Traitement des Signaux Aléatoires
Estimation de densités de probabilité

4 ETI – CPE Lyon

Travaux Pratiques TSA

Noms, Prénoms : LANGUILLE Antoine, BURNOT Jean-Christophe

Groupe : D

Date : 21 Novembre 2022

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 $(x[n])_{n=1,...N}$ soient identiquement distribués (i.e. suivent tous la même loi, quelque soit l'instant n) ?

réponse ci-dessous

Pour que les échantillons soient identiquement distribués, il faut que le signal soit strictement stationnaire. □

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

réponse ci-dessous

Il faut que $\mathbb{E}\{X(t_1)X(t_2)\} = 0$ □

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

réponse ci-dessous

La décorrélation entraine l'indépandance dans le cas d'un signal gaussien. □

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

La largeur des classes Δx illustre le compromis biais-variance

— Si Δx augmente, plus d'échantillons tombent dans la même classe, l'histogramme se rapproche plus d'un moyenneur. Le biais augmente donc, la densité de probabilité tendant vers l'espérance du signal. La variance se réduit alors.

$$\Delta x \text{ augmmente} \Rightarrow \text{biais augmente, variance diminue}$$

— Si Δx diminue, les classes sont plus nombreuses et plus petites, leur valeur dépend alors davantage des échantillons individuels, la variance augmente donc. Au contraire, le biais se réduit, la forme de l'histogramme se rapprochant plus de celle de la densité de probabilité.

$$\Delta x \text{ diminue} \Rightarrow \text{biais diminue, variance augmente}$$

□

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

Les quatres opérations élémentaires suffisent $(+, -, \times, \div)$. On procède en 2 étapes :

— On centre réduit la gaussienne $G_1 \sim Gau(m_1, \sigma_1)$ en $G_0 \sim Gau(0, 1)$ par l'opération

$$G_0 = \frac{G_1 - m_1}{\sigma_1}$$

— On dilate puis décale la loi gaussienne centrée réduite que l'on vient de calculer pour obtenir une loi gaussienne $G_2 \sim Gau(m_2, \sigma_2)$ grâce à l'opération

$$G_2 = G_0 \times \sigma_2 + m_2$$

L'opération finale s'exprime donc

$$G_2 = \frac{G_1 - m_1}{\sigma_1}\sigma_2 + m_2$$

□

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 : $K = \frac{\mathbb{E}\left\{\mathbf{x}^4\right\}}{\mathbb{E}^2\left\{\mathbf{x}^2\right\}}$

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

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

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

réponse ci-dessous

On a

$$K = \frac{\mathbb{E}\{x^4\}}{\mathbb{E}^2\{x^2\}} = \frac{\mathbb{E}\{x^2\} + \mathbb{E}\{x^2\} \times \mathbb{E}\{x^2\} + \mathbb{E}\{x^2\} \times \mathbb{E}\{x^2\} + \mathbb{E}\{x^2\} \times \mathbb{E}\{x^2\}}{\mathbb{E}^2\{x^2\}} = \frac{3\mathbb{E}^2\{x^2\}}{\mathbb{E}^2\{x^2\}} = 3$$

□

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 $x(t) \sim Gau(m_B, \sigma_B)$.

$$\text{et } y(t) = \begin{cases} \text{Amplitude : } \text{A} \\ \text{Centré : } \text{m=0} \\ \text{Période : } \text{T}_0 \\ \alpha = 1 \text{(fonction constante)} \\ \text{Retard : } \tau \end{cases} \Leftrightarrow y(t) = A.$$

La densité de probabilité de $z(t) = x(t) + y(t)$ n'est autre qu'une gaussienne de même écart-type σ_B que $x(t)$ mais d'espérance $m_B + A$.

□

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

Travaux Pratiques TSA

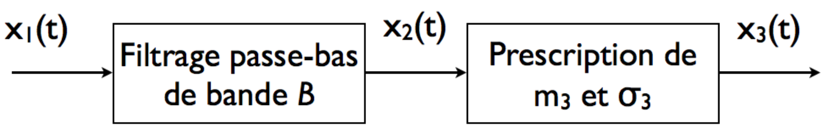
Noms, Prénoms : LANGUILLE Antoine, BURNOT Jean-Christophe

Groupe : D

Date : 21 Novembre 2022

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, A, B] = synthesize(N, B, m, s)

    Fs = 1e3;    % Fréquence d'échantillonnage 1kHz
```

```
    X1 = randn(N,1);            % Bruit gaussien centré réduis
    fc = B/(Fs/2);              % Fréquence normalisée pour filtre
    [B, A] = butter(8, fc);     % Génération du filtre
    X2 = filter(B, A, X1);      % Filtrage

    % X2 n'est plus centré réduis après filtrage
    X2 = (X2-mean(X2))/std(X2); % On centre réduis X2

    % On inverse l'opération de centrage-réduction
    X3 = s*X2 + m;      % Dilatation décalage du bruit centré réduis
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 [dp,cbins] = drawhist(X,M)
    if ~exist('M','var')
        % Valeur par défaut si M non fourni

        % Largeur ayant le meilleur compromis biais-variable
        dx = 3.49 * std(X) * length(X) ^ (-1/3);

        M = ceil((max(X) - min(X)) / dx); %nombre de bins associé à dx
        msg = "optimale";
    else
        % M fourni
        dx = (max(X) - min(X)) / M;     % On calcule dx à partir de M
        msg = "imposée";
```

```
    end

    [counts, cbins] = hist(X, M);    % On précise ici le nombre de bins de l'histogramme

    dp = counts./(length(X)*dx);    % On normalise l'histogramme en densité de probabilités

