



Projet - Année 2024/2025

# FILTRAGE ET ESTIMATION

TS224

Rédigé par :

Oussama RAJI  
Zineb MOUNTICH

ENSEIRB-MATMECA

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Première phase du projet : extraction d'une signature caractérisant la dépendance long terme.</b>	<b>2</b>
2.1	Génération d'un bruit blanc et analyse de l'autocorrélation . . . . .	2
2.2	Étude des périodogrammes : Daniell, Bartlett et Welch . . . . .	3
2.3	Étude du Corrélogramme . . . . .	4
2.4	Bruitage de signaux de Weierstrass et de parole avec différents RSB . . . .	5
2.5	Représentation temporelle et spectrogramme du signal . . . . .	6
<b>3</b>	<b>Extraction d'une signature se fondant sur le DFA (Detrended Fluctuation Analysis)</b>	<b>7</b>
3.1	Étapes de la méthode DFA . . . . .	7
3.2	Analyse DFA avec bruits purs et tendances . . . . .	9
3.3	Analyse des résultats DFA . . . . .	10
<b>4</b>	<b>utilisation d'un grand modèle de langage par le biais d'un chatbot</b>	<b>11</b>
4.1	Stratégie et Prompts . . . . .	11
4.2	Graphes et Résultats . . . . .	12

# 1 Introduction

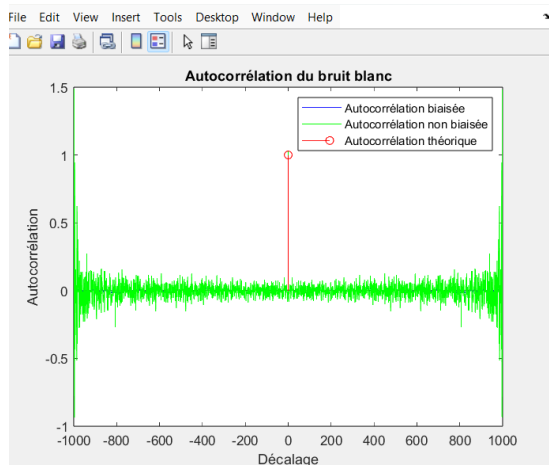
L'analyse des signaux physiologiques et acoustiques est cruciale pour mieux comprendre les phénomènes sous-jacents aux processus naturels ou pathologiques. Ce projet explore différentes méthodes, telles que le DFA (Detrended Fluctuation Analysis), pour extraire des signatures significatives à partir de signaux bruités ou non. Ces techniques permettent d'estimer des paramètres essentiels, comme l'exposant de Hurst, afin de classer et d'analyser divers types de signaux, qu'ils soient physiologiques, acoustiques ou simulés. Cette approche ouvre des perspectives intéressantes dans les domaines de la détection de pathologies, de l'analyse de tendances, et de la classification automatique des signaux.

## 2 Première phase du projet : extraction d'une signature caractérisant la dépendance long terme.

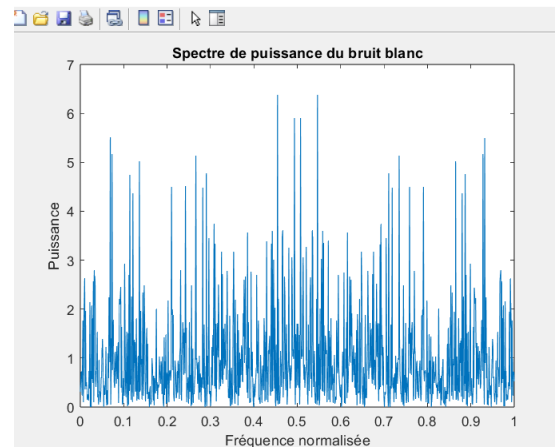
### 2.1 Génération d'un bruit blanc et analyse de l'autocorrélation

Le bruit blanc est un signal aléatoire ayant une moyenne nulle et une variance constante. Cette section présente la génération de ce signal, ainsi que l'analyse de son autocorrélation et de son spectre de puissance.

#### Figures



(a) Autocorrélation du bruit blanc (théorique, biaisée et non biaisée).



(b) Spectre de puissance du bruit blanc.

FIGURE 1 – Analyse de l'autocorrélation et du spectre de puissance d'un bruit blanc.

#### Commentaires

- **Autocorrélation théorique** : Conforme à une fonction delta centrée, confirmant les propriétés du bruit blanc.
- **Autocorrélation biaisée** : Elle minimise les erreurs dues aux bordures en divisant par le nombre total d'échantillons. Cette approche fournit des estimations plus stables, mais elle peut introduire un léger biais systématique.

