

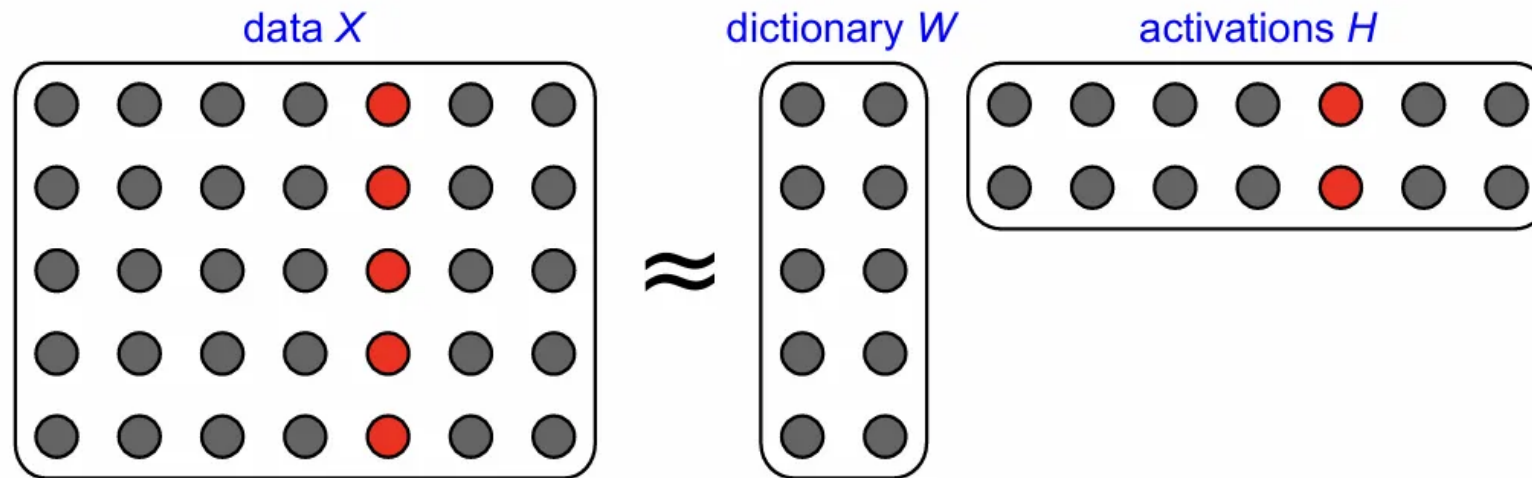


# Plan

---

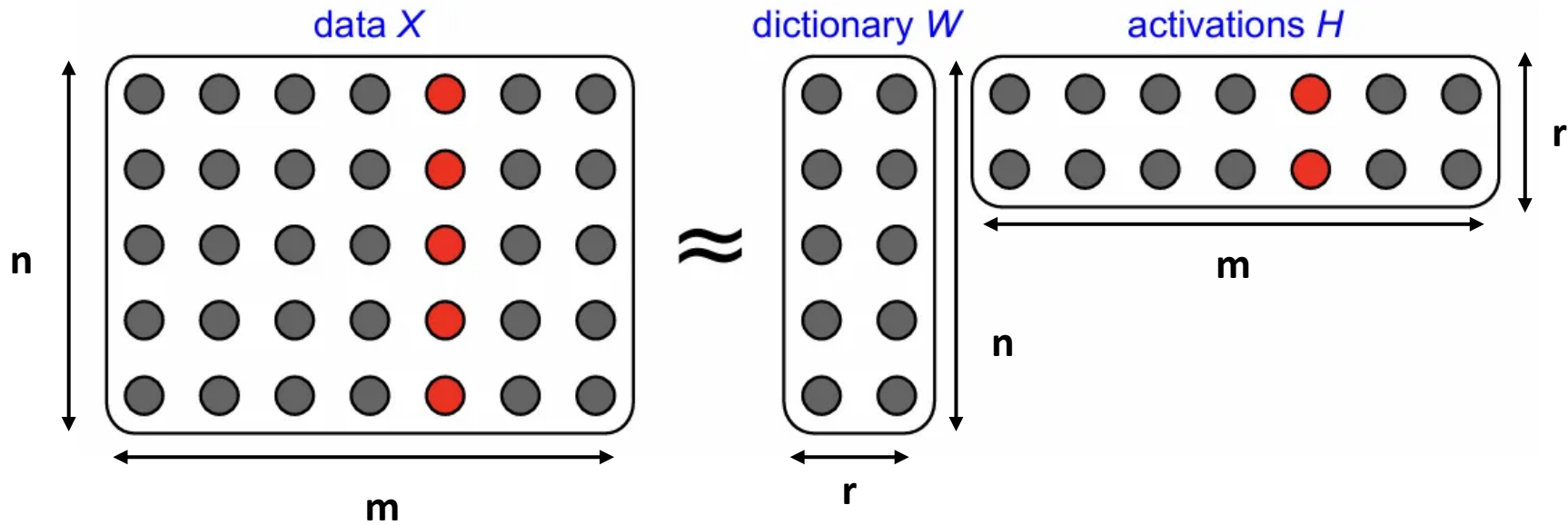
- **Principe de NMF**
- **Pratique de NMF**
  - Simulation sur une onde sinus
  - Échec sur le problème du mélange de musique et parole
- **Principe simple et Pratique de MNMF**
  - Principe générale
  - Simulation sur des mélange de parole
  - Simulation sur des mélange de musique
  - Expeience réelle

## **NMF: Factorisation en Matrice Non-Négative**



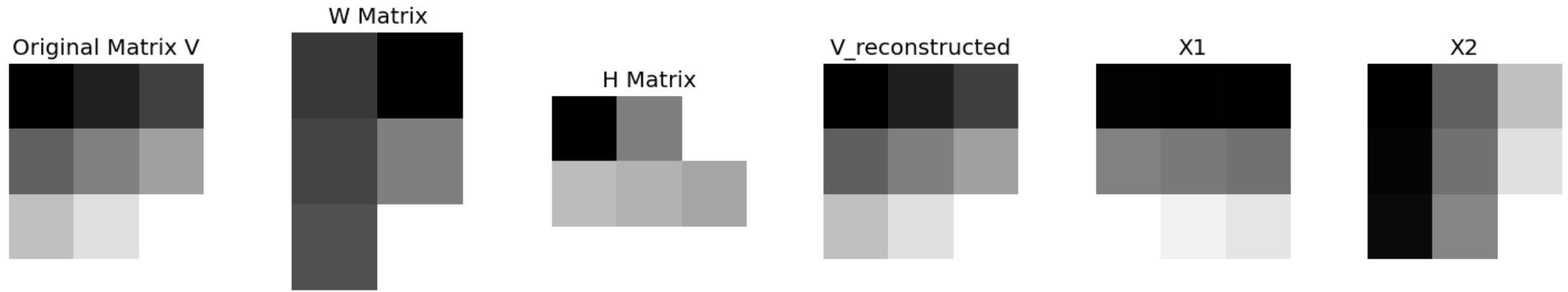
C'est une méthode de décomposition linéaire qui décompose une matrice non-négative  $X$  en le produit de deux matrices non-négatives.

