Traitement des Signaux Aléatoires
Estimation de densités de probabilité

4 ETI – CPE Lyon

Travaux Pratiques TSA

Noms, Prénoms : Beaucamp Elisa, Deschamps Corto

Groupe : C

Date : Lundi 4 octobre

Objectifs du TP

- Synthèse et filtrage de processus aléatoires
- Estimation empirique de densités de probabilités de différents processus aléatoires
- Filtrage passe-bas de processus non gaussiens.

Consignes :

- Le répertoire de travail sera exclusivement sur le compte d'un des membres du binôme (changer le répertoire courant de Matlab®). Mais pour certains traitements, on fera appel à des fonctions pré-programmées. Les fonctions utiles sont accessibles sur CPe-campus dans le cours `Traitement des signaux aléatoires`, rubrique `Travaux Pratiques`. Récupérer les fichiers `.m`.
- Utiliser la trame de `compte-rendu` fournie en répondant directement aux questions dans les espaces ménagés à cet effet.
- Regrouper dans un fichier annexe (type `word` ou `text`) les Codes Matlab® développés ainsi que les Figures obtenues. **Veiller à associer systématiquement une légende explicite à chaque Figure ou Tableau**.
- **Préparation obligatoire** (une seule par binôme) à rédiger directement sur le `compte-rendu` et à fournir en début de séance

1 Préparation

Il faudra avoir pris connaissance de la totalité de l'énoncé et de la documentation des diverses fonctions Matlab fournie en Annexe.

Pour estimer la densité de probabilité d'un signal aléatoire $\mathbf{x}$, on s'appuie ici sur l'histogramme d'une seule réalisation échantillonnée du signal aléatoire. Soient $(x[n] = x(n \cdot T_s))_{n=1,\ldots,N}$, la série temporelle correspondante échantillonnée à la fréquence $F_s = T_S^{-1}$:

$$\widehat{p_{\mathbf{x}}}(x) = \frac{\text{Nbre d'échantillons compris dans l'intervalle } \left[x - \frac{\Delta x}{2}, x + \frac{\Delta x}{2}\right]}{N\,\Delta x}.$$

Question 1 Quelles propriétés le signal aléatoire $\mathbf{x}$ doit il vérifier : :

- pour que les échantillons $(x[n])_{n=1,\dots N}$ soient identiquement distribués (i.e. suivent tous la même loi, quelque soit l'instant n)

réponse ci-dessous

Il faut que le signal soit stationnaire ☐

- pour que les échantillons $(x[n])_{n=1,\dots N}$ soient décorrélés ?

réponse ci-dessous

Il faut que E(x[1]x([2])) = 0 et ce pour tous les x[n] deux à deux ☐

- pour que la décorrélation des échantillons $(x[n])_{n=1,\dots N}$ entraine également leur indépendance ?

réponse ci-dessous

On veut ici que E(x[1]x[2])=E(x[1])E(x[2]) et ce pour tous les x[n] deux à deux ☐

Question 2 Sans calculs, indiquer quelle est l'influence du choix de Δx sur le biais et sur la variance de l'estimation $\widehat{p_{\mathbf{x}}}(x)$.

réponse ci-dessous

Lorsque Δx diminue, le biais diminue et la variance augmente. L'inverse est vrai aussi. Il faut utiliser un compromis bais variance pour trouver le bon équilibre. ☐

Question 3 Quelles opérations (arithmétiques simples, il ne s'agit pas de filtrage ici !) permettent de synthétiser un processus gaussien de moyenne m_2 et d'écart-type σ_2 à partir d'un processus gaussien stationnaire de moyenne $m_1 \neq m_2$ et d'écart-type $\sigma_1 \neq \sigma_2$?

réponse ci-dessous

Soit X notre variable aléatoire suivant la loi N(m1,σ1). Soit Y notre nouvelle variable aléatoire. Dans un premier temps, il faut centrer réduire la variable Y grace à la formule Y = $\frac{X-m1}{\sigma 1}$. Puis nous devons remultiplier et additionner Y avec le nouvel écart-type et la nouvelle moyenne. Y = Y.σ2 + m2. Ainsi, on a Y suit la loi N(m2 ;σ2). ☐

Question 4 Le **Kurtosis** est un indice qui permet de mesurer le caractère normal (gaussien) d'une série d'échantillons d'une variable aléatoire. Il est défini par le rapport : $K = \dfrac{\mathbb{E}\left\{\mathbf{x}^4\right\}}{\mathbb{E}^2\left\{\mathbf{x}^2\right\}}$

On rappelle que si $\mathbf{x}$ est gaussien et centré, alors

$$\begin{aligned}\mathbb{E}\{x(t_1)x(t_2)x(t_3)x(t_4)\} &= \mathbb{E}\{x(t_1)x(t_2)\}\,\mathbb{E}\{x(t_3)x(t_4)\} + \mathbb{E}\{x(t_1)x(t_3)\}\,\mathbb{E}\{x(t_2)x(t_4)\}\dots \\ &\quad + \mathbb{E}\{x(t_1)x(t_4)\}\,\mathbb{E}\{x(t_2)x(t_3)\}\end{aligned}$$

Montrer alors que dans le cas d'un signal aléatoire gaussien, centré et stationnaire, le Kurtosis vaut 3.