    % Affichage
    bar(cbins, dp);
    hold on;
    xlabel(sprintf("\\Deltax = %f (%s)",dx, msg));
    title(sprintf("Estimation DP (\\Deltax=%d)", dx));
end
```

□

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

La fréquence d'échantillonnage est importante, car si elle est trop faible, il y aura repliement spectral (Théorème de Shannon-Nyquist). □

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

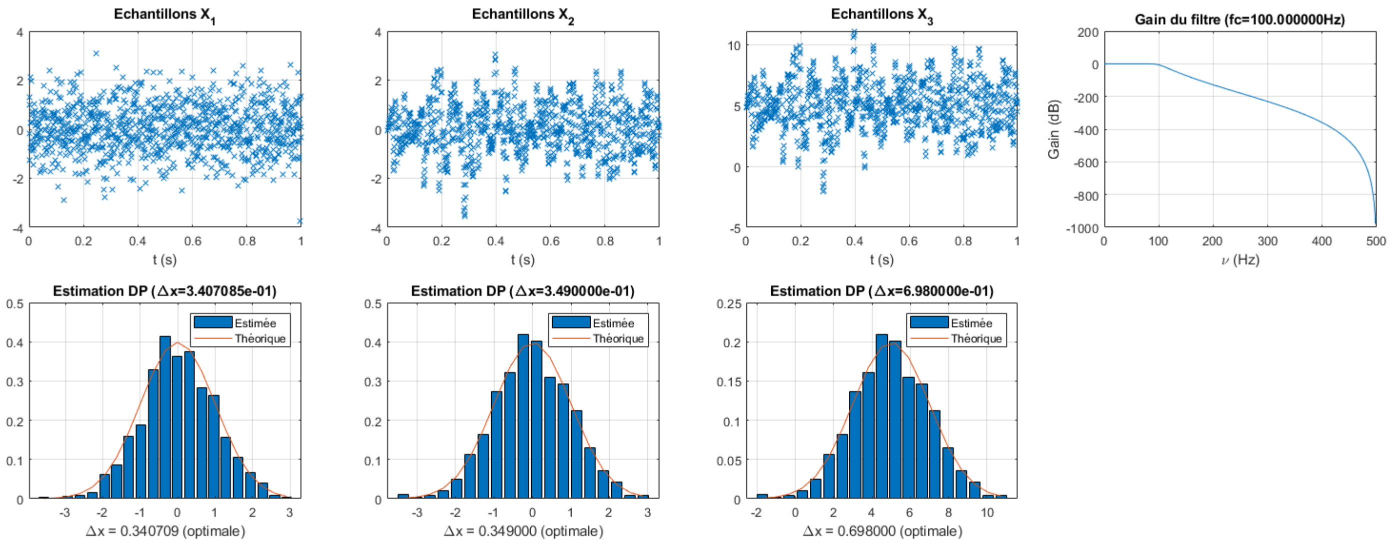


FIGURE 1 – Signaux et histogrammes associés

□

code ci-dessous

```
%% TP 1: Estimation de densités de probabilité

clear variables;
close all;

Fs = 1e3;   % 1kHz

N = 1e3;    % 1000 échantillons
B = 100;    % 100Hz

fc = B;     % Fréquence de coupure du filtre

% Espéarance et écart-type du bruit X3
m = 5;
s = 2;

% On charche dx optimal (on ne précise donc pas M)
[X1, X2, X3, A, B] = synthesize(N, B, m, s);

% Affichage
figure(1);
subplot(2,4,1);
x = 0 : 1/Fs : (length(X1)-1)/Fs;
plot(x, X1, 'x');
title("Echantillons X_1");
xlabel("t (s)");

subplot(2,4,2);
x = 0 : 1/Fs : (length(X2)-1)/Fs;
plot(x, X2, 'x');
```

```
title("Echantillons X_2");
xlabel("t (s)");


subplot(2,4,3);
x = 0 : 1/Fs : (length(X3)-1)/Fs;
plot(x, X3, 'x');
title("Echantillons X_3");
xlabel("t (s)");


subplot(2,4,4);
[h,f] = freqz(B, A, N, Fs);
plot(f, 20*log(abs(h)));
title(sprintf("Gain du filtre (fc=%fHz)", fc));
xlabel("\nu (Hz)");
ylabel("Gain (dB)");

% Gaussienne
gauss = @(x, mu, sig) (1/(sig*sqrt(2*pi)) * exp( -(x-mu).^2/(2*sig^2) ));

subplot(2,4,5);
[dp1, cbins1] = drawhist(X1);
plot(cbins1, gauss(cbins1, 0, 1));
legend('Estimée','Théorique');

subplot(2,4,6);
[dp2, cbins2] = drawhist(X2);
plot(cbins2, gauss(cbins2, 0, 1));
legend('Estimée','Théorique');