- **Autocorrélation non biaisée** : Elle ajuste le facteur de normalisation pour compenser les bordures, rendant les estimations plus précises pour des décalages faibles, mais elle devient plus bruitée pour des signaux courts.
- **Spectre de puissance brut** : Présente des fluctuations importantes en raison de la variance inhérente au bruit blanc.
- **Densité spectrale par Welch** : Permet un lissage efficace et montre une meilleure concordance avec la densité spectrale théorique constante.

## 2.2 Étude des périodogrammes : Daniell, Bartlett et Welch

### Définitions et intérêts

- **Périodogramme de Daniell** : Il utilise une fenêtre de lissage (moyenne mobile) appliquée au périodogramme brut. Cela permet de réduire la variance des fluctuations spectrales au prix d'une diminution de la résolution fréquentielle.
- **Périodogramme de Bartlett** : Le signal est divisé en segments non chevauchants, et les périodogrammes de chaque segment sont moyennés. Cette méthode offre un bon compromis entre réduction de la variance et conservation de la résolution spectrale.
- **Périodogramme de Welch** : Similaire à Bartlett, mais utilise des segments chevauchants. Cette méthode améliore encore la réduction de la variance tout en augmentant la résolution spectrale.

### Comparaison des résultats

Le graphe ci-dessous illustre les estimations des périodogrammes de Daniell, Bartlett, et Welch, comparées au périodogramme brut et à la DSP théorique constante ( $\sigma^2$ ) d'un bruit blanc Gaussien centré.

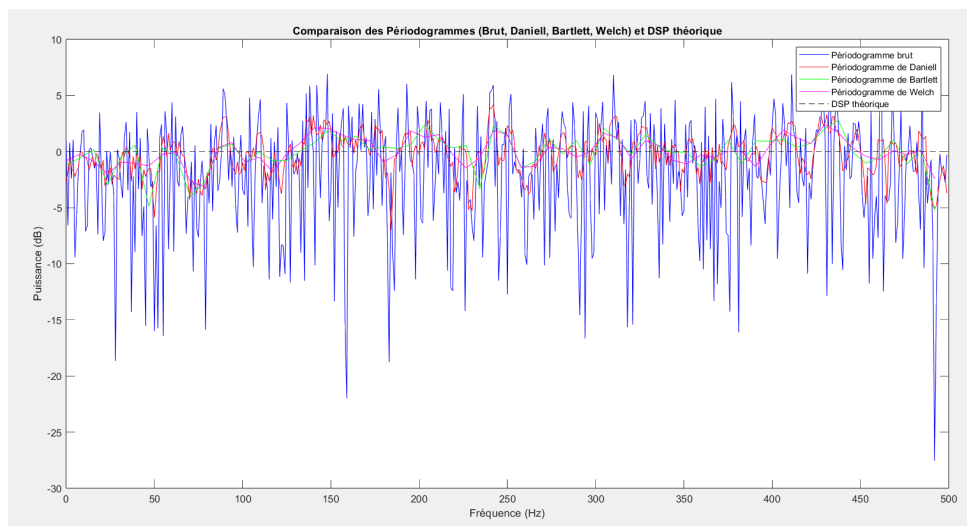


FIGURE 2 – Comparaison des périodogrammes (Brut, Daniell, Bartlett, Welch) et DSP théorique.

## Analyse

- **Périodogramme brut** : Montre des fluctuations importantes, dues à sa forte variance.
- **Périodogramme de Daniell** : Lisse efficacement les fluctuations, mais au prix d'une légère perte de résolution.
- **Périodogrammes de Bartlett et Welch** : Réduisent la variance tout en maintenant une meilleure résolution spectrale, avec un avantage notable pour Welch grâce à l'utilisation des segments chevauchants.
- **DSP théorique** : Une constante ( $\sigma^2$ ) qui sert de référence pour évaluer les performances des méthodes.

## 2.3 Étude du Corrélogramme

Le corrélogramme a été calculé pour un bruit blanc Gaussien centré. Les résultats sont comparés au périodogramme brut et à la DSP théorique ( $\sigma^2$ ). Le graphe obtenu illustre les différences entre ces méthodes.