réponse ci-dessous

Pour que l'intégrale de la dsp vale 1, on a un signal gaussien centré réduit et stationnaire. On utilise le théorème de transfert :

$$K = \hat{m_4} = E((\frac{X-\mu}{\sigma})^4 = \frac{1}{\sigma^4}E((X-\mu)^4) = E(X^4)$$

car $\mu = 0$ et $\sigma = 1$ par centrage réduction. On obtient alors :

$$K = \int_{-\infty}^{\infty} x^4 \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^3 x e^{\frac{-x^2}{2}}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}} [-x^3 e^{-\frac{x^2}{2}}]_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 3x^2 e^{\frac{-x^2}{2}}\,\mathrm{d}x$$

$$= \frac{3}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x x e^{\frac{-x^2}{2}}\,\mathrm{d}x = \frac{3}{\sqrt{2\pi}} [e^{-\frac{x^2}{2}}]_{-\infty}^{\infty} + \frac{3}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}}\,\mathrm{d}x = 3.\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}}\,\mathrm{d}x$$

On obtient 3 fois l'intégrale de la ddp d'une gaussienne. Or l'intégrale d'une ddp vaut 1. Donc finalement, on a :

$$K = 3$$

□

Question 5 Soit $\mathbf{x}(t)$ un bruit gaussien de valeur moyenne m_B et d'écart-type σ_B.
Soit $\mathbf{y}(t)$ un signal carré d'amplitude A, centré, périodique de période T_0, de rapport cyclique égal à 1 et retardé par rapport à l'origine d'un retard τ uniformément distribué entre 0 et T_0.
Donner l'expression de la densité de probabilité de la somme $\mathbf{z}(t) = \mathbf{x}(t) + \mathbf{y}(t)$.

réponse ci-dessous

On a $f_{U+V}(x) = f_U(x) * f_V(x)$Donc

$$p_Z(x) = p_X(x) * p_Y(x)$$

On sait que

$$p_X(x) = \frac{1}{\sigma_B\sqrt{2\pi}} e^{-\frac{1}{2\sigma_B^2}(x-m_B)^2}$$

Et $p_Y(x)$ est représenté par deux demis diracs en A et -A. Donc

$$p_Z(x) = \frac{1}{\sigma_B\sqrt{2\pi}} e^{-\frac{1}{2\sigma_B^2}(x-m_B)^2} * \frac{1}{2}\delta(t-A) + \frac{1}{\sigma_B\sqrt{2\pi}} e^{-\frac{1}{2\sigma_B^2}(x-m_B)^2} * \frac{1}{2}\delta(t+A)$$

Ainsi, $p_Z(x)$ est de la forme de deux gaussiennes, en $m_B - a$ et $m_B + A$, dont la hauteur est divisée par deux par rapport à la gaussienne d'origine. □

Traitement des Signaux Aléatoires
Estimation de densités de probabilité **4 ETI – CPE Lyon**

Travaux Pratiques TSA

Noms, Prénoms : Beaucamp Elisa, Deschamps Corto

Groupe : C

Date : Lundi 4 octobre

2 Bruit gaussien filtré, échantillonné

On souhaite générer un bruit gaussien $x_3(t)$ blanc dans la bande $[-B, B]$, de moyenne m_3 non nulle et d'écart-type $\sigma_3 > 1$. Pour cela, on applique la procédure décrite dans la préparation (Question 3) et schématisée ci-dessous :

où $x_1(t)$ est un bruit blanc gaussien, centré, d'écart-type $\sigma_1 = 1$.

2.1 Programmation

Programmer deux fonctions Matlab distinctes dont vous reproduirez les codes ci-dessous.

2.1.1 Fonction synthèse des signaux aléatoires

- Paramètres d'entrée :
  - le nombre N d'échantillons à générer
  - la largeur de bande B du filtre passe-bas
  - la moyenne m_3 et l'écart-type σ_3 du bruit $x_3(t)$.
- Traitements à effectuer dans la fonction :
  - génération d'une séquence $x_1(t)$ de bruit gaussien échantillonné (à la fréquence F_s), centré et d'écart-type $\sigma_1 = 1$
  - synthèse d'un filtre de *Butterworth* de type passe-bas, de fréquence de coupure f_c correspondant à la largeur de bande B et d'ordre $m = 8$
  - filtrage du bruit $x_1(t)$ par le filtre passe-bas pour obtenir le bruit filtré $x_2(t)$
  - transformation de $x_2(t)$ pour obtenir $x_3(t)$ de valeur moyenne m_3 et d'écart-type σ_3.
- Variables de sortie :
  - les vecteurs des échantillons de x_1, x_2 et x_3
  - les coefficients de la fonction de transfert du filtre passe-bas (coefficients des polynômes $A(z)$ et $B(z)$).

code ci-dessous

```
function [x1, x2, x3, Az, Bz] = synthese(N, B, m3, sigma3)

m = 8;
Fs = 1000;

[Bz, Az] = butter(m, (2*B)/Fs);
```

```
x1 = randn(1, N);
x2 = filter(Bz, Az, x1);
m2 = mean(x2);
x3 = ((x2 - m2)./std(x2)).*sigma3 + m3;

end
```

□

2.1.2 Fonction Calcul d'histogramme

- Paramètres d'entrée :
  - le vecteur des N échantillons d'un signal aléatoire $x(t)$
  - paramètre **optionnel** : M le nombre d'intervalles imposés pour le calcul de l'histogramme
- Traitements à effectuer :
  - si le nombre d'intervalles M n'est pas spécifié :
    - appliquer la règle empirique de calcul *optimal* de Δx (vue en TD)
    - calculer le centre de chaque intervalle de l'histogramme correspondant à ce choix de Δx
    - calculer l'histogramme correspondant
  - si le nombre d'intervalles M est spécifié :
    - déterminer la largeur des intervalles Δx correspondant à ce choix de M
    - calculer l'histogramme correspondant
  - déduire de l'histogramme calculé une estimation de la densité de probabilité de $\mathbf{x}$
  - afficher dans la figure et le graphe courants la densité de probabilité estimée
  - labéliser les axes en indiquant la valeur de Δx utilisée (et préciser si celle-ci est *optimale* ou *imposée*). Donner un titre pertinent (distinctif) au graphe.
- Variables de sortie :
  - le vecteur des valeurs de la densité de probabilité estimée
  - le vecteur des centres d'intervalles calculés