# **NMF:** Factorisation en Matrice Non-Négative



# NMF: Factorisation en Matrice Non-Négative

---



un exemple sans signification réelle

## **NMF: Méthode d'optimisation**

---

$$\min_{\mathbf{W}, \mathbf{H} \geq 0} D(\mathbf{V} | \mathbf{WH}) = \sum_{fn} d([\mathbf{V}]_{fn} | [\mathbf{WH}]_{fn}),$$

## NMF: Méthode d'optimisation

$$\min_{\mathbf{W}, \mathbf{H} \geq 0} D(\mathbf{V} | \mathbf{WH}) = \sum_{fn} d([\mathbf{V}]_{fn} | [\mathbf{WH}]_{fn}),$$

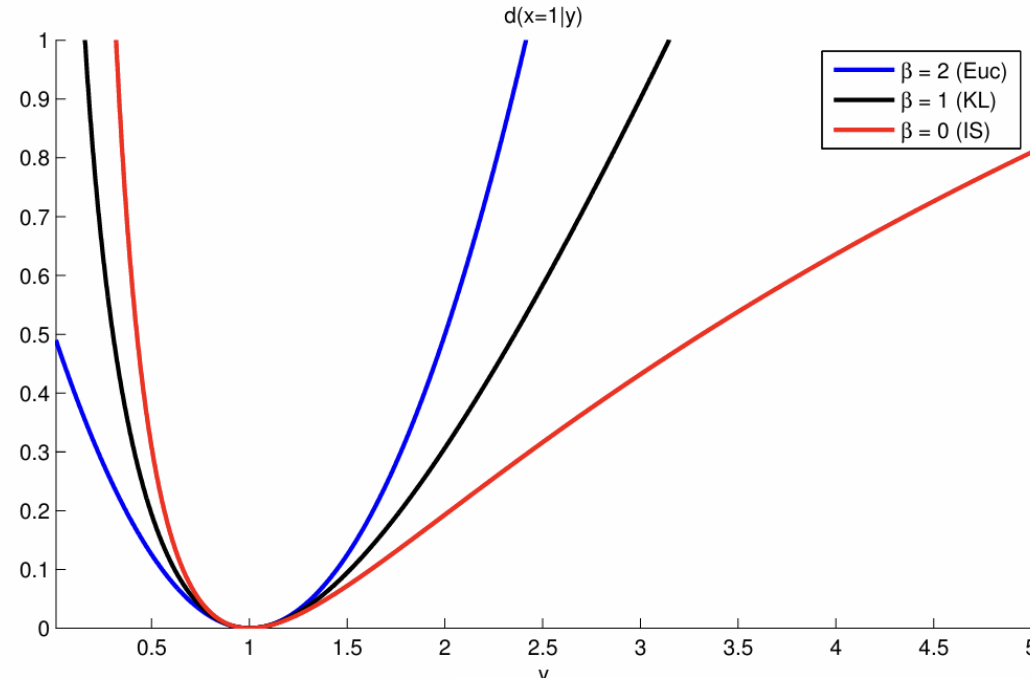
↓ **Beta-divergence**

$$\min D_E(V \| WH) = \frac{1}{2} \|V - WH\|_2^2$$

↑ **Distance euclidienne**

$$d_\beta(x|y) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{\beta(\beta-1)} (x^\beta + (\beta-1)y^\beta - \beta x y^{\beta-1}) & \beta \in \mathbb{R} \setminus \{0, 1\} \\ x \log \frac{x}{y} + (y-x) & \beta = 1 \quad \rightarrow \text{KL-divergence} \\ \frac{x}{y} - \log \frac{x}{y} - 1 & \beta = 0 \quad \rightarrow \text{IS-divergence} \end{cases}$$

# NMF: Méthode d'optimisation



$$\min_{\mathbf{W}, \mathbf{H} \geq 0} D(\mathbf{V} | \mathbf{WH}) = \sum_{fn} d([\mathbf{V}]_{fn} | [\mathbf{WH}]_{fn}),$$

$$\min D_E(V \| WH) = \frac{1}{2} \|V - WH\|_2^2$$

↓ **Beta-divergence**

↑ **Distance euclidienne**

$$d_\beta(x|y) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{\beta(\beta-1)} (x^\beta + (\beta-1)y^\beta - \beta x y^{\beta-1}) & \beta \in \mathbb{R} \setminus \{0, 1\} \\ x \log \frac{x}{y} + (y-x) & \beta = 1 \quad \blackrightarrow \text{KL-divergence} \\ \frac{x}{y} - \log \frac{x}{y} - 1 & \beta = 0 \quad \blackrightarrow \text{IS-divergence} \end{cases}$$



## **NMF: Méthode d'optimisation**

---

$$w_{ik} \leftarrow w_{ik} - \mu_{ik} \frac{\partial D_E(V \| WH)}{\partial w_{ik}}$$
$$h_{kj} \leftarrow h_{kj} - \eta_{kj} \frac{\partial D_E(V \| WH)}{\partial h_{kj}}$$

Algorithmes de descente de gradient  
couramment utilisés



Ne peut pas satisfaire **la contrainte de non-négativité de la matrice**, une optimisation sous contrainte est nécessaire.

## **NMF: Méthode d'optimisation**

---

$$w_{ik} \leftarrow w_{ik} - \mu_{ik} \frac{\partial D_E(V \| WH)}{\partial w_{ik}}$$

$$h_{kj} \leftarrow h_{kj} - \eta_{kj} \frac{\partial D_E(V \| WH)}{\partial h_{kj}}$$



$$\frac{\partial D_E(V \| WH)}{\partial w_{ik}} = -[(V - WH)H^T]_{ik}$$

$$\frac{\partial D_E(V \| WH)}{\partial h_{kj}} = -[W^T(V - WH)]_{kj}$$

**Distance euclidienne**

**Algorithmes de descente de gradient  
couramment utilisés**



**Ne peut pas satisfaire la contrainte de non-négativité de la matrice, une optimisation sous contrainte est nécessaire.**

# NMF: Méthode d'optimisation

$$w_{ik} \leftarrow w_{ik} - \mu_{ik} \frac{\partial D_E(V \| WH)}{\partial w_{ik}}$$

$$h_{kj} \leftarrow h_{kj} - \eta_{kj} \frac{\partial D_E(V \| WH)}{\partial h_{kj}}$$



$$\frac{\partial D_E(V \| WH)}{\partial w_{ik}} = -[(V - WH)H^T]_{ik}$$

$$\frac{\partial D_E(V \| WH)}{\partial h_{kj}} = -[W^T(V - WH)]_{kj}$$

Distance euclidienne



$$\mu_{ik} = \frac{w_{ik}}{[WHH^T]_{ik}}, \quad \eta_{kj} = \frac{h_{kj}}{[W^TWH]_{kj}}$$

Coefficient de descente de gradient



$$w_{ik} \leftarrow w_{ik} \frac{[VH^T]_{ik}}{[WHH^T]_{ik}}$$

$$h_{kj} \leftarrow h_{kj} \frac{[W^TV]_{kj}}{[W^TWH]_{kj}}$$

Algorithme d'optimisation multiplicative

Algorithmes de descente de gradient couramment utilisés



Ne peut pas satisfaire la **contrainte de non-négativité de la matrice**, une optimisation sous contrainte est nécessaire.