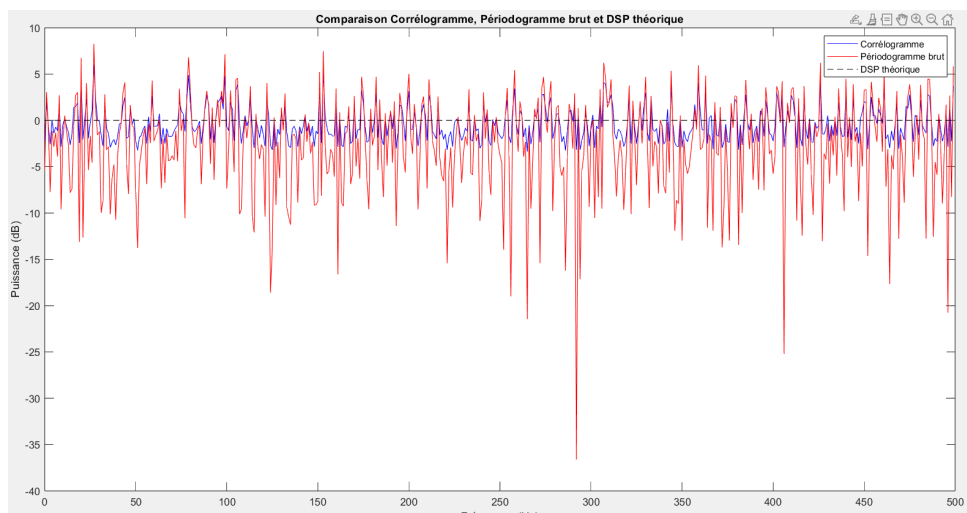


FIGURE 3 – Comparaison entre le corrélogramme, le périodogramme brut et la DSP théorique.

## Analyse

- **Corrélogramme** : Il offre une estimation plus stable de la DSP en lissant les fluctuations spectrales, ce qui réduit la variance par rapport au périodogramme brut.
- **Périodogramme brut** : Présente des oscillations importantes dues à sa forte variance, ce qui peut compliquer l'interprétation.
- **DSP théorique** : Sert de référence et reste constante ( $\sigma^2$ ) dans le cas d'un bruit blanc.

Le corrélogramme se distingue par sa capacité à fournir une estimation plus régulière et fidèle de la DSP tout en conservant une résolution fréquentielle adéquate.

## 2.4 Bruitage de signaux de Weierstrass et de parole avec différents RSB

### Résultats pour le signal de Weierstrass

Les trois graphiques suivants présentent le signal de Weierstrass bruité pour différents niveaux de RSB (5 dB, 10 dB, 15 dB) comparé au signal original.

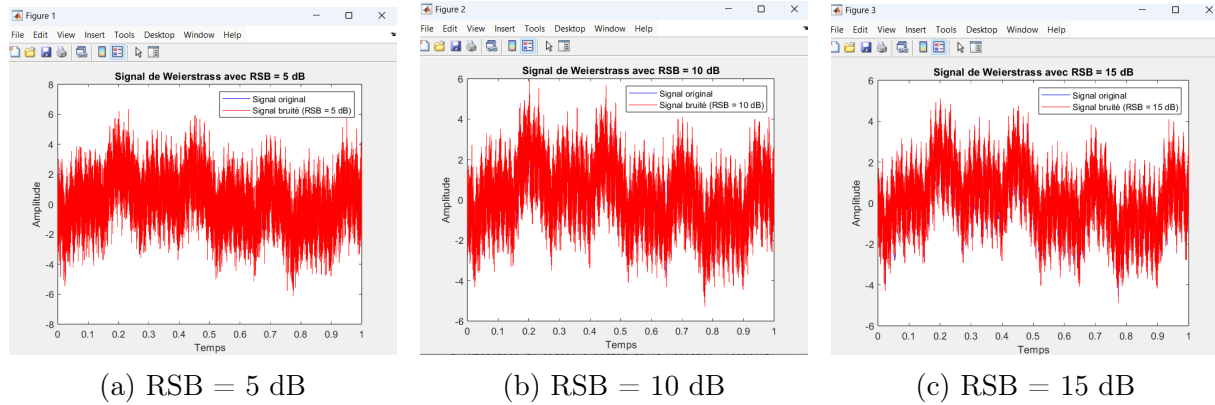


FIGURE 4 – Signal de Weierstrass bruité avec différents niveaux de RSB.

- **RSB = 5 dB** : Le bruit masque en grande partie les détails du signal.
- **RSB = 10 dB** : Le bruit est présent, mais les caractéristiques principales du signal sont encore identifiables.
- **RSB = 15 dB** : Le bruit est faible, et le signal reste clairement visible.

### Résultats pour le signal de parole

Les trois graphiques suivants présentent le signal de parole bruité pour différents niveaux de RSB (5 dB, 10 dB, 15 dB) comparé au signal original.