subplot(2,4,7);
[dp3, cbins3] = drawhist(X3);
plot(cbins3, gauss(cbins3, m, 2));
legend('Estimée','Théorique');
```

□

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. Compléter la **Table 1** avec les valeurs mesurées.

<table>
<tr><td></td><td>$\widehat{m_1}$</td><td>$\widehat{m_2}$</td><td>$\widehat{m_3}$</td></tr>
<tr><td>Décrire une 1ère méthode de mesure de la moyenne</td><td colspan="3">réponse et mesures ci-dessous
On prend la valeur centrale de la classe ayant la plus grande probatilité. □</td></tr>
<tr><td>Mesure de la moyenne par la méthode 1</td><td>$\widehat{m_1} \approx -0.334$</td><td>$\widehat{m_2} \approx -0.2403$</td><td>$\widehat{m_3} \approx 4.519$</td></tr>
<tr><td>Décrire une 2ème méthode de mesure de la moyenne</td><td colspan="3">réponse et mesures ci-dessous
On prend la valeur médianne des valeurs de nos échantillons. □</td></tr>
<tr><td>Mesure de la moyenne par la méthode 2</td><td>$\widehat{m_1} \approx -0.334$</td><td>$\widehat{m_2} \approx -0.2403$</td><td>$\widehat{m_3} \approx 4.519$</td></tr>
</table>

TABLE 1 – Estimations de la valeur moyenne des signaux

b) idem pour les écart-type (avec <u>au moins deux méthodes</u> de mesure distinctes que l'on détaillera). Compléter la **Table 2** avec les valeurs mesurées.

<table>
<tr><td></td><td>$\widehat{\sigma_1}$</td><td>$\widehat{\sigma_2}$</td><td>$\widehat{\sigma_3}$</td></tr>
<tr><td>Décrire une 1ère méthode de mesure de l'écart-type</td><td colspan="3">réponse et mesures ci-dessous
On regarde les deux valeurs pour lesquelles la densité de probabilité atteint la moitié de la hauteur maximale de la courbe. L'écart entre ces deux valeurs constitue la "Largeur à mi hauteur" et est égale à environ 2.355σ. $\square$</td></tr>
<tr><td>Mesure de l'ecart-type par la méthode 1</td><td>$\widehat{\sigma_1} \approx 1.023628$</td><td>$\widehat{\sigma_2} \approx 1.144940$</td><td>$\widehat{\sigma_3} \approx 2.073252$</td></tr>
<tr><td>Décrire une 2ème méthode de mesure de l'écart-type</td><td colspan="3">réponse et mesures ci-dessous
On regarde la valeur pour laquelles, l'aire sous la courbe centrée entre cette valeur et son opposée atteint 68.27% de l'aire totale. Cette valeur est alors l'écart-type σ. $\square$</td></tr>
<tr><td>Mesure de l'écart-type par la méthode 2</td><td>$\widehat{\sigma_1} \approx 1.240000$</td><td>$\widehat{\sigma_2} \approx 1.236000$</td><td>$\widehat{\sigma_3} \approx 2.398000$</td></tr>
<tr><td>Décrire une 3ème méthode de mesure de l'écart-type</td><td colspan="3">réponse et mesures ci-dessous
$\square$</td></tr>
<tr><td>Mesure de l'écart-type par la méthode 3</td><td>$\widehat{\sigma_1} \approx$</td><td>$\widehat{\sigma_2} \approx$</td><td>$\widehat{\sigma_3} \approx$</td></tr>
<tr><td>Décrire une 4ème méthode de mesure de l'écart-type</td><td colspan="3">réponse et mesures ci-dessous
$\square$</td></tr>
<tr><td>Mesure de l'écart-type par la méthode 4</td><td>$\widehat{\sigma_1} \approx$</td><td>$\widehat{\sigma_2} \approx$</td><td>$\widehat{\sigma_3} \approx$</td></tr>
</table>

Table 2 – Estimations de l'écart-type

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

réponse ci-dessous

Pour l'espérance, la méthode de la médianne semble plus précise car au contraire de la méthode de la barre la plus haute. En effet la médianne tiens compte de la totalié des échantillons, là où la méthode de la barre la plus haute ne dépends que des échantillons tombant dans la classe considérée.
Pour l'écart-type, la méthode faisait appel à la l'aire sous la courbe semble la plus précise car elle requiert une intégrale de la densité de probabilité estimée et fait donc intervenir tous les échantillons dans toutes les classes. En revanche la méthode de la largeur à mi hauteur ne fait intervenir au plus que 3 échantillons (valeur maximale atteinte et les deux échantillons pour mesurer la largeur). □

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 $\mathbb{E}\{\widehat{p_\mathbf{x}}(x)\} \pm \mathrm{std}(\widehat{p_\mathbf{x}}(x))$ 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

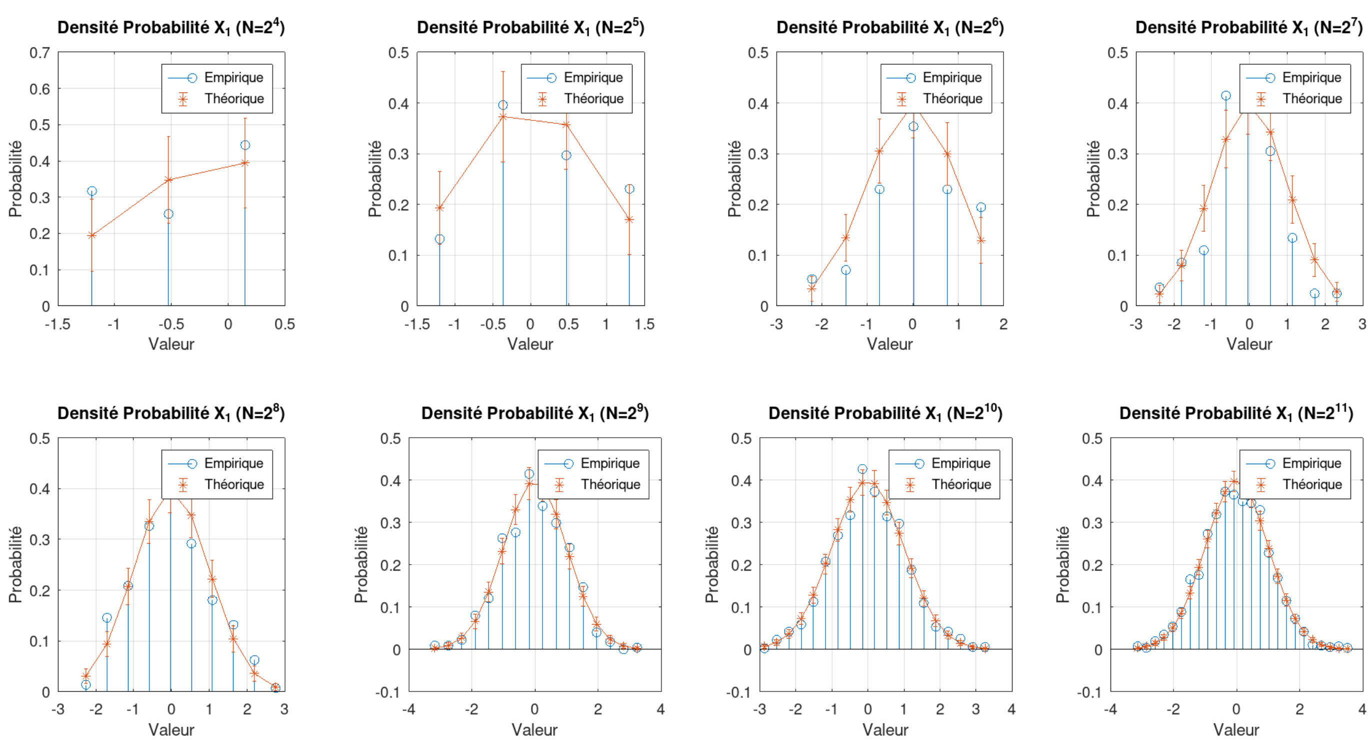


FIGURE 2 – Variation du nombre d'échantillons N

□

code ci-dessous

