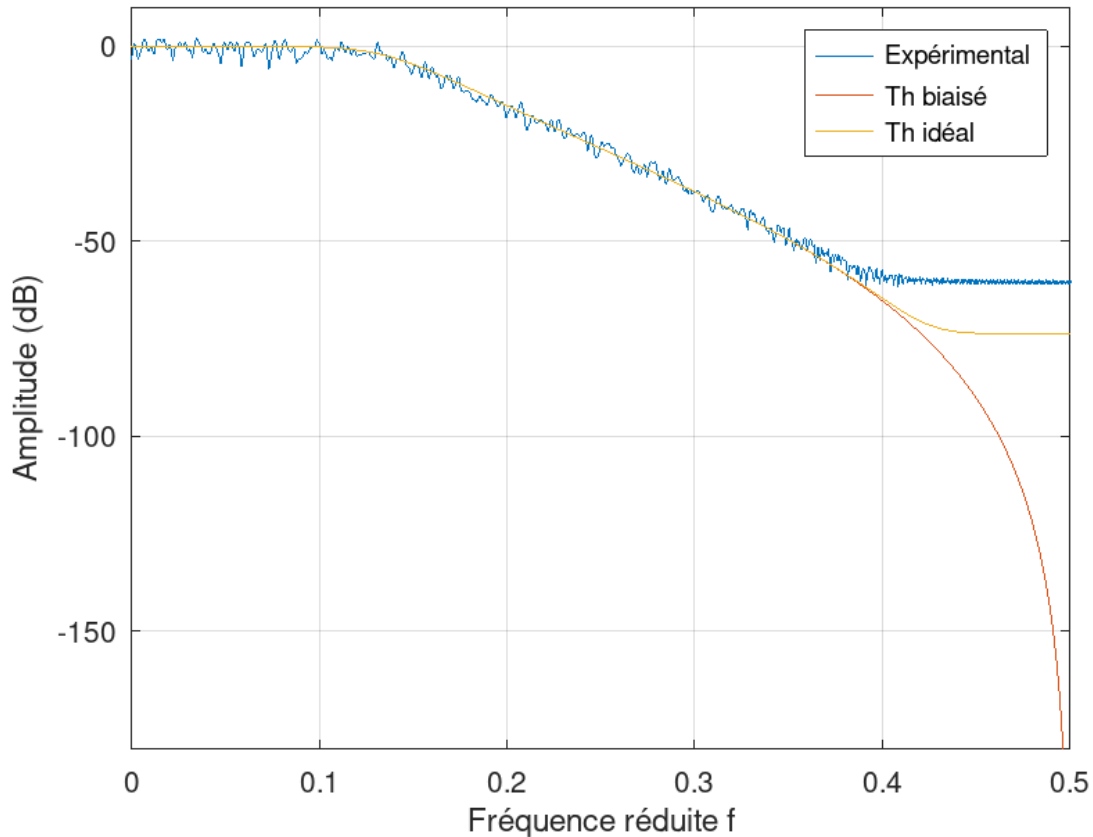


Figure 13:  $N = 4096 - M = 819$ ,  $NFFT = 1638$ . Fenêtre Gaussienne – Recouvrement de 50%

□

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

réponse ci-dessous

Une fenêtre autre que rectangulaire permet d'éviter la convolution par un sinus cardinal dans le domaine spectral, augmentant alors la précision de l'estimation de la DSPM.

En observant la densité spectrale d'énergie, on peut aisément déduire que le spectre se rapprochant le plus d'un Dirac (élément neutre de la convolution) est l'estimateur dont la limitation temporelle affecte le moins la DSPM du signal.

On observe alors une amélioration des caractéristiques telles que la variance ou l'écart-type mais une diminution de la résolution fréquentielle, la fenêtre rectangulaire ayant le "pic" le plus fin sur sa DSPM. □

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  $N = 20000$

Longueur des tranches  $M = 0.2 * N = 4000$

Type de fenêtre Hanning  
Taux de recouvrement = 50%  
Nombre de points de transformée de Fourier  $NFFT = 2M = 8000$