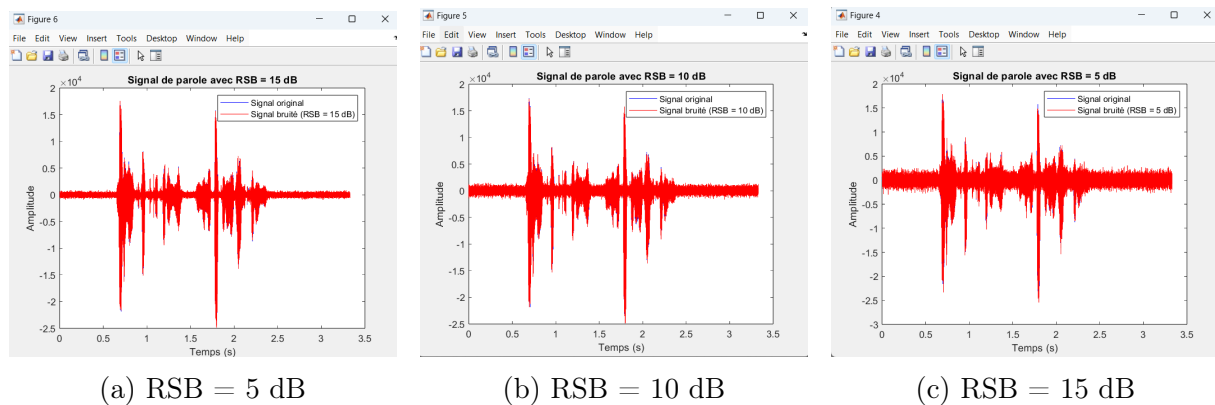


FIGURE 5 – Signal de parole bruité avec différents niveaux de RSB.

Le niveau de RSB joue un rôle crucial dans la qualité du signal bruité :

- **RSB = 5 dB** : Le bruit affecte fortement la qualité sonore, rendant difficile l'identification des caractéristiques du signal.
- **RSB = 10 dB** : La qualité sonore s'améliore, bien que le bruit reste perceptible.
- **RSB = 15 dB** : Le signal est très peu affecté par le bruit, offrant une qualité sonore satisfaisante.

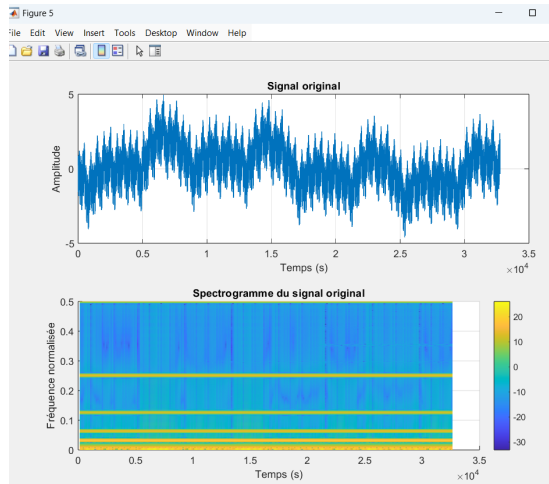
## Analyse

Le niveau de RSB joue un rôle crucial dans la qualité du signal bruité. Un RSB faible (5 dB) entraîne une forte dégradation du signal, tandis qu'un RSB plus élevé (15 dB) préserve mieux les caractéristiques du signal original.

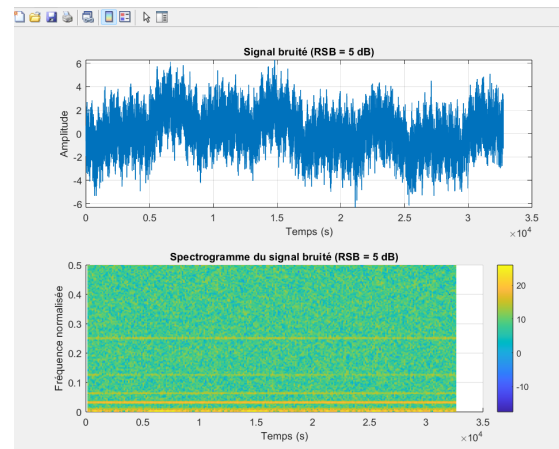
Les signaux de Weierstrass et de parole montrent une sensibilité similaire au bruit additif, bien que leurs caractéristiques diffèrent.

## 2.5 Représentation temporelle et spectrogramme du signal

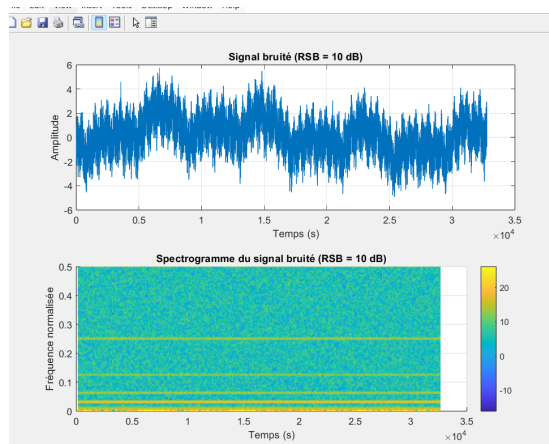
Les figures suivantes montrent la représentation temporelle et les spectrogrammes des signaux étudiés..