```
% Variation du paramètre M

figure(2);

for i=4:11

  subplot(2,4,i-3);

  N = 2^i;
  [X1, X2, X3, Ac, Bc] = synthesize(N, B, m, s);


  % Affichage de la densité empirique
  [dp1, cbins1] = drawhist(X1);

  % Affichagege de ala densité théorique
  dp1_th = gauss(cbins1, 0, 1);

  % Calcul de l'espérance et de l'écart-type de chaque classe
  dx = 3.49 * std(X1) * length(X1) ^ (-1/3);

  Ex = dp1_th;                     % Formules de TD
  Vx = dp1_th/N .* (1/dx - dp1_th);

  Incert1 = sqrt(Vx);

  errorbar(cbins1, dp1_th, Incert1, '-*');

  title(sprintf("Densité Probabilité X_1 (N=2^{%d})", i));
  xlabel("Valeur");
  ylabel("Probabilité");
  legend("Empirique","Théorique");
  grid on;


end
```

□

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

réponse ci-dessous

On observe que plus le nombre d'échantillons augmente, plus la densité de probabilité empirique se rapproche de la densité théorique. Les barres d'erreur se réduisent signifiant que l'espérance ne change pas (comme attendu) mais que l'écart-type et donc la variance diminue. □

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

réponse ci-dessous

À partir de cette expérience seule, on ne peut conclure définitivement sur le biais ou la variance de cet estimateur. En effet l'estimateur ne dépend pas directement du paramètre N. Le seul paramètre inhérent à l'estimateur est le paramètre M. Les paramètres N, B, Fc... dépendent des signaux synthétisés. □

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