code ci-dessous

```
 function [ddp, c] = histo(N, m, sigma, M)

if nargin==3
  delta_x=3.49*std(N)^2*length(N)^(-1/3);
  c1=min(N)+delta_x/2;
  cx=max(N)-delta_x/2;
  c=c1:delta_x:cx;

  [ddp,c]=hist(N,length(c));
  ddp=ddp/(length(N)*delta_x);

else nargin == 4
  delta_x=(max(N)-min(N))/M;
  c1=min(N)+delta_x/2;
  cx=max(N)-delta_x/2;
  c=c1:delta_x:cx;

  [ddp,c]=hist(N,M);
  ddp=ddp/(length(N)*delta_x);
```

```
end
```

$\square$

2.2 Expérimentation

2.2.1 Cas général

On supposera que le signal est échantillonné à la fréquence $F_s = 1\,KHz$. Ce choix est il important ? Pourquoi ?

réponse ci-dessous

Non, le choix de la fréquence d'échantillonnage n'est pas déterminant pour la suite du TP. En effet, elle n'influence pas du tout le signal, qui reste le même car il s'agit de valeurs prises aléatoirement. La seule chose qui change sera la précision, car la fréquence d'échantillonnage correspond au nombre de points que nous gardons lorsqu'on étudie un signal. $\square$

Dans les conditions suivantes :

- $N = 1000$ échantillons de signal
- Filtre passe-bas avec $B = 100\,Hz$ (ordre $m = 8$)
- $m_3 \neq 0$ et $\sigma_3 > 1$ (choix libres que l'on précisera clairement dans le compte-rendu)
- choix empirique *optimal* de la largeur Δx des intervalles,

afficher ci-dessous, sur une même figure partagée en 2×4 **sous-graphes** (*subplots*) :

- sur la première ligne : les séries temporelles $x_1(k.T_s)$, $x_2(k.T_s)$ et $x_3(k.T_s)$, ainsi que le module du gain complexe du filtre passe-bas
- sur la deuxième ligne : sous chacune des 3 séries temporelles, les densités de probabilité estimées auxquelles on superposera les densités théorique correspondantes. **Donner aussi le code utilisé pour calculer et afficher ces d.d.p. théoriques.**

figure ci-dessous

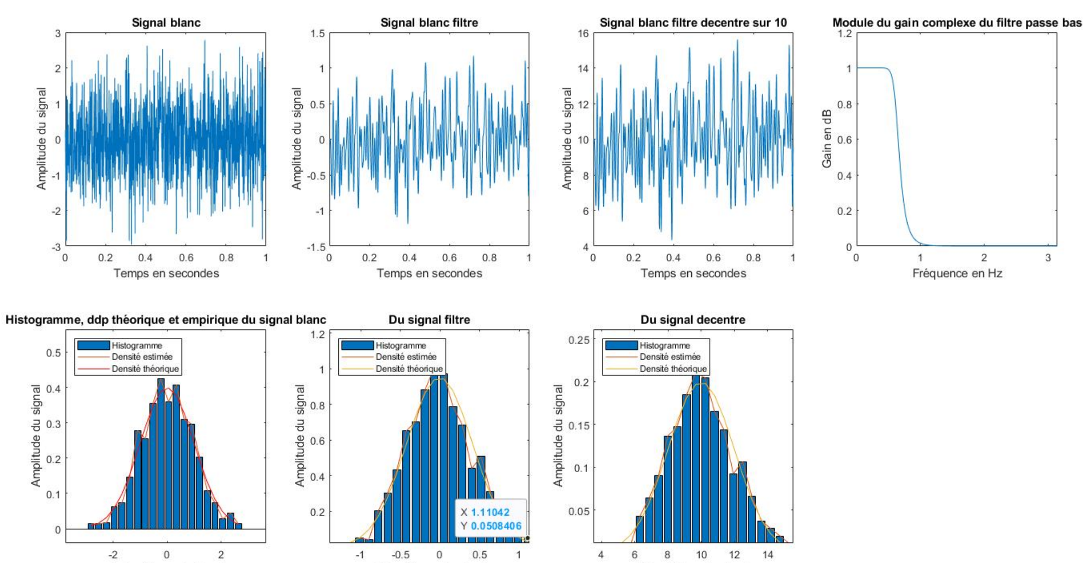

Figure 1 - Signaux temporels et ddp associées

□

code ci-dessous

```
m1 = 0;
sigma1 = 1;
m2 = mean(x2);
sigma2 = std(x2);
m3 = 10;
sigma3 = 2;

[ddp1,c1] = histo(x1,m1,sigma1,M);
[ddp2,c2] = histo(x2,m2,sigma2,M);
[ddp3,c3] = histo(x3,m3,sigma3,M);

ddp1_th = (1/(sigma1*sqrt(2*pi)))*exp((-(c1-m1).^2)/(2*sigma1^2));
ddp2_th = (1/(sigma2*sqrt(2*pi)))*exp((-(c2-m2).^2)/(2*sigma2^2));
ddp3_th = (1/(sigma3*sqrt(2*pi)))*exp((-(c3-m3).^2)/(2*sigma3^2));

subplot(245);
bar(c1,ddp1);
hold on;
plot(c1,ddp1);
plot(c1,ddp1_th,'r','Linewidth',0.5);
title('Histogramme, ddp théorique et empirique du signal blanc');
legend(["Histogramme","Densité estimée", "Densité théorique"],"FontSize", 6, "Location", "northwest");
xlabel("Echantillons du signal");
ylabel("Amplitude du signal");

subplot(246);
bar(c2, ddp2);
hold on;
plot(c2,ddp2);
plot(c2,ddp2_th);
title('Du signal filtre');
legend(["Histogramme","Densité estimée", "Densité théorique"],"FontSize", 6, "Location", "northwest");
xlabel("Echantillons du signal");
ylabel("Amplitude du signal");

subplot(247);
bar(c3, ddp3);
hold on;
plot(c3, ddp3);
plot(x_out3,ddp3_th);
title('Du signal decentre');
legend(["Histogramme","Densité estimée", "Densité théorique"],"FontSize", 6, "Location", "northwest");
xlabel("Echantillons du signal");
ylabel("Amplitude du signal");
```