# **NMF: Méthode d'optimisation**

$$w_{ik} \leftarrow w_{ik} - \mu_{ik} \frac{\partial D_E(V \| WH)}{\partial w_{ik}}$$

$$h_{kj} \leftarrow h_{kj} - \eta_{kj} \frac{\partial D_E(V \| WH)}{\partial h_{kj}}$$

Algorithmes de descente de gradient couramment utilisés

Tant que les valeurs initiales sont non négatives, il garantit la non-négativité des matrices dans l'optimisation, simplifiant le problème en **une optimisation sans contrainte**.



$$\frac{\partial D_E(V \| WH)}{\partial w_{ik}} = -[(V - WH)H^T]_{ik}$$

$$\frac{\partial D_E(V \| WH)}{\partial h_{kj}} = -[W^T(V - WH)]_{kj}$$

Distance euclidienne



$$\mu_{ik} = \frac{w_{ik}}{[WHH^T]_{ik}}, \quad \eta_{kj} = \frac{h_{kj}}{[W^TWH]_{kj}}$$

Coefficient de descente de gradient



$$w_{ik} \leftarrow w_{ik} \frac{[VH^T]_{ik}}{[WHH^T]_{ik}}$$

$$h_{kj} \leftarrow h_{kj} \frac{[W^TV]_{kj}}{[W^TWH]_{kj}}$$



Algorithme d'optimisation multiplicative

## **NMF: Modélisation acoustique**

**Hypothèse** : le signal reçu est composé de la somme d'une série de variables aléatoires complexes indépendantes.

$$y_{fn} = \sum_{k \in \mathcal{K}} x_{fn} \quad \text{with} \quad x_{fn} \sim \mathcal{N}_c(0, w_{fk} h_{kn})$$

Les signaux sont représentés par des spectrogrammes, où  $f$  représente l'axe des fréquences et  $n$  l'axe du temps.

## **NMF: Modélisation acoustique**

**Hypothèse** : le signal reçu est composé de la somme d'une série de variables aléatoires complexes indépendantes.

$$y_{fn} = \sum_{k \in \mathcal{K}} x_{fn} \quad \text{with} \quad x_{fn} \sim \mathcal{N}_c(0, w_{fk} h_{kn})$$



$$y_{fn} \sim \mathcal{N}_c\left(0, \sum_{k \in \mathcal{K}} w_{fk} h_{kn}\right)$$

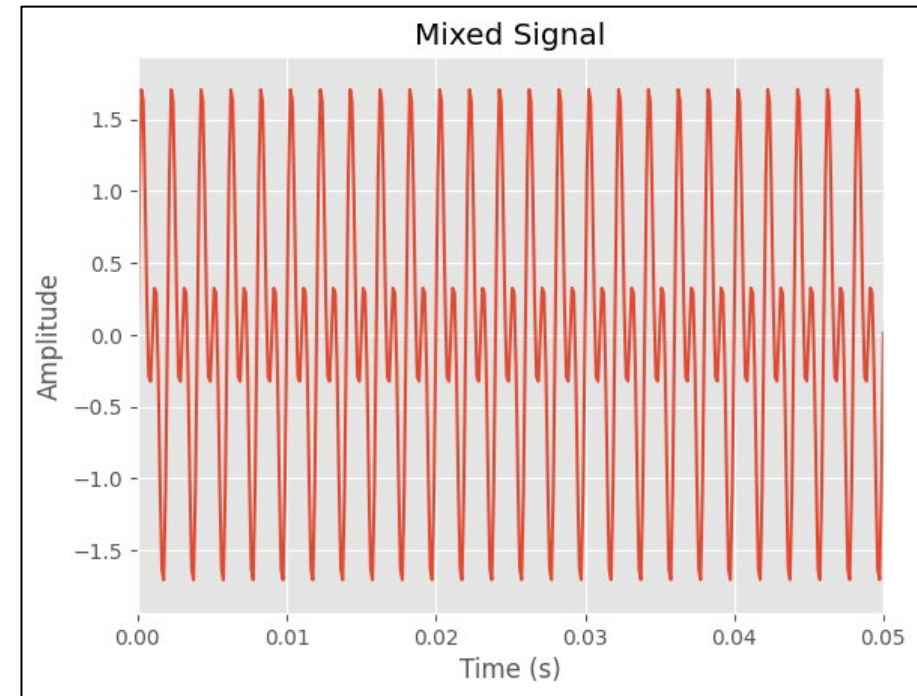
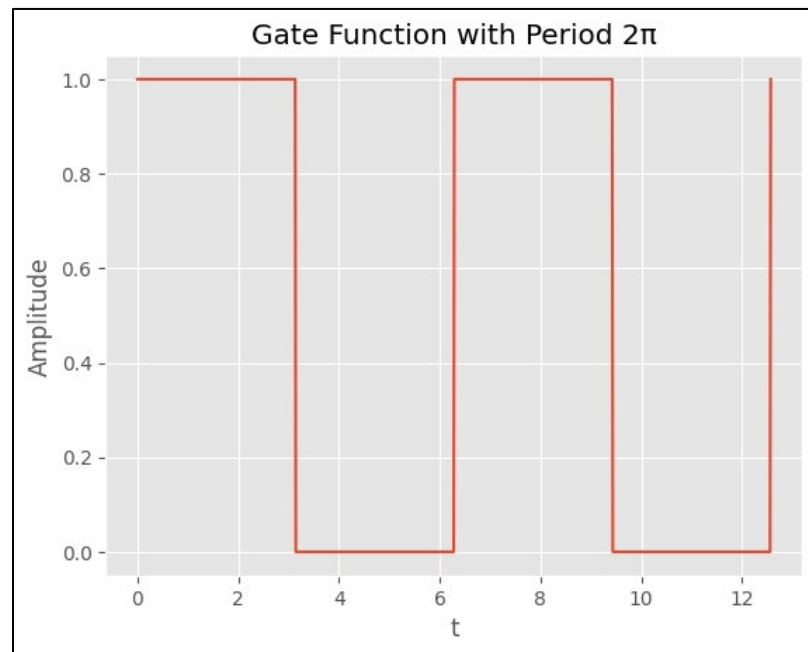


$$-\log p(\mathbf{Y} \mid \mathbf{W}, \mathbf{H}) = \sum_{fn} d_{IS}\left(|s_{fn}|^2 \mid [\mathbf{WH}]_{fn}\right) + \text{constante}$$

la négative log-vraisemblance de  $y$  peut s'écrire comme une somme de divergences IS