réponse ci-dessous

Pour caractériser le biais et la variance de l'estimateur, il faudrait cette fois-ci faire varier le paramètre M (et donc Δx). Pour chaque valeur de M, on effectue plusieurs estimation du même signal aléatoire pour estimer le biais et la variance. □

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

Les deux situations extrêmes sont :

— $M = 1$, une seule classe récupère tous les échantillons, l'estimateur se résume alors à un moyenneur. On s'attend donc à avoir une erreur d'estimation maximale (densité de probabilité constante très éloignée d'une gaussienne) mais une variance très faible.

— $M = N$, autant de classe que d'échantillons. L'estimation se rapproche au plus près de la densité théorique. Le biais est donc minimum mais chaque classe ne contenant que peu d'échantillons, la valeur de l'estimation varie beaucoup entre chaque estimation, la variance sera alors importante.

Ansi, faire un balayage de $M = 1$ à $M = N$ permet de bien voir l'évolution du biais et de la variance de l'estimateur. □

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

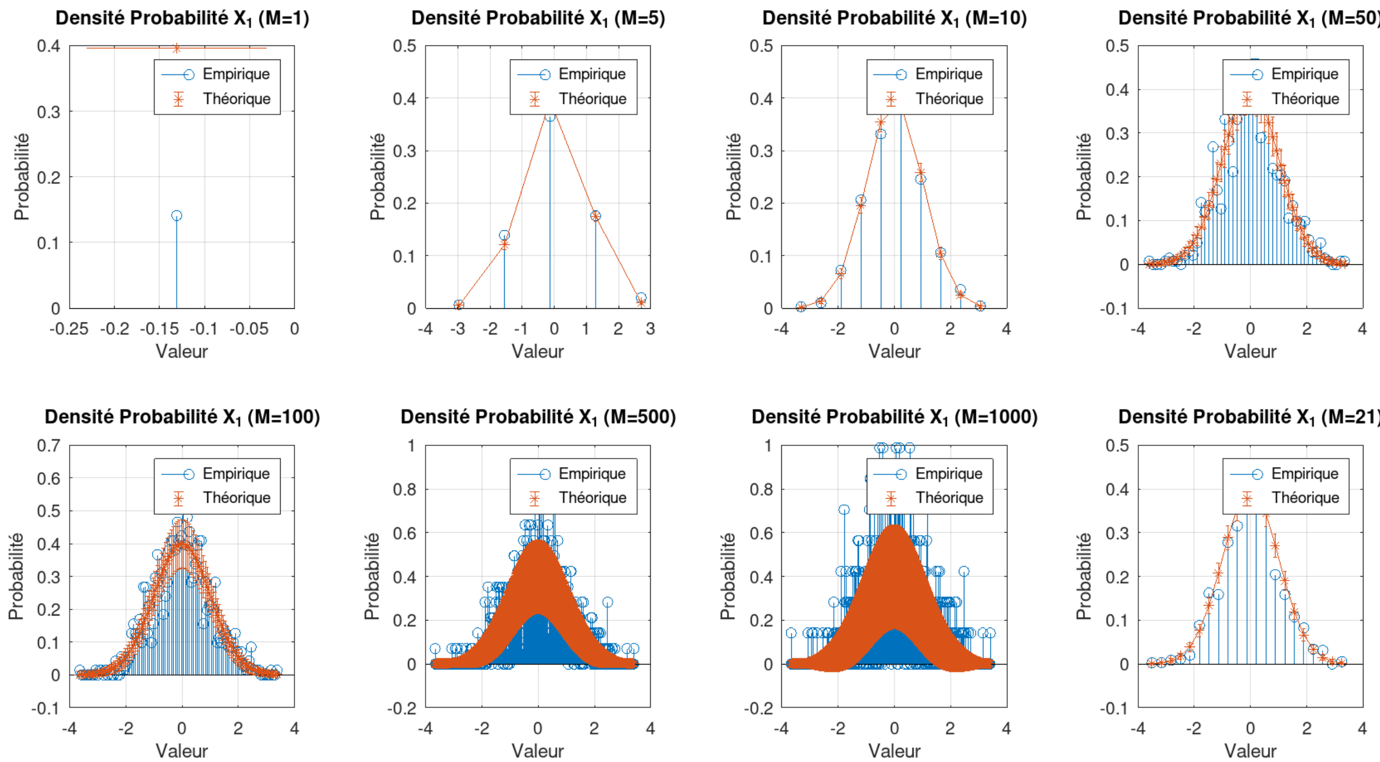


FIGURE 3 – Variation du nombre de classe M

□

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

On observe comme prévu une augmentation de la variance de l'estimation (quantifiée par les barres rouges) lorsque M augmente (donc Δx diminue). On observe dans l'autre cas etrême un écrt significatif entre la valeur théorique et la valeur estimée. □