□

Pour chacun des 3 processus, vérifier par la mesure sur les densités estimées et en utilisant des estimateurs empiriques (disponibles sous Matlab) :

a) la conformité entre moyennes mesurées et théoriques

	$\widehat{m_1}$	$\widehat{m_2}$	$\widehat{m_3}$
Décrire une 1ère méthode de mesure de la moyenne	**réponse et mesures ci-dessous** Matlab possède une fonction mean() qui permet de mesurer la moyenne d'un signal. □		
Mesure de la moyenne par la méthode 1	$m1_{emp} = 0.0080$	$m2_{emp} = 0.010$	$m3_{emp} = 9.97$
Décrire une 2ème méthode de mesure de la moyenne	**réponse et mesures ci-dessous** Nous avons vu en probabilité que pour une gaussienne, la valeur maximale de la ddp est atteinte en x = la moyenne. □		
Mesure de la moyenne par la méthode 2	$m1_{emp} = 0.05$	$m2_{emp} = 0.06$	$m3_{emp} = 9.7$

b) idem pour les écart-type (avec au moins deux méthodes de mesure distinctes que l'on détaillera)

	$\widehat{\sigma_1}$	$\widehat{\sigma_2}$	$\widehat{\sigma_3}$
Décrire une 1ère méthode de mesure de l'écart-type	**réponse et mesures ci-dessous** Matlab possède une fonction std() qui permet de mesurer l'écart-type d'un signal. □		
Mesure de l'ecart-type par la méthode 1	$\widehat{\sigma_1} = 1.02$	$\widehat{\sigma_2} = 0.42$	$\widehat{\sigma_3} = 2.03$
Décrire une 2ème méthode de mesure de l'écart-type	**réponse et mesures ci-dessous** A mi-hauteur de la gaussienne en x $= \mu$ la moyenne du signal, on trouve la courbe à x $\approx \mu + 1.17\sigma$ et x $\approx \mu - 1.17\sigma$. Donc la largeur de la courbe à mi-hauteur vaut 2.35σ. □		
Mesure de l'écart-type par la méthode 2	$\widehat{\sigma_1} = 0.97$	$\widehat{\sigma_2} = 0.40$	$\widehat{\sigma_3} = 1.91$
Décrire une 3ème méthode de mesure de l'écart-type	**réponse et mesures ci-dessous** On peut mesurer la base, ou support, de la densité qui est égale à 6σ. En effet, 99% des réalisations d'une gaussienne $\in [m - 3\sigma; m + 3\sigma]$. □		
Mesure de l'écart-type par la méthode 3	$\widehat{\sigma_1} = 0.9$	$\widehat{\sigma_2} = 0.37$	$\widehat{\sigma_3} = 1.80$
Décrire une 4ème méthode de mesure de l'écart-type	**réponse et mesures ci-dessous** On peut également mesurer l'amplitude de la ddp pour x = la moyenne, qui correspont à $\frac{1}{\sigma\sqrt{2\pi}}$. □		
Mesure de l'écart-type par la méthode 4	$\widehat{\sigma_1} = 1.1$	$\widehat{\sigma_2} = 0.40$	$\widehat{\sigma_3} = 1.90$

Lesquelles de ces méthodes vous paraissent les plus précises ? Pourquoi ?

réponse ci-dessous

Les méthodes utilisant std et mean paraissent être les plus précises. En effet, elles ne s'appuyent que sur un calcul mathématique et non sur une estimation graphique pouvant varier selon la personne qui lit le graphique, ou même la précision disponible dans l'environnement de travail.

□

2.2.2 Influence de N

On ne considère ici que le signal aléatoire $x_1(t)$, le nombre d'intervalles pour le calcul des histogrammes restant constant et égal à $M = 20$.

a) Sur une même figure, afficher dans différents sous-graphes (pour une meilleure lisibilité des courbes, on pourra utiliser la commande `stem.m` en lieu et place de la commande `bar.m`), les densités de probabilité de $x_1(t)$ estimées pour plusieurs valeurs du nombre d'échantillons : pour cela faire varier dans une boucle `for...end`, le nombre N de 2^4 à 2^{11}. Superposer systématiquement les densités théoriques ainsi que les intervalles de précision théoriques $\mathbb{E}\{\widehat{p_{\mathbf{x}}}(x)\} \pm \text{std}(\widehat{p_{\mathbf{x}}}(x))$ calculés en TD. Veiller à commenter précisément chaque figure (légendes, labels,...)
Donner aussi le code Matlab de calcul de ces intervalles de confiance.