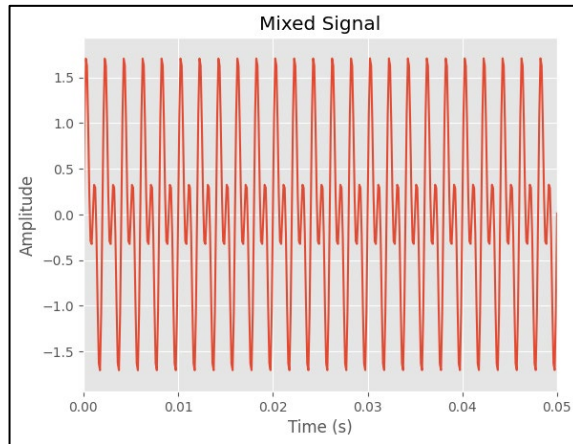
## Simulation sur une onde sinus

$$s(t) = g(\alpha t) \sin(\gamma t) + g(\beta t) \sin(\delta t)$$

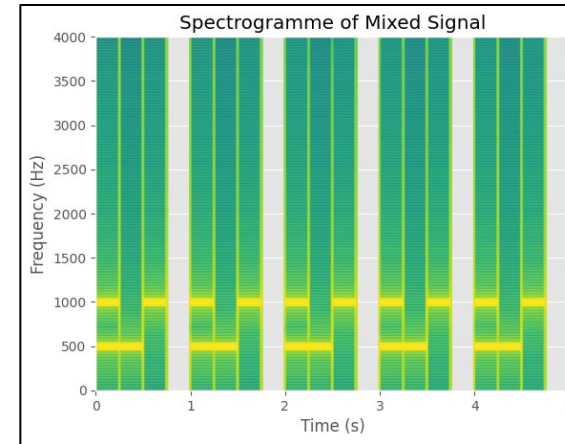


$$\alpha = 1, \beta = 2, \gamma = 500, \delta = 1000$$

# **Simulation** sur une onde sinus



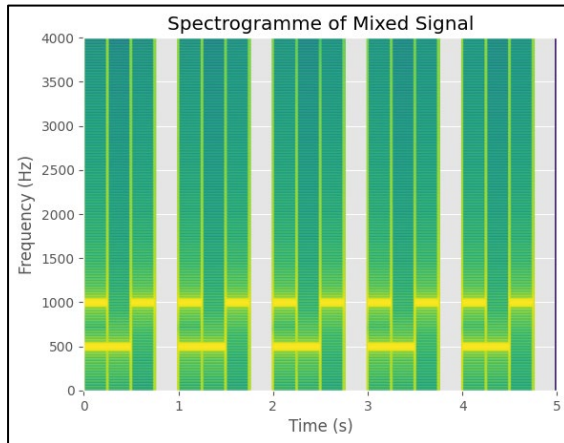
➔  