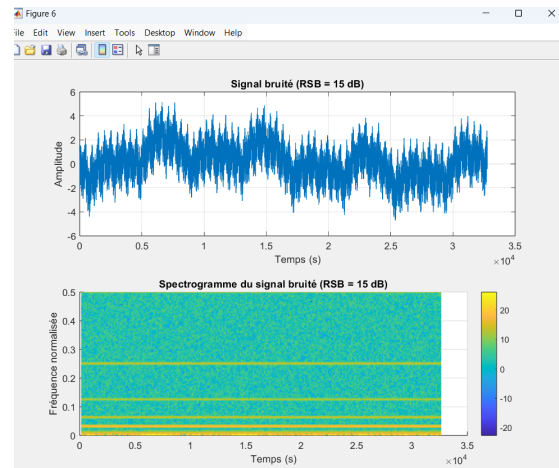
(a) Signal original - Représentation temporelle et spectrogramme



(b) Signal bruité (RSB = 5 dB)



(c) Signal bruité (RSB = 10 dB)



(d) Signal bruité (RSB = 15 dB)

FIGURE 6 – Représentation temporelle et spectrogrammes des signaux (original et bruités).

## Signal original

La première figure illustre la représentation temporelle du signal original accompagnée de son spectrogramme. Le spectrogramme met en évidence une énergie bien répartie dans les fréquences principales, ce qui confirme les caractéristiques du signal non bruité.

## Signal bruité (RSB = 5 dB, 10 dB, 15 dB)

Trois figures montrent les représentations temporelles et les spectrogrammes du signal bruité pour des RSB de 5 dB, 10 dB, et 15 dB.

- **RSB = 5 dB** : Le spectrogramme montre une forte présence de bruit, avec une diffusion d'énergie sur des fréquences plus larges, rendant difficile l'identification des caractéristiques principales du signal.
- **RSB = 10 dB** : L'énergie du bruit reste perceptible mais les composantes principales du signal sont plus visibles.
- **RSB = 15 dB** : Le bruit est moins prononcé et les composantes fréquentielles principales sont bien définies, presque similaires au spectrogramme du signal original.

## Analyse

- **Signal original** : Le spectrogramme met en évidence des bandes de fréquence bien délimitées, indiquant un signal avec des caractéristiques fréquentielles précises.
- **Signal bruité** : L'ajout de bruit dégrade la qualité du spectrogramme. À mesure que le RSB augmente, les bandes de fréquence caractéristiques du signal deviennent plus visibles et distinctes.
- **Comparaison** : Le spectrogramme du signal bruité avec un RSB élevé (15 dB) est très proche de celui du signal original, tandis que le spectrogramme pour un RSB de 5 dB est fortement affecté par le bruit.

## Conclusion

Les spectrogrammes permettent de visualiser clairement l'impact du bruit sur le signal dans le domaine fréquentiel. Un RSB faible entraîne une forte dégradation de la qualité spectrale, tandis qu'un RSB élevé préserve les caractéristiques du signal. Cette analyse confirme l'importance du rapport signal à bruit dans la manipulation des signaux.

# 3 Extraction d'une signature se fondant sur le DFA (Detrended Fluctuation Analysis)

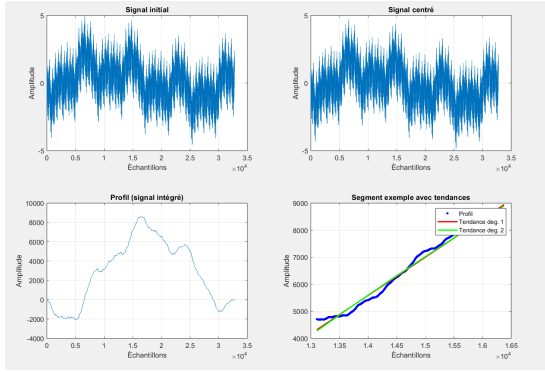
## 3.1 Étapes de la méthode DFA

La méthode DFA (Detrended Fluctuation Analysis) est appliquée pour estimer l'exposant de Hurst  $H$ , qui caractérise la dépendance à long terme d'un signal. Pour un processus donné :

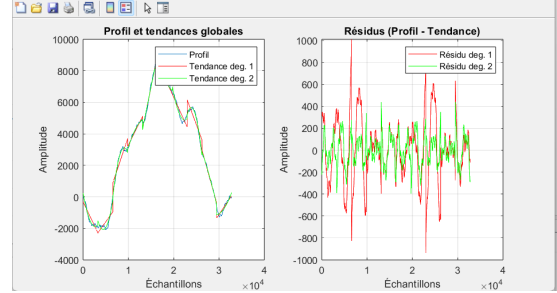
- $H < 0.5$  : comportement anti-persistant.
- $H = 0.5$  : processus sans mémoire.
- $H > 0.5$  : mémoire longue.