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

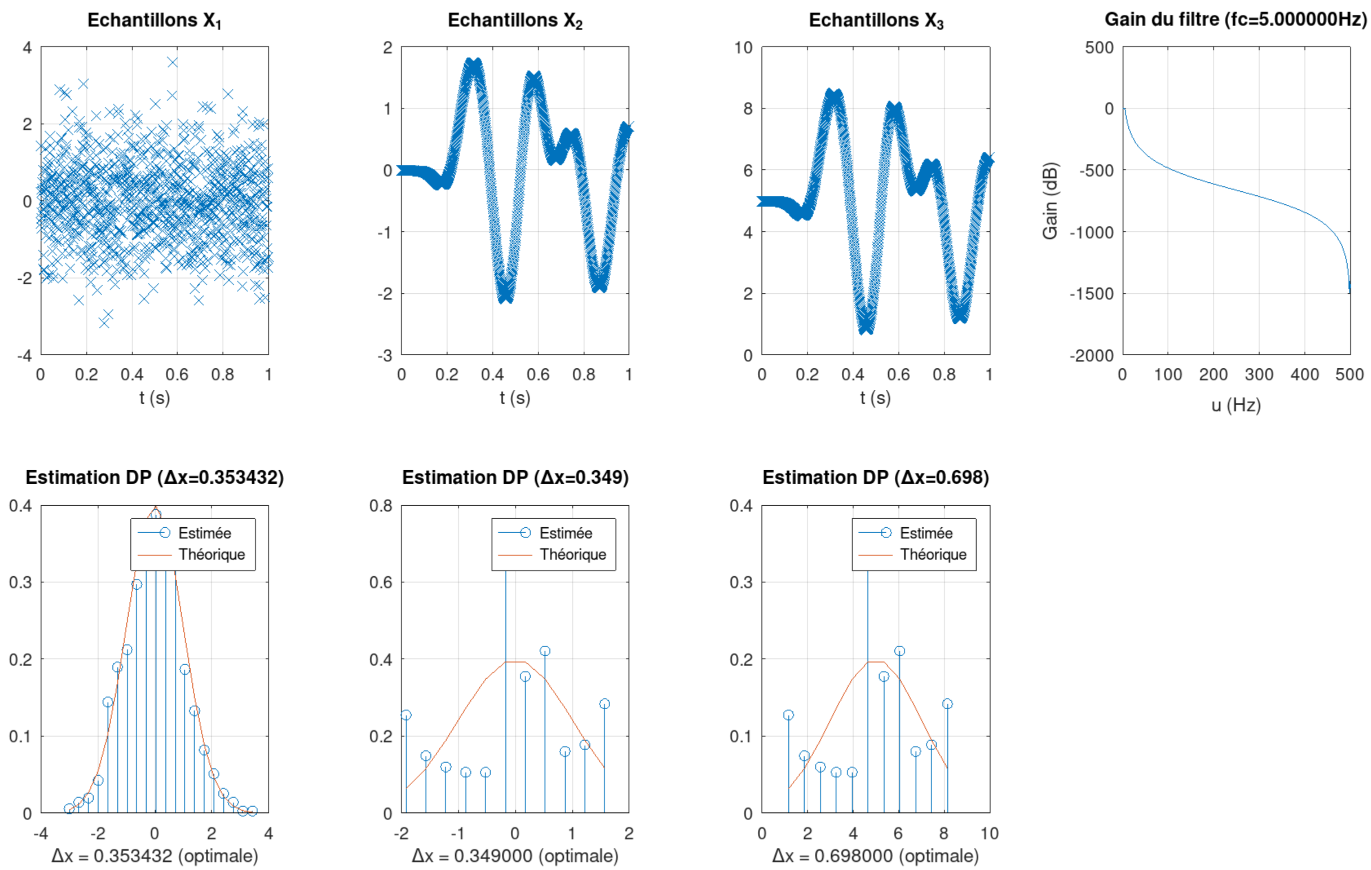


FIGURE 4 – Variation de la fréquence de coupure f_c

□

b) Le signal $x_2(t)$ est il gaussien ? Justifiez votre réponse (on pourra par exemple calculer le

réponse ci-dessous

Kurtosis sur la série temporelle $(x_2[n])_{n=1,...N}$). Au vu de la forme de la densité de probabilité au regarde de la densité théorique, on peut déduire facilement que le signal n'est pas gaussien. On peut également le confirmer en calculant le Kurtosis de cette estimation. Un signal gaussien a un Kurtosis valant 3. Le signal considéré lui a un Kurtosis égal à 2.0373. □

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

La différence vient à cause du filtrage. En effet, le filtrage passe-bas réduis la bande spectrale B du signal. Le rayon de corrélation augmente alors en conséquence. Les échantillons sont alors de plus en plus corrélés. Si le nombre d'échantillons est faible, on n'observe plus alors d'échantillons décorrelés et l'hypothèse d'ergodicité n'est alors plus respectée. L'estimation échoue alors.

Pour contrer l'effet du filtrage, on peut alors augmenter le nombre d'échantillons observés de sorte à continuer d'observer des échantillons décorrelés même si le rayon de corrélation augmente.

On simule ci-dessous ce cas avec $N = 1e5$ et $B = 5Hz$. On retrouve alors la bonne estimation. $\square$

figures ci-dessous

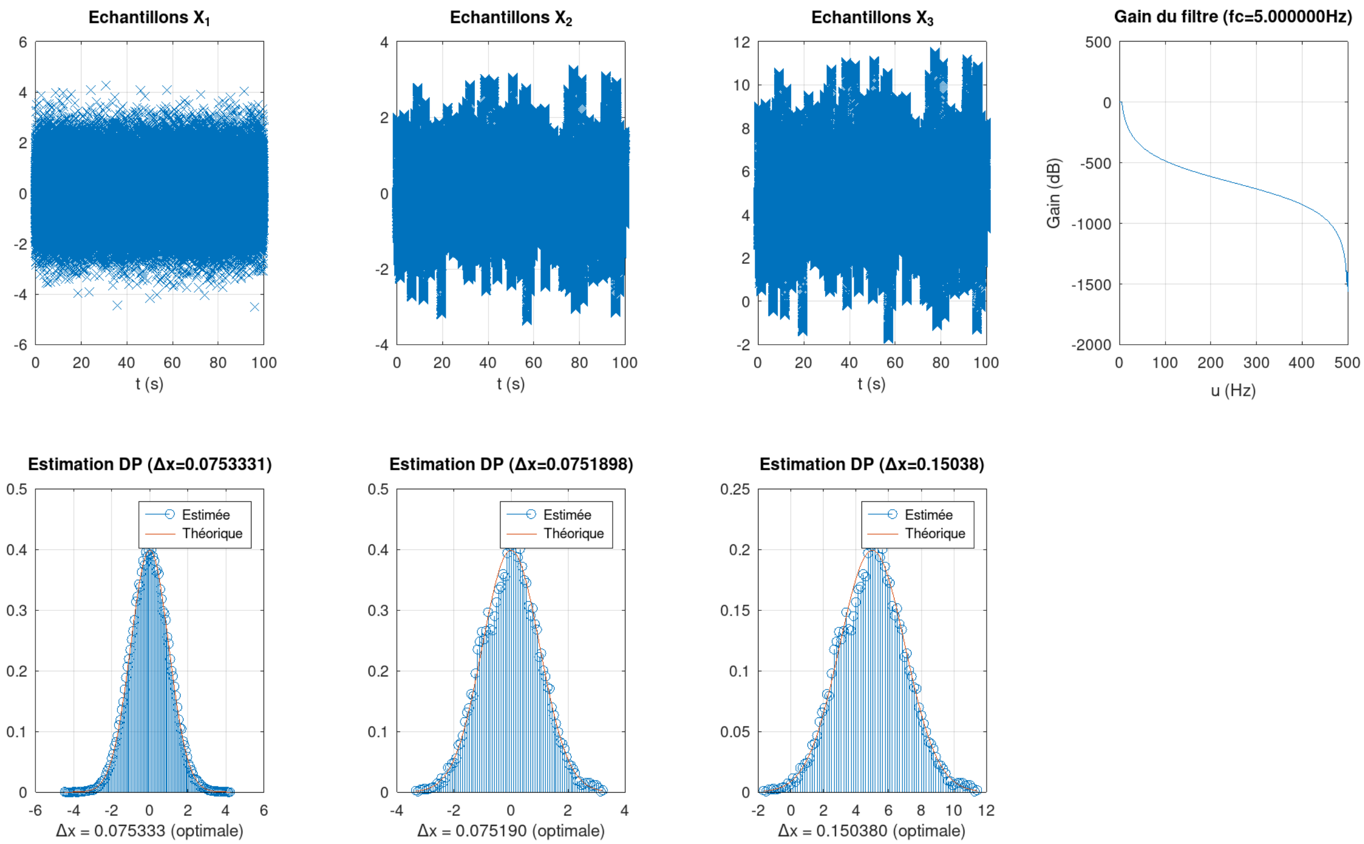