STFT



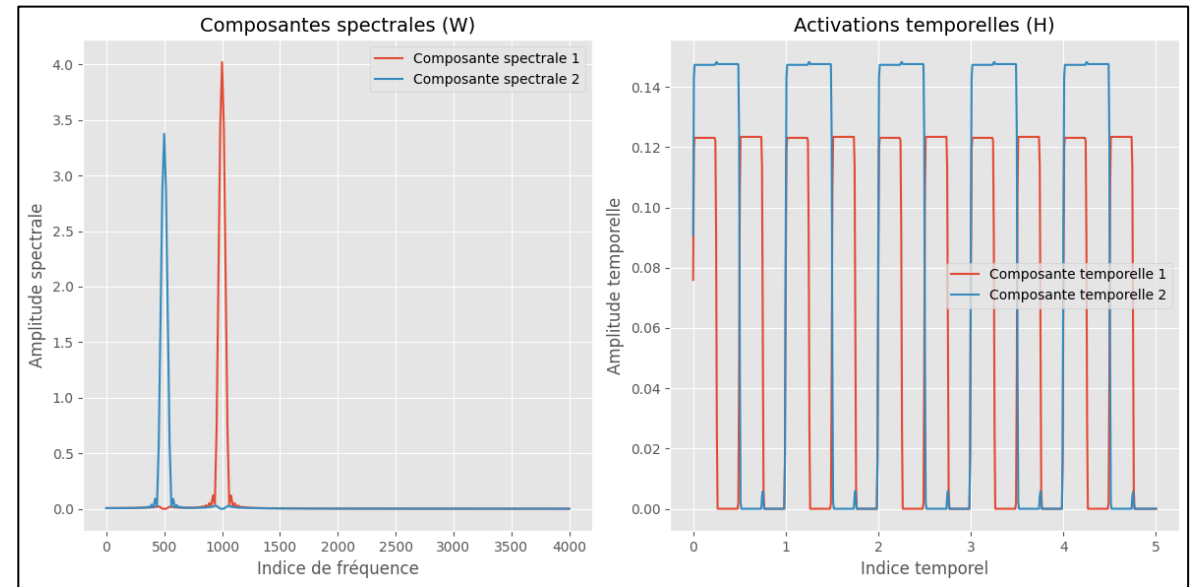


# Simulation sur une onde sinus

➔  
STFT

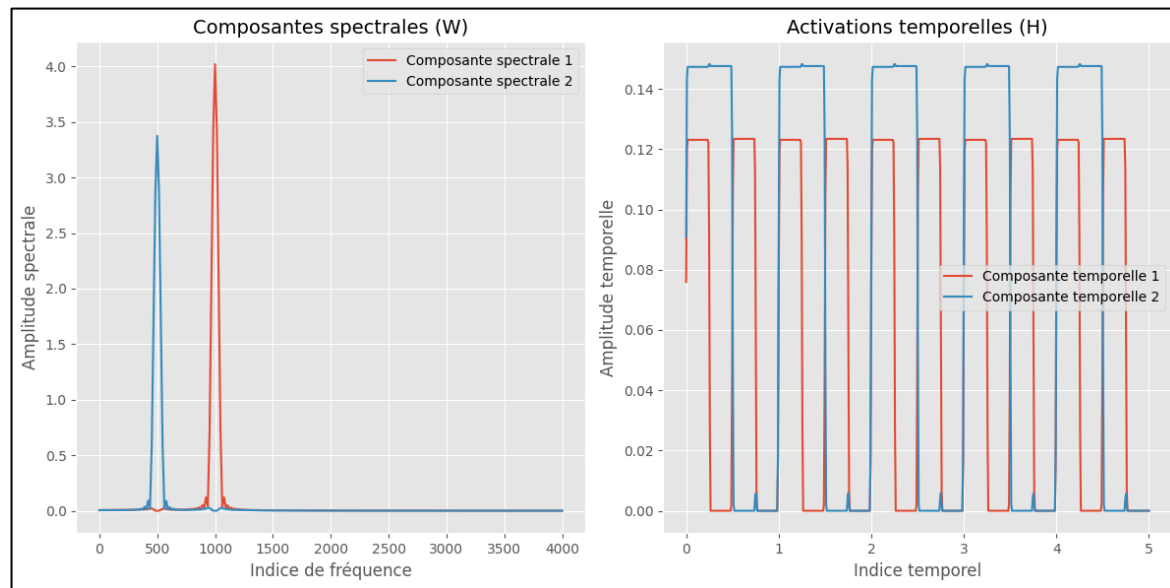


➔  
NMF



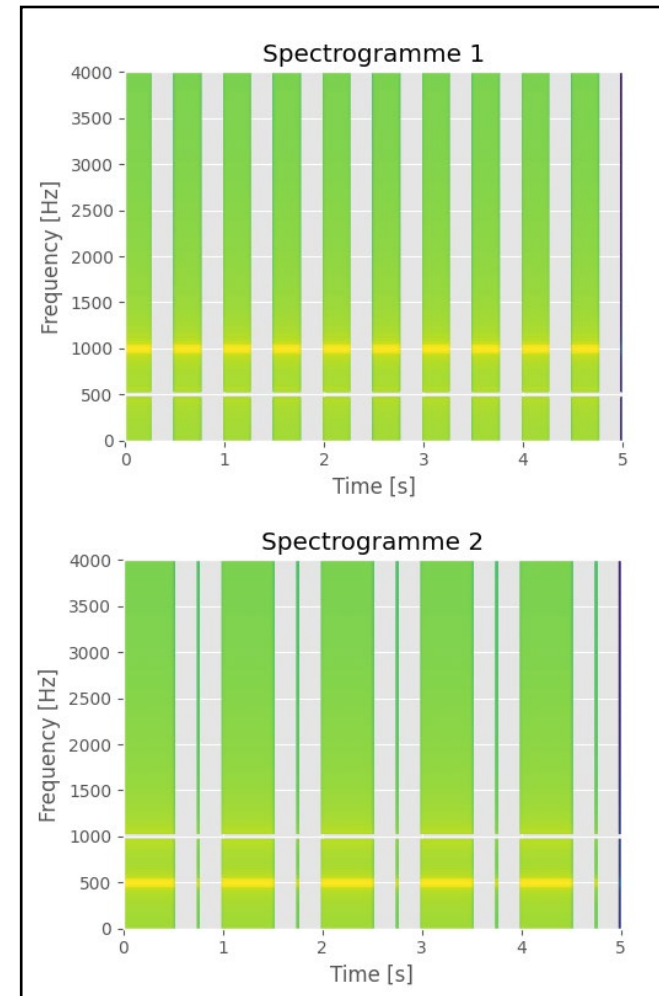
**Les colonnes de la matrice dictionnaire et les lignes de la matrice d'activation représentent les caractéristiques des deux signaux**