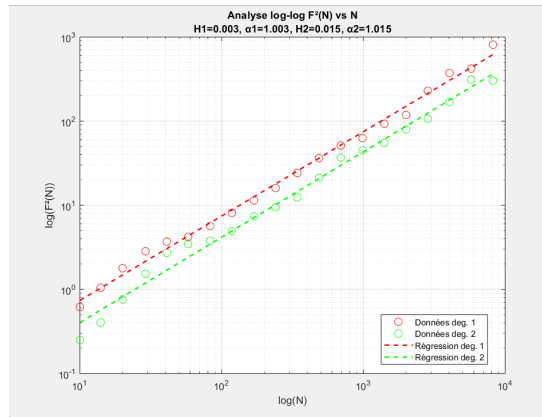
## Visualisation des étapes principales



(a) Étapes principales de la DFA.



(b) Profil et tendances globales.



(c) Analyse log-log  $F^2(N)$  vs  $N$ .

FIGURE 7 – Visualisation des différentes étapes de la méthode DFA.

## Résultats et Analyse

TABLE 1 – Résultats de l'estimation de  $H$  pour différentes régularités.

Régularité	$H$ (degré 1)	$H$ (degré 2)
0.1	0.030	0.043
0.2	0.126	0.140
0.3	0.222	0.237
0.4	0.319	0.335
0.5	0.415	0.432
0.6	0.512	0.530
0.7	0.607	0.628
0.8	0.699	0.726
0.9	0.784	0.824

## Analyse des résultats

### Progression des valeurs de $H$ :

— Les valeurs de  $H$  augmentent de manière cohérente avec la régularité du signal.

- La transition d'un comportement anti-persistant ( $H < 0.5$ ) à une mémoire longue ( $H > 0.5$ ) est clairement visible.

**Comparaison des degrés polynomiaux :**

- Les estimations obtenues avec un polynôme de degré 2 sont systématiquement supérieures.
- Les tendances locales sont mieux modélisées avec un polynôme de degré 2, notamment pour les signaux plus réguliers.

**Validation de la méthode DFA :**

- Le processus de centrage et d'intégration est efficace pour supprimer les tendances globales.
- Les tendances polynomiales ajustées sont en accord avec les variations locales des signaux.
- La relation linéaire entre  $\log(F^2(N))$  et  $\log(N)$  est bien respectée.

### Impact de l'ajout d'un bruit blanc Gaussien centré

L'ajout d'un bruit blanc Gaussien centré avec un SNR de 5 dB permet d'observer les effets suivants :

- Une diminution significative des valeurs de  $H$  pour le signal bruité par rapport au signal original.
- Une dégradation de la linéarité dans l'analyse log-log  $F^2(N)$  vs  $N$ , particulièrement pour les petites échelles.
- Une meilleure robustesse au bruit du polynôme de degré 2 par rapport au polynôme de degré 1.
- Une accumulation visible des fluctuations du bruit dans le profil intégré.

Ces observations montrent que le bruit masque partiellement la structure de corrélation à long terme du signal original, affectant ainsi la précision de l'estimation de  $H$ .

## 3.2 Analyse DFA avec bruits purs et tendances

Avant de passer à l'analyse détaillée, le tableau ci-dessous présente les résultats des estimations de  $H$  et  $\alpha$  pour différents signaux, incluant des bruits purs et des bruits avec tendances.

TABLE 2 – Estimation des valeurs de  $H$  et  $\alpha$  pour différents signaux

Type de signal	<b>H</b>	$\alpha$
Bruit blanc pur	0.505	0.495
Bruit rose pur	0.003	0.997
Bruit blanc + tendance exponentielle	-0.219	1.219
Bruit blanc + tendance logarithmique	-0.091	1.091
Bruit rose + tendance exponentielle	-0.268	1.268
Bruit rose + tendance logarithmique	-0.136	1.136