figures ci-dessous

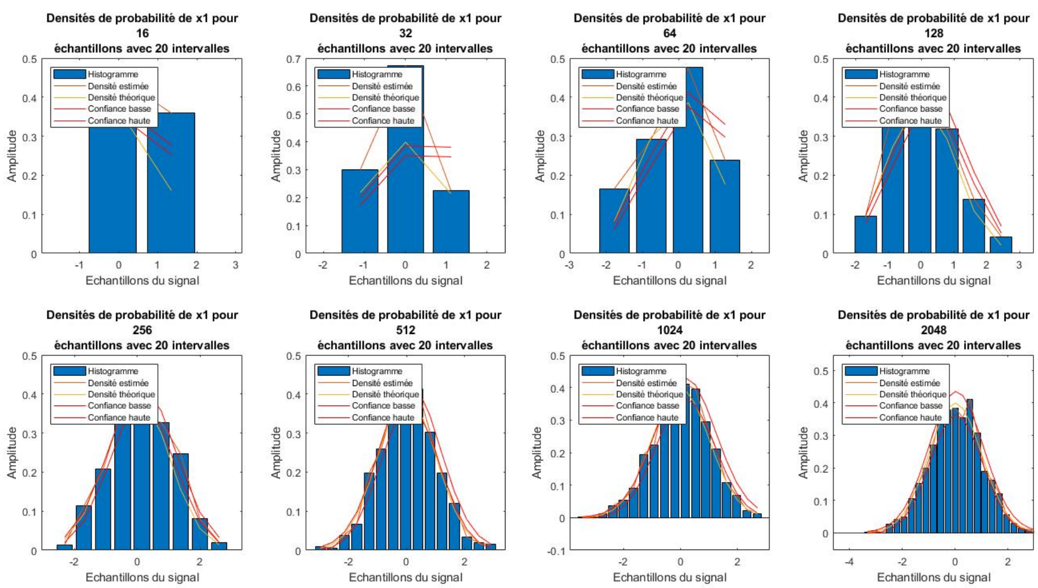

Figure 2 - Signal aléatoire x1 selon différentes valeurs de N

□

code ci-dessous

```
figure(2);
for k = 0:7
    N2 = 2^(4+k);
    [x1i, x2i, x3i, Azi, Bzi] = synthese(N2, B, m3, sigma3);
    [ddp1_i,c1_i] = histo(x1i,m1,sigma1);
    ddp1_th = (1/(sigma1*sqrt(2*pi)))*exp((-(c1_i-m1).^2)/(2*sigma1^2));
    delta_x = c1_i(2) - c1_i(1);
    int_min = ddp1_th - sqrt((ddp1_th/N1).*((1/delta_x)-ddp1_th));
    int_max = ddp1_th + sqrt((ddp1_th/N1).*((1/delta_x)-ddp1_th));
```

```
    subplot(2, 4, k+1);
    bar(c1_i, ddp1_i); hold on;
    plot(c1_i, ddp1_i);
    plot(c1_i, ddp1_i);
    plot(x_out1_i, int_min, 'r');
    plot(x_out1_i, int_max, 'r');
    title(["Densités de probabilité de x1 pour", {N2},"échantillons avec 20 intervalles"]);
    legend(["Histogramme","Densité estimée", "Densité théorique", "Confiance basse", "Confiance hau
    xlabel("Echantillons du signal");
    ylabel("Amplitude");

end
```

□

b) Qualitativement, expliquez à partir de ces tracés, l'évolution de la variance (ou de l'écart-type) d'estimation.

réponse ci-dessous

Nous pouvons constater que moins on a d'échantillons, plus la variance est grande et de mauvaise qualité. En effet, moins on a d'échantillons, plus les échantillons sont mal répartis dans l'histogramme ce qui impacte directement la variance puisque cette grandeur caractérise la dispersion des échantillons autour de la moyenne. Plus on en a et mieux l'histogramme sera fournis, avec des valeurs cohérentes. □

c) Peut on conclure sur le biais d'estimation à partir de cette seule expérience ? Expliquez.

réponse ci-dessous

Pour pouvoir conclure sur le biais de manière précise, il faudrait réaliser un grand nombre d'expériences. Ici, nous ne pouvons conclure que sur la variation de l'estimateur. En effet, nous pouvons réaliser l'expérience avec un nombre très élevé d'échantillons, ce qui devrait stabiliser la variation. □

d) Quelle expérience faudrait il mener pour caractériser empiriquement et précisément le biais et la variance d'estimation ?

réponse ci-dessous

On a M le nombre d'intervalles lorsqu'il est imposé. Si nous faisons varier M, nous jouons sur la largeur d'un intervalle comprenant une seul catégorie d'échantillons. En augmentant ou diminuant ce M, nous pourrons caractériser le biais. □

2.2.3 Influence de Δx

Ici encore, on ne s'intéresse qu'à $x_1(t)$ et à une de ses réalisations sur $N = 1000$ points.

a) En faisant varier M, le nombre d'intervalles de l'histogramme, sur une plage incluant les 2 situations extrêmes (**que l'on indiquera et justifiera**), calculer et afficher (sur une même figure partagée en sous-graphes) les densités de probabilité estimées. Superposer les densités théoriques ainsi que les intervalles de précision.

réponse ci-dessous

Comme cas extrêmes nous avons pris 2 et 1000 intervalles. 2 car c'est le nombre minimal d'intervalles que nous pouvons mettre et 1000 car cela correspond au nombre d'échantillons que nous avons générés. Nous ne pouvons pas avoir plus d'intervalles que d'échantillons. □

b) Dans un dernier sous-graphe de la même figure, représenter la densité de probabilité estimée avec un choix optimal de Δx.