# Simulation sur une onde sinus



Les colonnes de la matrice dictionnaire et les lignes de la matrice d'activation

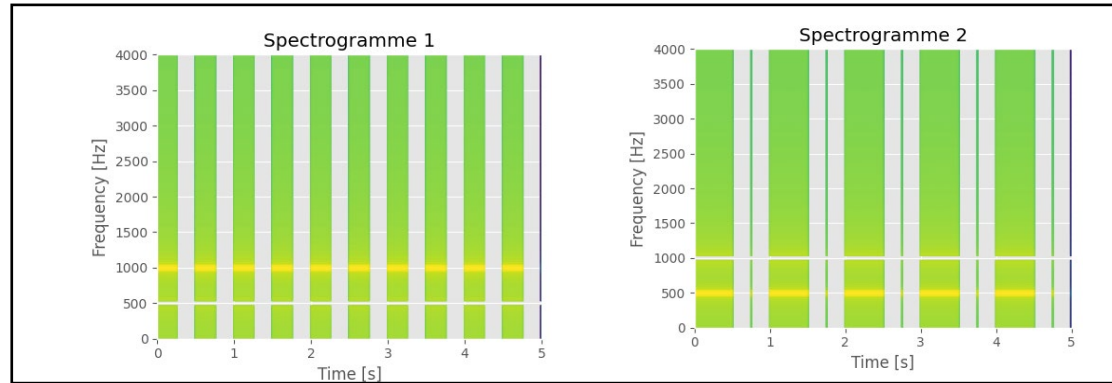
➔  
Produit



Spectrogrammes séparés

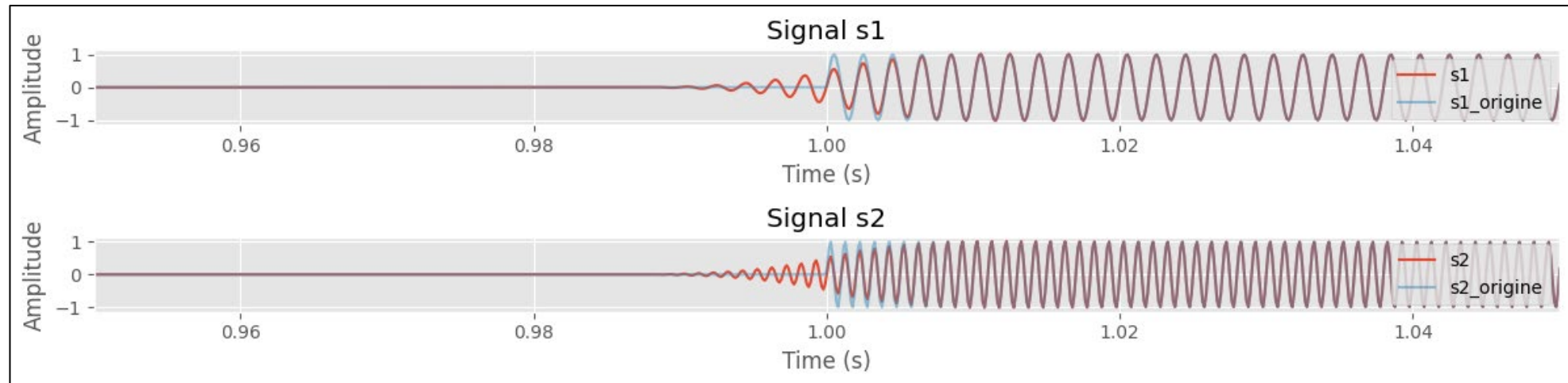
# Simulation sur une onde sinus

➔  
NMF



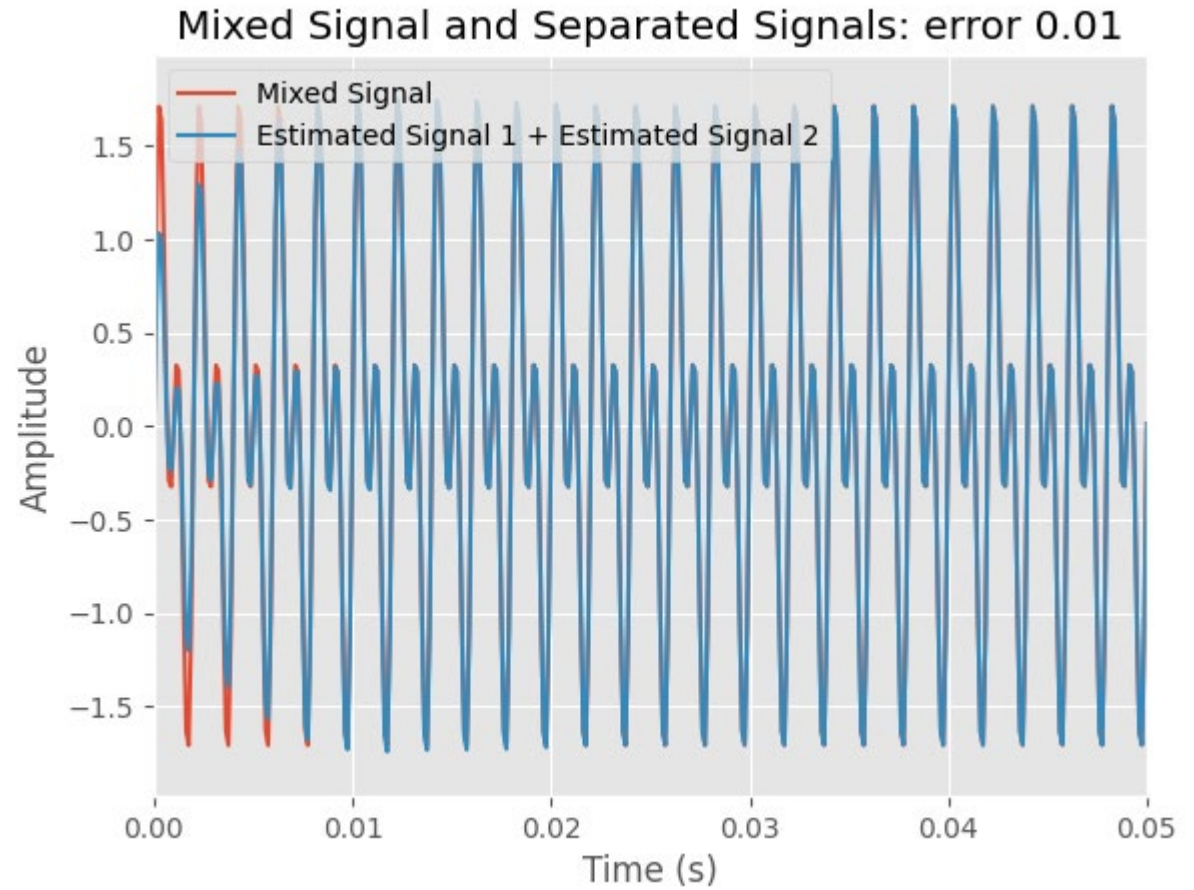
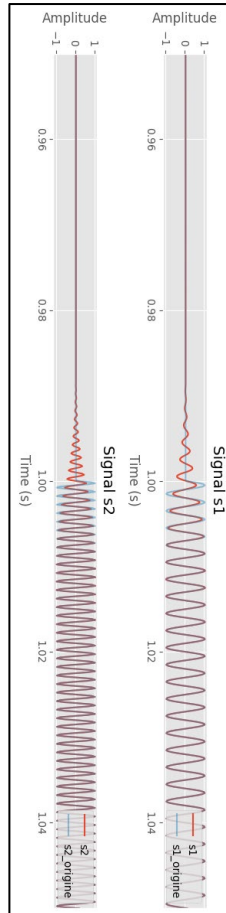
Spectrogrammes séparés