FIGURE 5 – Augmentation du nombre d'échantillons N avec filtrage très sélectif

$\square$

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

On nous dit dans l'énoncé que le retard à l'origine de notre signal carré est géré par τ qui est une variable aléatoire. Celà signifie concrètement que pour plusieurs essais de notre signal carré, on peut ne pas avoir la même valeur pour un instant t donné. Celà est due au fait qu'on ne "démarre" pas le signal au même endroit de notre période.
Notre signal carré est donc un signal aléatoire. □

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

On échantillone a bien plus que $2\nu_0$ donc il n'y aura donc pas de repliement spectral. Le signal étant de carré d'amplitude ± 1 et de rapport cyclique 0.5, sa moyenne est nulle. Le bruit gaussien ayant une moyenne nulle, notre signal a donc une moyenne de $\sigma = 0$.
On a donc $\mathbb{P}\{z \in [-0.5, 0.5]\} = 2\mathbb{P}\{z \in [0, 0.5]\}$

- Si $y = 1$:

$$\text{On a } \mathbb{P}\{z \in [0, 0.5]\} = \mathbb{P}\{x < -0.5\}$$

- si $y = -1$:

$$\text{On a } \mathbb{P}\{z \in [0, 0.5]\} = \mathbb{P}\{x > 0.5\}$$

Comme x est un bruit blanc et que le rapport.
$\mathbb{P}\{x < -0.5\} = \mathbb{P}\{x > 0.5\}$
Le rapport cyclique est de 0,5 ces deux cas sont similaires en terme de probabilités, en sommant on obtient finalement
$\mathbb{P}\{z \in [-0.5, 0.5]\} = \mathbb{P}\{x < -0.5\}$
On pose T la variable réduite :
$\mathbb{P}\{T < \frac{-0.5}{\sigma_B}\} = 0,5$