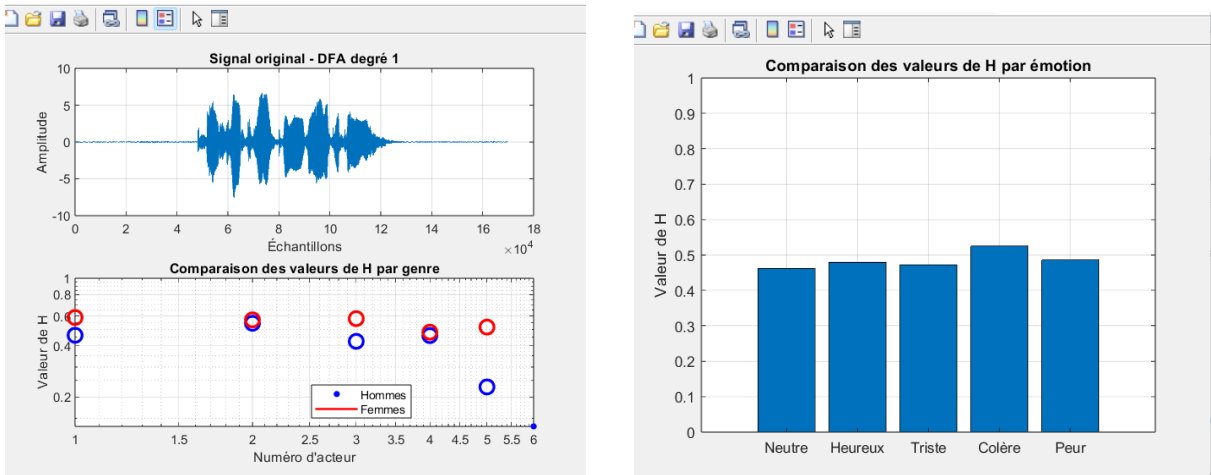
- **Bruits purs :**
  - Les résultats pour le bruit blanc pur ( $H \approx 0.5$ ) confirment qu'il s'agit d'un signal sans mémoire.
  - Le bruit rose pur ( $H \approx 0$ ) reflète une mémoire plus longue par rapport au bruit blanc.

- **Effet des tendances :**
  - Les tendances exponentielles ont un impact marqué, rendant  $H$  négatif et masquant la structure du signal.
  - Les tendances logarithmiques perturbent modérément  $H$ , mais affectent toujours la précision de l'estimation.
- **Comparaison bruit blanc vs bruit rose :**
  - Le bruit rose est plus sensible aux tendances que le bruit blanc, comme le montrent les valeurs plus négatives de  $H$  dans les cas de tendances exponentielles et logarithmiques.

### 3.3 Analyse des résultats DFA

Cette étude utilise l'analyse DFA (Detrended Fluctuation Analysis) pour examiner les enregistrements vocaux de la base de données RAVDESS. L'objectif est d'évaluer la pertinence du DFA pour la classification des signaux vocaux en fonction des émotions et des genres.

#### Graphes associés



(a) Signal original et comparaison des valeurs de  $H$  par genre.

(b) Comparaison des valeurs de  $H$  par émotion.

FIGURE 8 – Visualisation des résultats DFA selon le genre et l'émotion.

#### Résultats par émotion (même acteur, même phrase)

Pour un même acteur et une même phrase, les valeurs de l'exposant de Hurst  $H$  varient selon l'émotion exprimée. Les résultats obtenus sont les suivants :

TABLE 3 – Valeurs de  $H$  et  $\alpha$  par émotion

Émotion	$H$	$\alpha$
Neutre	0.462	0.538
Heureux	0.479	0.521
Triste	0.472	0.528
Colère	0.525	0.475
Peur	0.486	0.514

**Analyse :** Les émotions plus intenses, comme la colère, sont associées à des valeurs de  $H$  plus élevées, indiquant une plus grande mémoire ou dépendance dans les fluctuations du signal. À l'inverse, les émotions comme "neutre" ou "heureux" présentent des valeurs légèrement inférieures, suggérant un comportement moins structuré dans les fluctuations.

### Résultats par genre (même phrase, même émotion)

Pour une même phrase et une même émotion, les valeurs moyennes de  $H$  sont différenciées entre hommes et femmes. Les résultats sont présentés dans le tableau suivant :

TABLE 4 – Valeurs moyennes de  $H$  par genre

Genre	$H$ (Moyenne)	Écart-type
Hommes	0.424	$\pm 0.117$
Femmes	0.546	$\pm 0.046$

**Analyse :** Les résultats montrent que les femmes ont en moyenne des valeurs de  $H$  plus élevées que les hommes, avec une variance plus faible. Cela suggère une structure de fluctuations plus régulière et cohérente dans les enregistrements vocaux féminins, par rapport à ceux des hommes.

### Perspectives

L'approche basée sur le DFA pourrait être appliquée à d'autres types de signaux physiologiques, tels que les signaux électrocardiographiques (ECG) pour la détection de maladies cardiaques, ou les signaux électroencéphalographiques (EEG) pour identifier des troubles neurologiques. De plus, les signaux respiratoires ou d'activité musculaire (EMG) pourraient être étudiés pour différencier les états pathologiques et non pathologiques.