□

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

### 5.1 Modification des programmes

Script d'une des fonctions modifiée

---

code ci-dessous

---

```
close all;
clear variables;
clc;
%% Pour charger la bibliothèque signal
pkg load signal;

%% Charge le vecteur S
load sig

%% Première fonction
N = 100000;
NFFT = 2^17;
figure()
%appel de la fonction simple
[Gamma1,VecteurFreq1, N] = simpleDSPM(s,1,N, NFFT);
semilogy(VecteurFreq1,Gamma1)
axis([0 0.5 10 10^7])
figure()

%% Deuxième fonction
N = 100000;
M = 5000;
NFFT = 2^17;
[Gamma2, VecteurFreq2] = moyenneurDSPM(s, N, M, NFFT);
semilogy(VecteurFreq2,Gamma2)
axis([0 0.5 10 10^7])

%% Troisième fonction
N = 100000;
M = 5000;
NOVERLAP = 0.5;
NFFT = 2^17;
[Gamma3,VecteurFreq3] = welchDSPM(s,N,'hanning',M,NOVERLAP,NFFT);
figure()
semilogy(VecteurFreq3,Gamma3)
axis([0 0.5 10 10^7])
```

□

## 5.2 Expérimentation

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

[figure ci-dessous](#)

Figure 14: Estimateur spectral simple.  $N = 100000$ ,  $NFFT = 2^{17}$ .

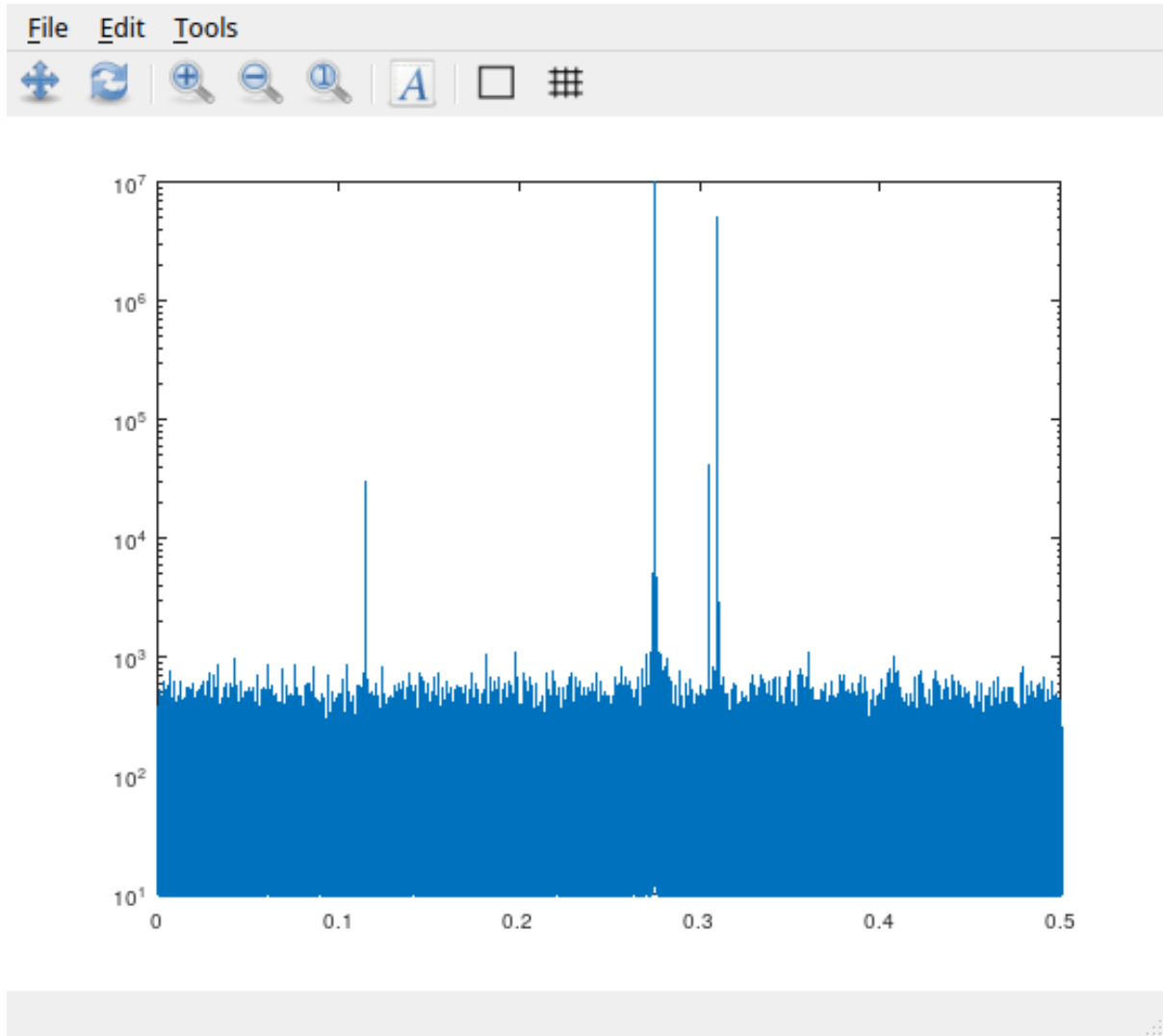


Figure 15: Estimateur spectral moyenné.  $N = 100000$ ,  $NFFT = 2^{17}$ ,  $M = 5000$ .