Et d'après le tableau $\sigma_B = 1/(2 * 3) = 1/6$

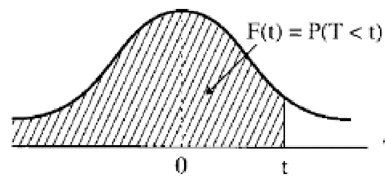
En utilisant le tableau de la loi normale ci dessous

Fonction de répartition de la loi centrée-réduite de Laplace-Gauss

$T \sim N(0,1)$

Probabilité d'une valeur inférieure à t :

$$F(t) = P(T < t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2}\, dx$$



t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7290	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0 9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0 9162	0 9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9649	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9779	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0 9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

FIGURE 6 – Densité de probabilité d'une loi normale

□

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

réponse ci-dessous

On raisonne avec la largeur a mis hauteur.

$$p_{\mathbf{z}}(0) = \frac{1}{2}(p_{\mathbf{x}}(1) + p_{\mathbf{x}}(-1)) = p_{\mathbf{x}}(1)$$

On peut maintenant utiliser les propriétés de la mi hauteur de la gaussienne sachant que :

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

On a donc $\sigma_b = \frac{2}{2.35} = 0.85$

□

Afficher sur une même figure, les deux densités correspondantes.

figures ci-dessous

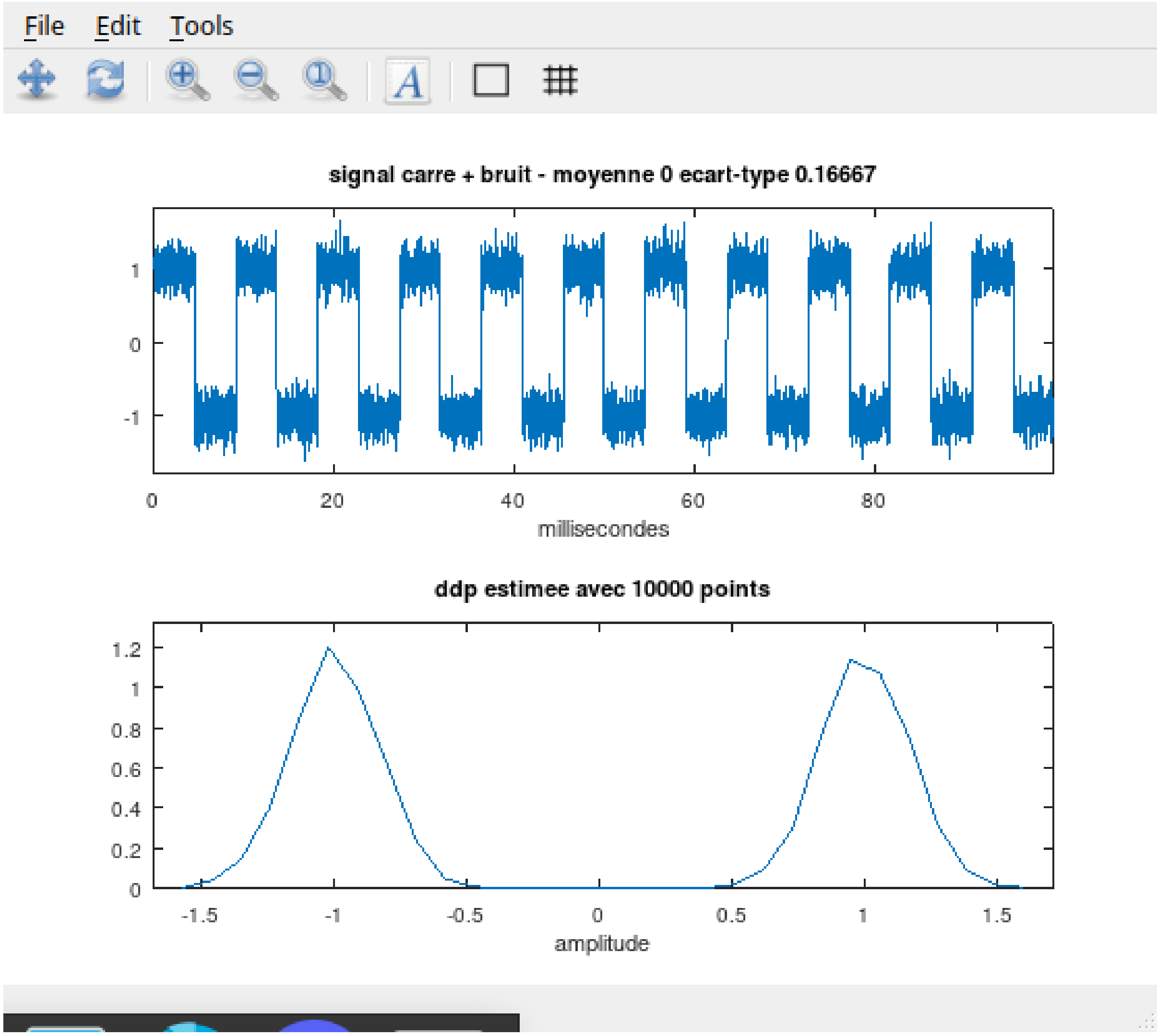


FIGURE 7 – Densité de probabilité d'une loi normale

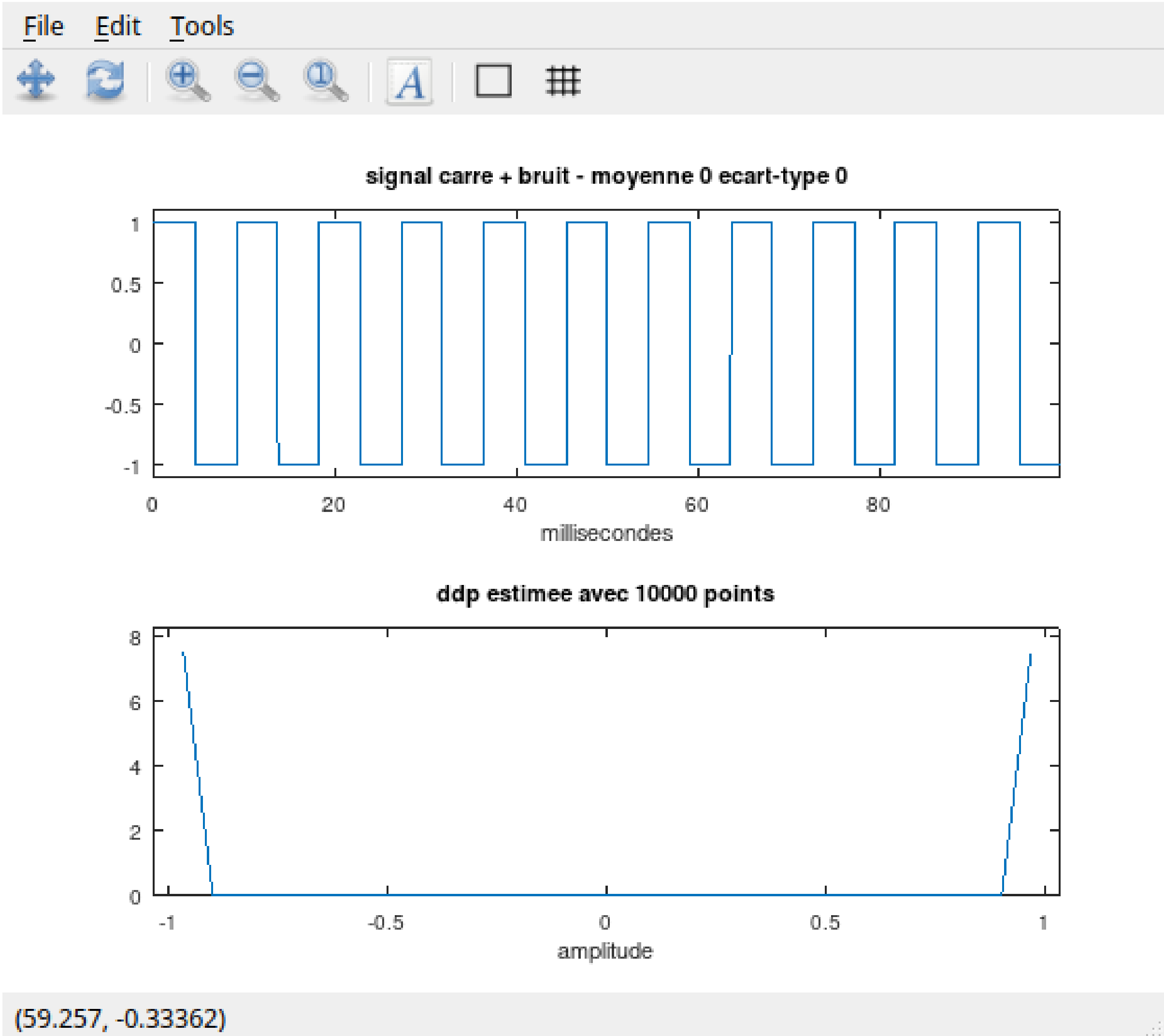


Figure 8 – Densité de probabilité d'une loi normale