figures ci-dessous

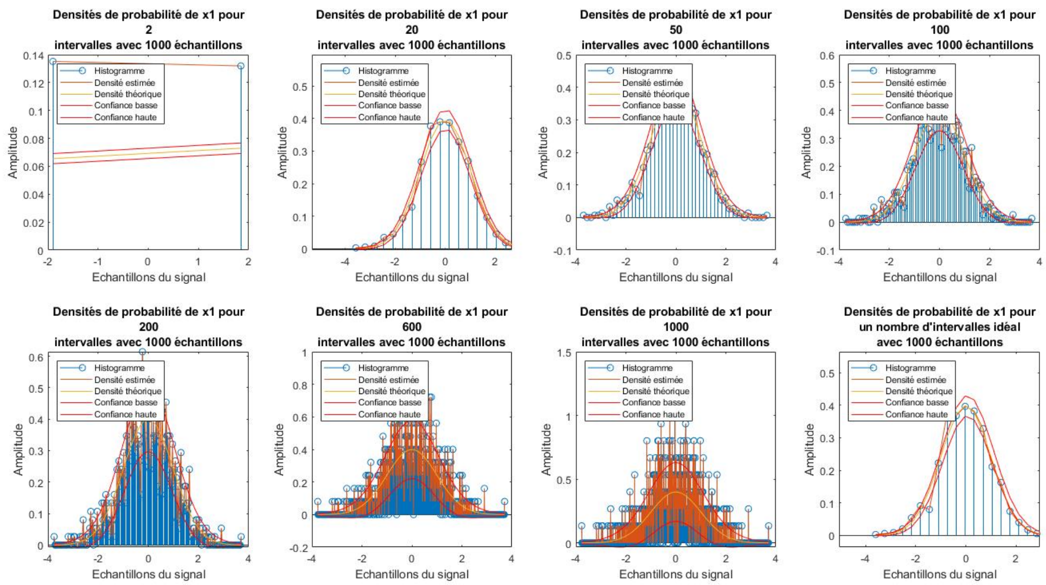

Figure 3 - Ddp estimées selon Δx

□

c) Comme pour la question précédente, décrivez qualitativement en l'expliquant, l'évolution de la variance et du biais d'estimation en fonction de Δx.

réponse ci-dessous

Nous avons vu en TD que la variance d'estimation s'écrivait :

$$\epsilon^2 = \frac{Var(h_i^{\Delta x})}{E^2(h_i^{\Delta x})} = \frac{1}{N}\left(\frac{1}{\Delta x f_X(c_i)} - 1\right)$$

Donc pour Δx fixé, si N augmente, la variance augmente. Le biais s'écrit :

$$Biais(\hat{f_X}(c_i)) = E(\hat{f_X}(c_i)) - f_X(c_i)$$

Or quand Δx tend vers 0, l'espérance de $\hat{f_X}(c_i)$ tend vers : $f_X(c_i)$. Donc nous avons le biais qui tend vers

$$Biais(\hat{f_X}(c_i)) = f_X(c_i) - f_X(c_i) = 0$$

Ainsi pour un Δx petit, nous avons un estimateur qui est non-biaisé, cependant la variance augmente. Lorsque Δx augmente, le biais augmente également, mais la variance diminue. Il faut donc mettre au point un compromis biais-variance. □

2.2.4 Influence de B

On se place dans les conditions suivantes :

- $N = 1000$ échantillons
- $m_3 \neq 0$ et $\sigma_3 > 1$ (garder les mêmes valeurs que celles choisies pour la première expérience)
- choix empirique *optimal* des largeurs d'intervalles Δx
- Filtre de Butterworth passe-bas, d'ordre $m = 8$ et de **bande** $B = 5\,Hz$.

a) Afficher sur une même figure dans différents sous-graphes, le gabarit (gain complexe) du filtre passe-bas correspondant, le processus filtré $x_2(t)$ et la densité de probabilité estimé sur le processus filtré $x_2(t)$. **Superposer la densité théorique**.