↓  
ISTFT



# Simulation sur une onde sinus

➔  
ISTFT

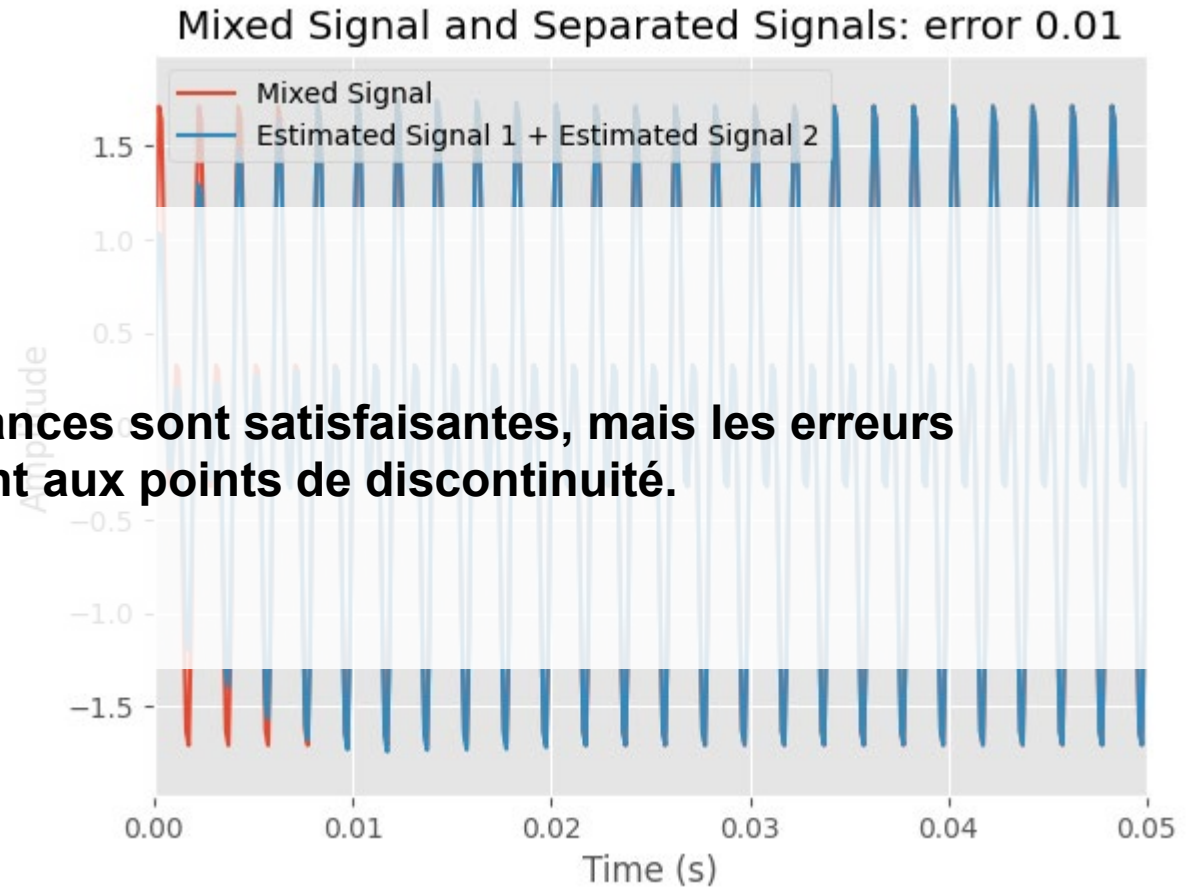


# Simulation sur une onde sinus

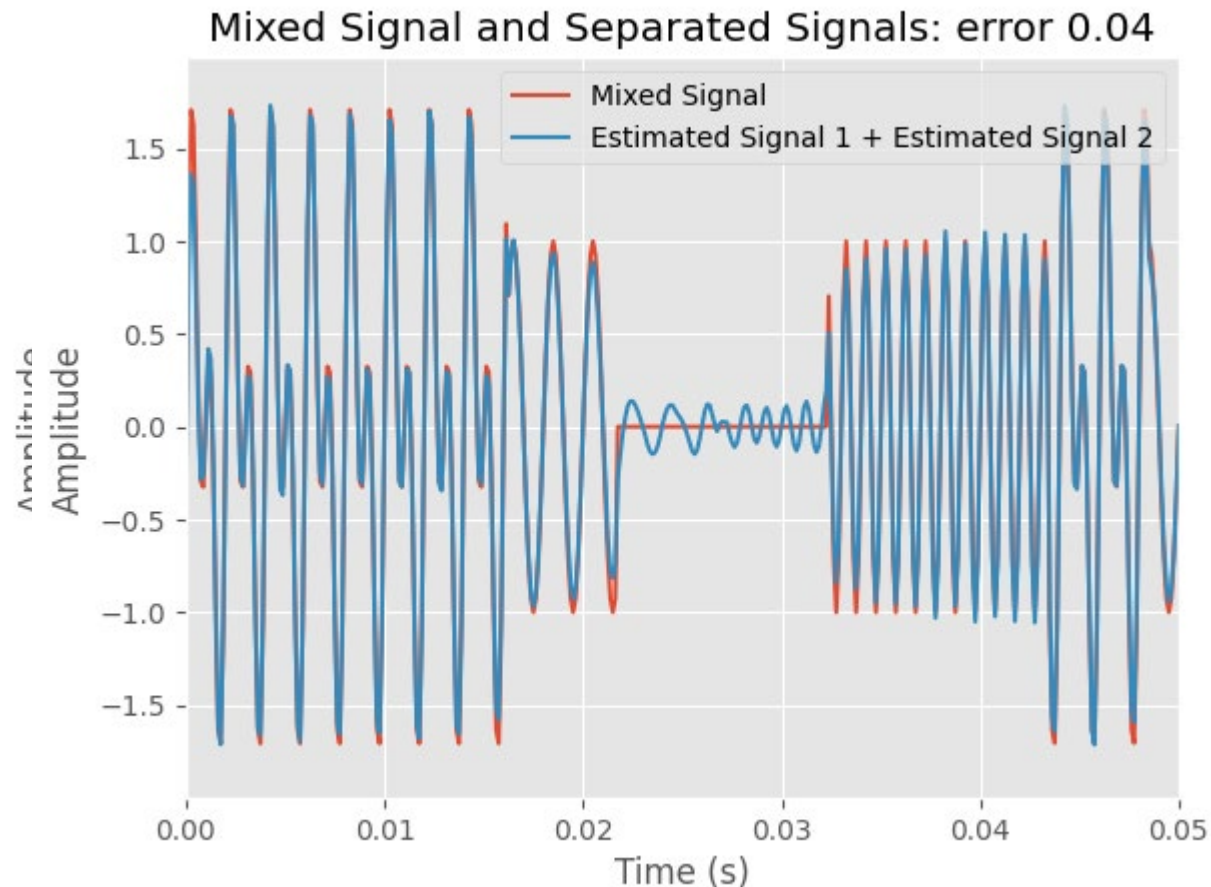
→  
ISTFT



Totalement, les performances sont satisfaisantes, mais les erreurs se présentent aux points de discontinuité.



## Simulation sur une onde sinus



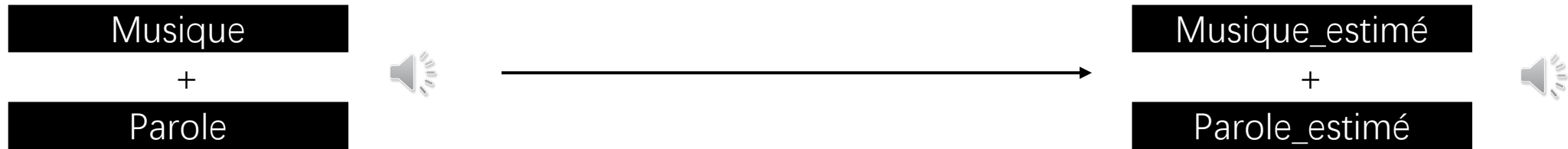
### Performance aux points de discontinuité

$$\alpha = 23, \beta = 31, \gamma = 500, \delta = 1000$$

$$\text{erreur} = \frac{1}{N} \sum_{i=1}^N \left| y^{(i)} - \left( x_{e1}^{(i)} + s_{e2}^{(i)} \right) \right|$$