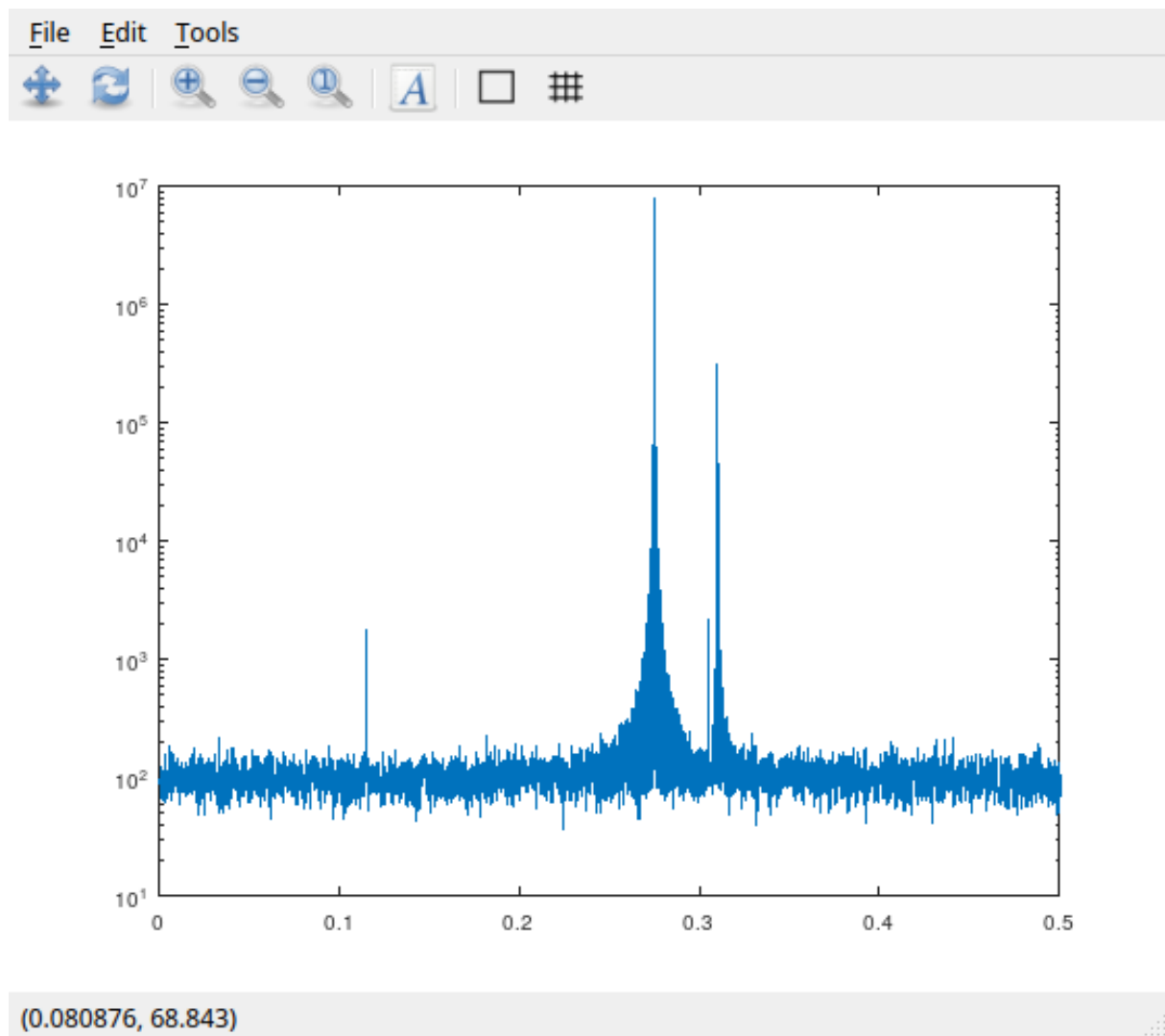
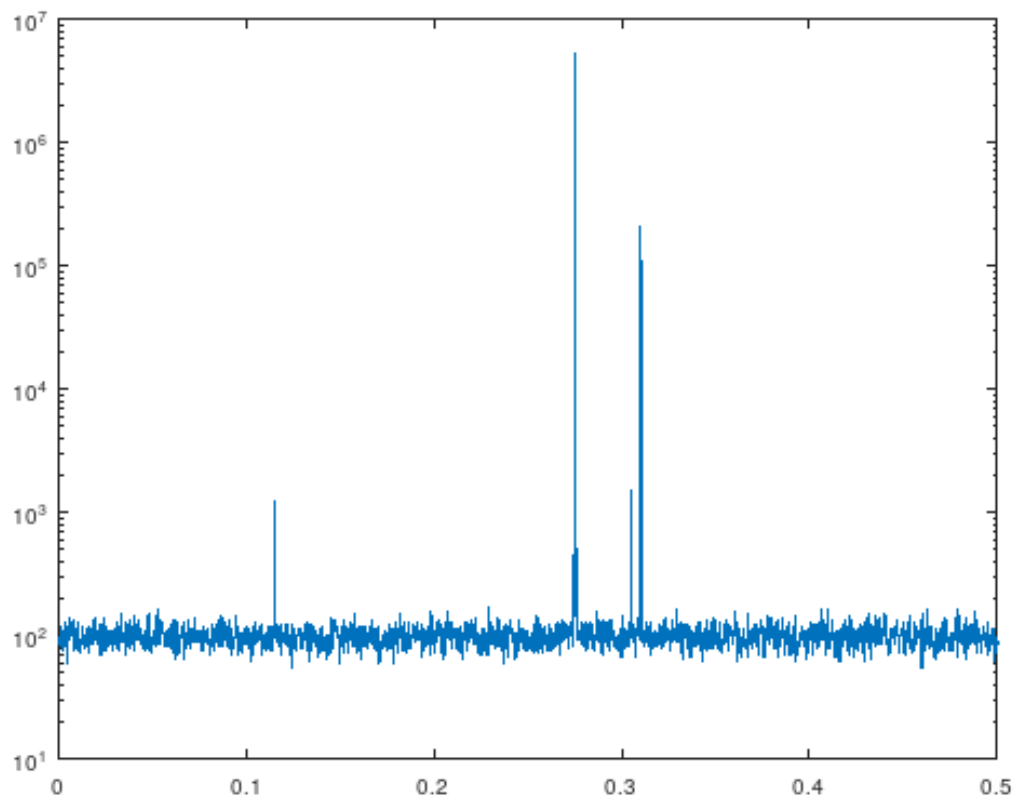
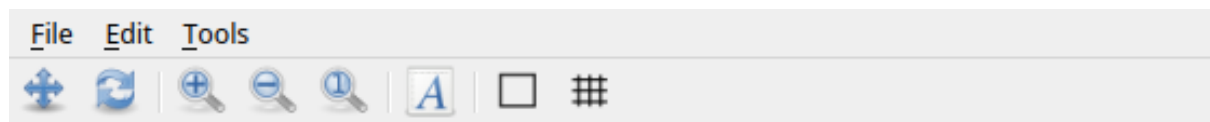


Figure 16: Estimateur spectral de Welch.  $N = 100000$ ,  $NFFT = 2^{17}$ ,  $M = 5000$ , fenetre de hanning, overlap de 0,5%.



(0.2076, 53711)



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

La méthode la plus adéquate pour faire l'étude du signal est celle de welsh, c'est celle qui est la moins bruitée



### 5.2.1 Interprétations

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

réponse ci-dessous

Elle diminue la résolution fréquentielle et engendre plus de calculs.



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

\_\_\_\_\_ **réponse ci-dessous** \_\_\_\_\_

On a 0,275 pour le premier pic

0,31 pour le second pic

0,115 pour le troisième

0,305 pour quatrième

☐

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

\_\_\_\_\_ **réponse ci-dessous** \_\_\_\_\_

La DSPM permet d'avoir des information sur les fréquences du signal et sur la forme du bruit ☐

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

\_\_\_\_\_ **réponse ci-dessous** \_\_\_\_\_

On peut savoir où sont les fréquences du signal pour faire un filtrage adéquat

☐