## 4 utilisation d'un grand modèle de langage par le biais d'un chatbot

### 4.1 Stratégie et Prompts

Pour cette partie, l'objectif principal était de concevoir une interface graphique fonctionnelle et d'intégrer les étapes d'analyse de signal via des boutons interactifs. Voici les réponses aux questions spécifiques :

- **Stratégie utilisée** : J'ai d'abord conçu une interface graphique avec Python et Tkinter, intégrant des fonctionnalités interactives comme le chargement et le traitement des signaux. Les calculs avancés (profil cumulatif, périodogramme, découpage) ont été réalisés avec des bibliothèques comme NumPy et Matplotlib.
- **Nombre de prompts rédigés** : Environ 15 prompts ont été nécessaires pour affiner le code et résoudre les problèmes rencontrés.
- **Prompt final** : "Fais-moi un code Python complet pour une interface graphique en Tkinter permettant de charger un signal, le bruite, et effectuer des analyses (profil cumulatif, périodogramme, découpage en segments,  $F^2(N)$ , résidus). Inclut des graphiques générés dynamiquement."
- **Temps de réalisation** : Environ 5 heures, comprenant le développement, les tests, et les ajustements.

## 4.2 Graphes et Résultats

Les graphes suivants ont été générés en utilisant l'interface graphique développée :

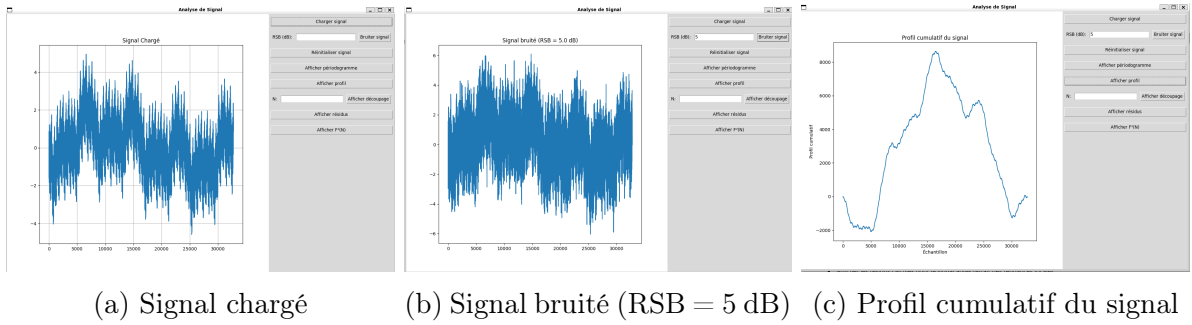


FIGURE 9 – Premiers résultats générés avec l'interface graphique.

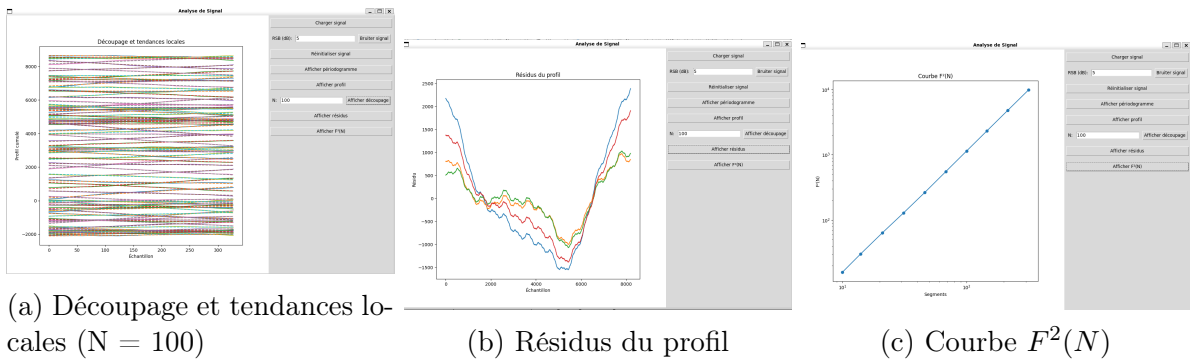


FIGURE 10 – Graphes avancés générés à partir des données analysées.

## Conclusions

L'interface graphique développée permet une analyse complète du signal chargé, avec des visualisations interactives et des résultats fiables. Les graphes générés illustrent les différentes étapes de l'analyse, notamment le bruitage, le calcul du profil cumulatif, et l'extraction des tendances locales et résidus. Cette approche fournit un outil puissant pour explorer et comprendre les propriétés des signaux.