figures ci-dessous

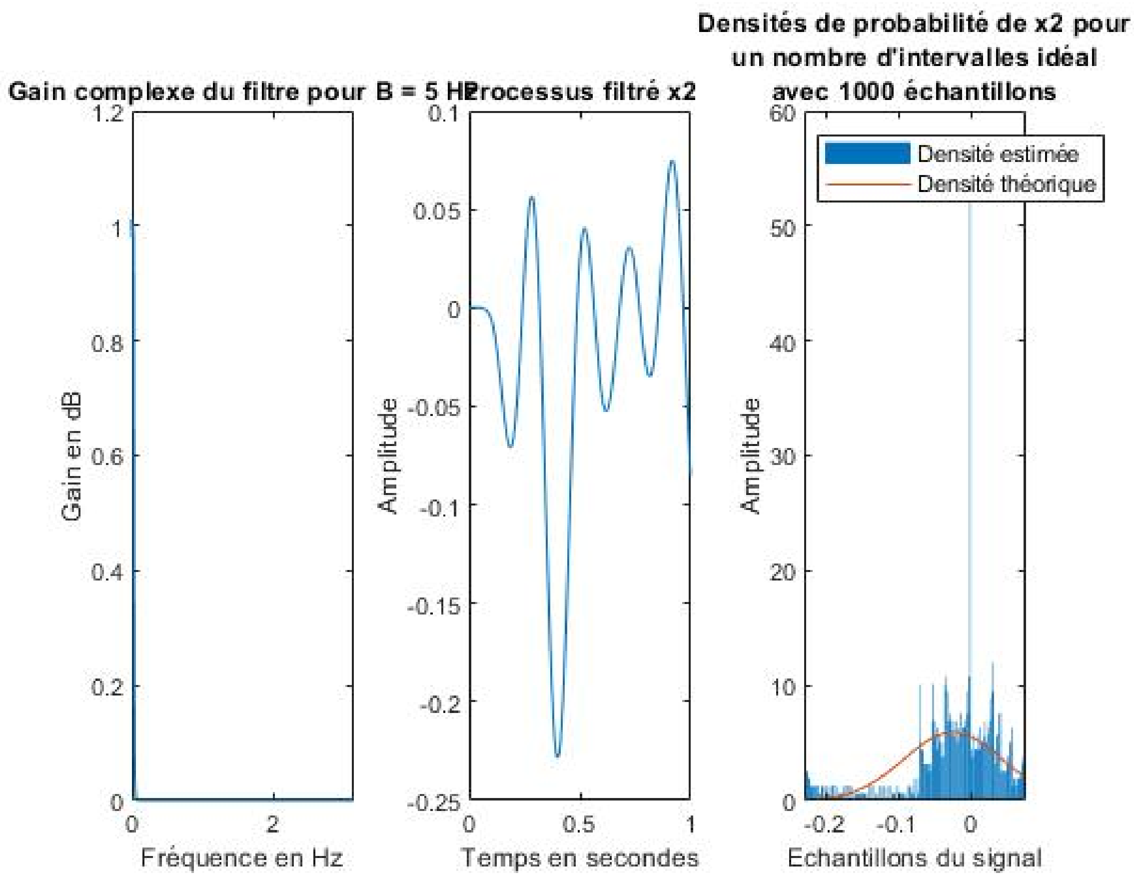

Figure 4 - Gain complexe du filtre et courbes du processus filtré

□

b) Le signal $x_2(t)$ est il gaussien ? Justifiez votre réponse (on pourra par exemple calculer le Kurtosis sur la série temporelle $(x_2[n])_{n=1,\dots N}$).

réponse ci-dessous

En utilisant la fonction kurtosis fourni pour le tp, nous trouvons en faisant une dizaine d'essais un Kurtosis variant entre 2.1 et 3.9 or pour une gaussienne K = 3. Nous pouvons donc en déduire que le signal $x_2(t)$ est grossièrement gaussien. □

c) Pourquoi l'estimation de la densité de probabilité de x_2 est elle aussi différente de la densité gaussienne $\mathcal{N}(m_2, \sigma_2)$? En gardant $B = 5\,Hz$, proposer une nouvelle configurations de paramètres pour corriger cet effet. Vérifier la solution proposée, **en affichant la densité de probabilité ainsi estimée**.

réponse ci-dessous

Lorsque l'on observe le spectre d'un bruit gaussien, on se rend compte que ce dernier se répartit sur l'ensemble du spectre. En appliquant un filtre avec une fréquence de coupure aussi basse, on perd à fortiori beaucoup d'échantillons contenant les informations du signal. Pour palier cela nous pouvons augmenter le nombre d'échantillons du signal. En prenant N = 100 000 on obtient la figure ci-dessous, on remarque que la densité estimée est beaucoup plus proche de la théorique comparé au graphe précédent. De plus sur une dizaine d'essais le Kurtosis ne fluctue plus qu'entre 2.93 et 3.15. On peut donc maintenant affirmer que x2(t) est quasi gaussien. □

figures ci-dessous

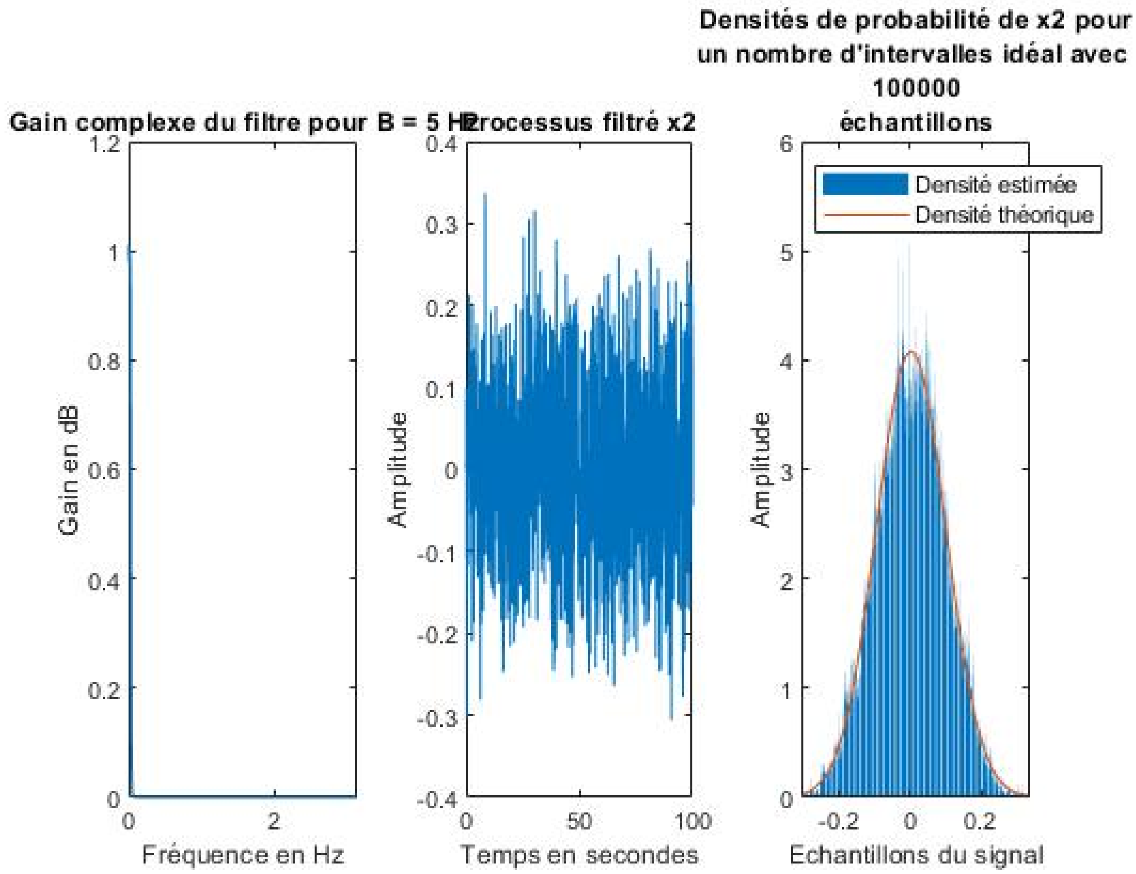

Figure 5 - Gain complexe du filtre et courbes du processus filtré, influence de N

□

3 Somme d'un signal carré à retard équiparti et d'un bruit gaussien

On veut étudier la densité de probabilité de la somme d'un signal carré à retard équiparti **y** et d'un bruit gaussien **x** de valeur moyenne m_B et d'écart-type σ_B.

Pour cela, utiliser la fonction Matlab `carbr(moy,ecartype,N)`, où :

`moy`: moyenne du bruit

`ecartype`: écart-type du bruit

`N`: nombre de points de signal à analyser

Le signal carré, de fréquence $\nu_0 = 110\,Hz$, d'amplitude ± 1, a pour retard à l'origine, une variable aléatoire τ distribuée uniformément sur l'intervalle $[0, T_0[$, où $T_0 = 1/\nu_0$ est la période du signal carré.

En quelques mots, expliquer alors, en quoi le signal carré est un signal aléatoire ?

réponse ci-dessous

Le signal carré possède une phase à l'origine τ comprise $[0; T_0[$. Ainsi, on observe une variation de l'amplitude moyenne qui dépend de l'intervalle de temps sur lequel nous travaillons. □

La somme **z** des 2 signaux aléatoires est échantillonnée à 100 kHz.
La fonction affiche le mélange signal carré + bruit et la d.d.p. estimée $\widehat{P_{\mathbf{z}}}(z)$.

En choisissant la moyenne du bruit $m_B = 0$, trouver, en la justifiant, la valeur de l'écart-type σ_B correspondant à chacune des 2 situations suivantes :

1) $\mathbb{P}\{z \in [-0.5, 0.5]\} \leq 0.5\%$

réponse ci-dessous

Nous avons déterminé à la question 5 de la préparation que $p_Z(T)$ était constitué de deux demi-gaussiennes situées en +A et -A par rapport à la valeur de la moyenne du signal. Donc nous avons

$$P(Z \in [-0.5; 0.5]) = 0.5\% \Longleftrightarrow 2P(\in [0; 0.5])$$

Avec le calcul de la distribution $\frac{1}{2}p_X(T - A)$ on peut dire :

$$2.\frac{1}{2}P(Z < 0.5) = 0.5\% \Longleftrightarrow P(Z < 0.5) = 0.5\%$$

On peut pose X = Z - A et obtenir le résutat suivant :

$$P(X < 0.5 - 1) = 0.5\%$$

Or X est centré donc on peut écrire :

$$P(X < -0.5) = P(X < -3\sigma_B) \Longleftrightarrow -\frac{1}{2} = -3\sigma_B \Longleftrightarrow \sigma_B = \frac{1}{6}$$

□

2) $p_{\mathbf{z}}(0) = \dfrac{1}{2}\,p_{\mathbf{x}}(0)$

réponse ci-dessous

$$p_Z(0) = \frac{1}{2}(p_X(1) + p_X(-1)) = 2.\frac{1}{2}p_X(1) = p_X(1)$$

Cela revient donc à chercher σ_B pour

$$p_X(1) = \frac{1}{2}p_X(0)$$

Or on connait la valeur de la largeur à mi-hauteur d'une gaussienne, qui vaut 2.35σ. Donc finalement :

$$\sigma_B = \frac{2}{2.35} \approx 0.85$$

□

Afficher sur une même figure, les deux densités correspondantes.

figures ci-dessous

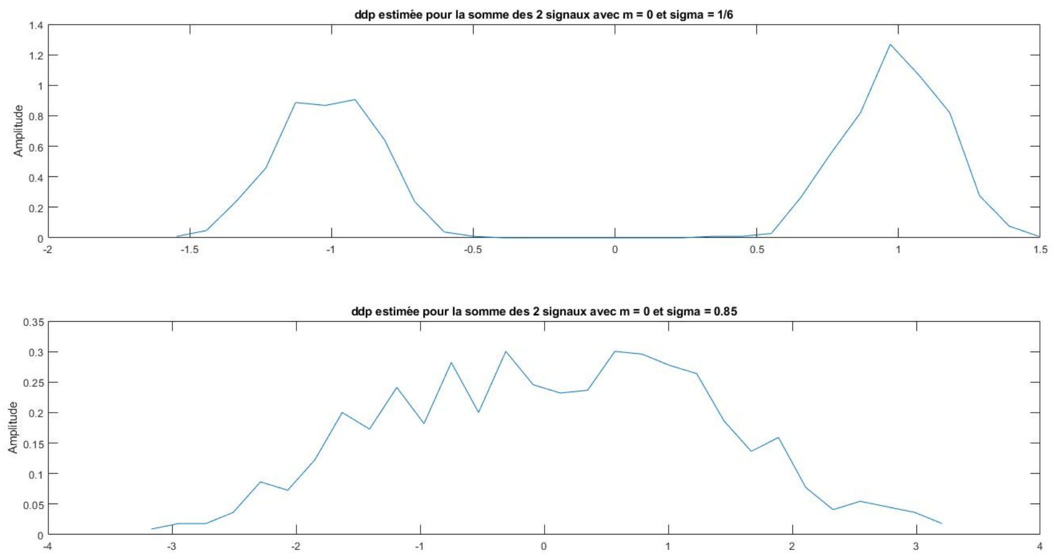

Figure 6 - Ddp pour les variances $\sigma_B = \frac{1}{6}$ et $\sigma_B = 0.85$

□