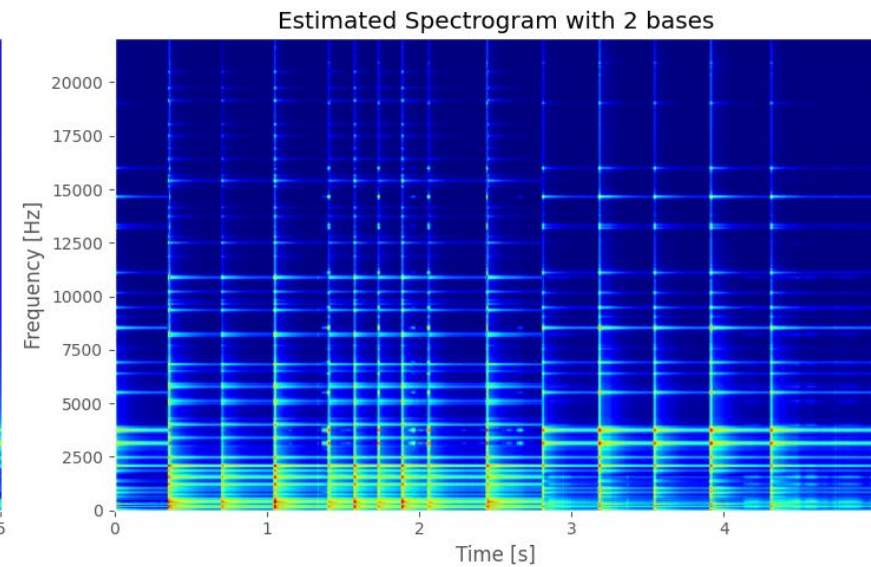
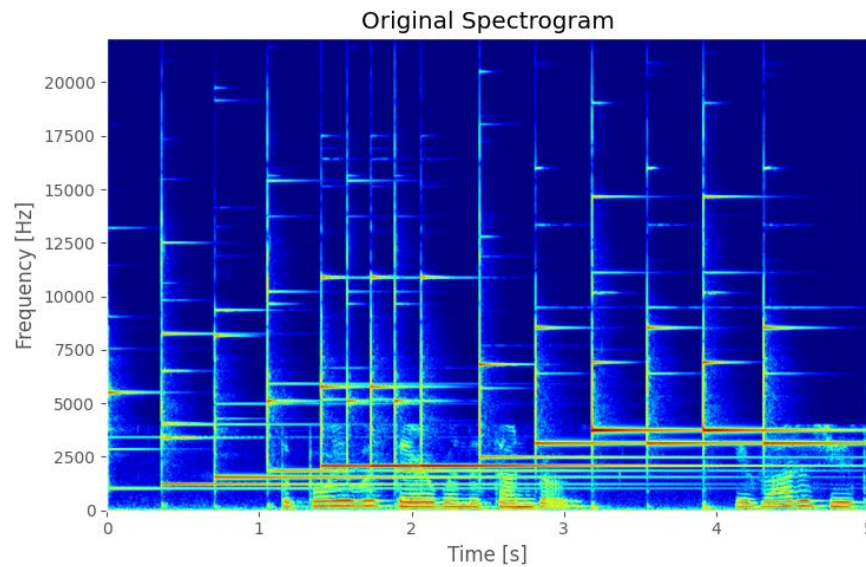
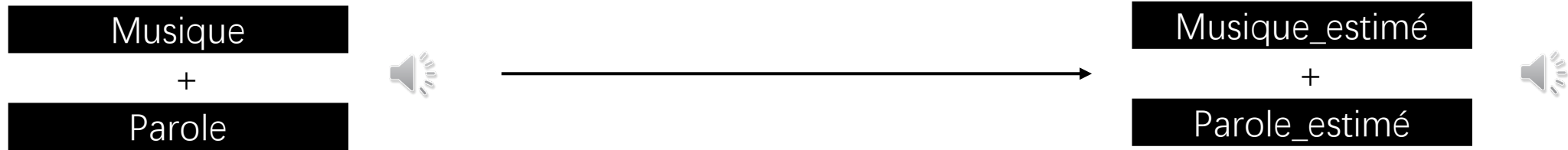
# Échec sur la simulation du mélange de musique et parole

---



$V \neq HW$

# Échec sur le problème du mélange de musique et parole

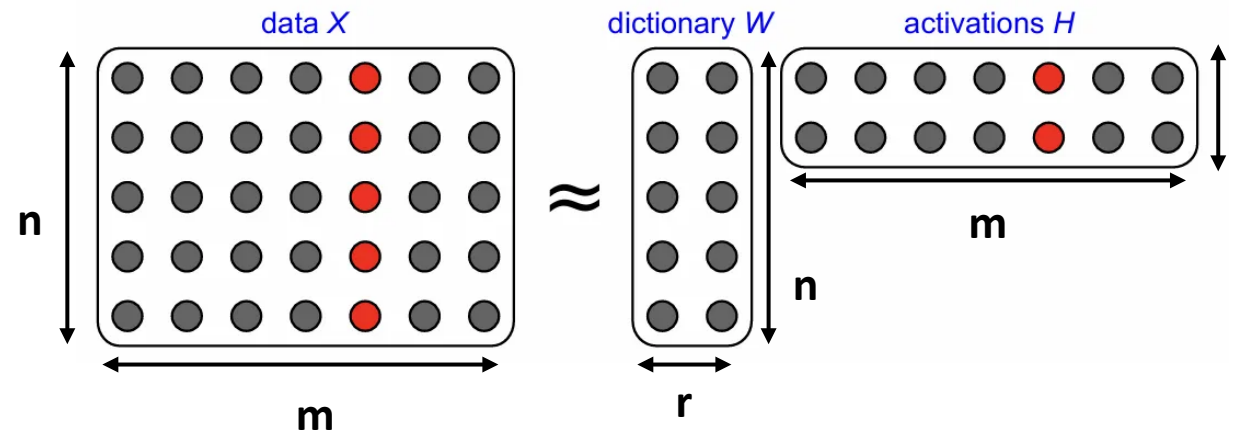


**NMF n'a pas réussi à reconstruire correctement le signal.**



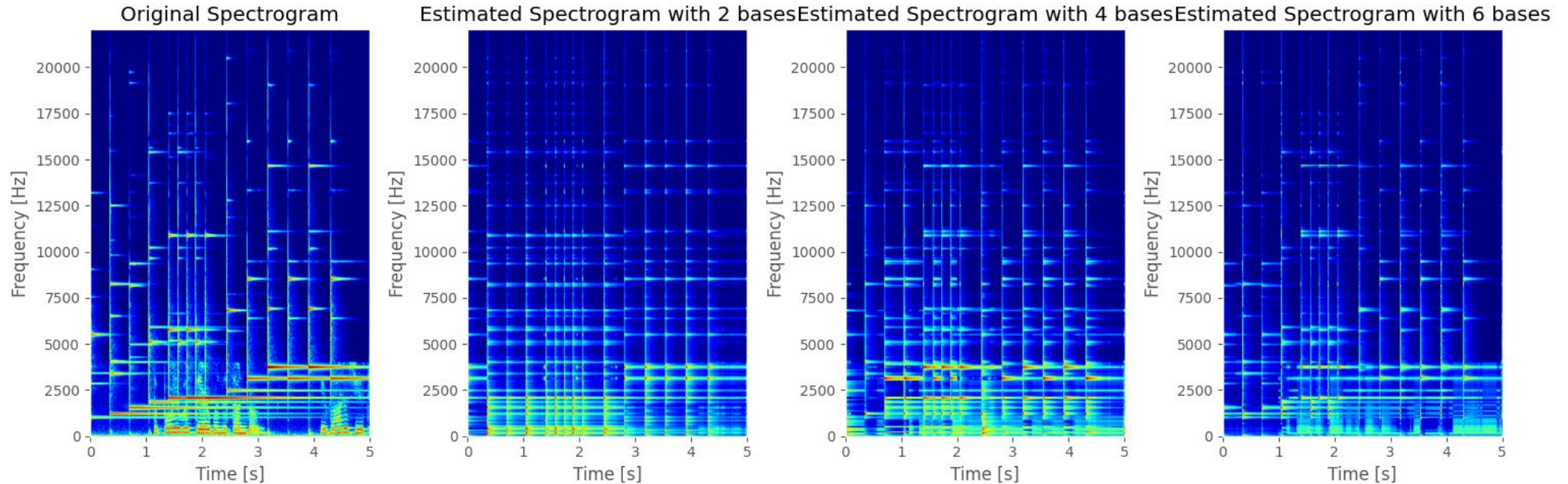
# Échec sur le problème du mélange de musique et parole

---



NMF comprime l'information de la matrice originale, et le taux de compression dépend de  $r$ .

# Échec sur le problème du mélange de musique et parole



Corrélation =  $\langle \text{signal}, \text{signal estimé} \rangle$       Corrélation = -17161      Corrélation = 11517      Corrélation = 25898

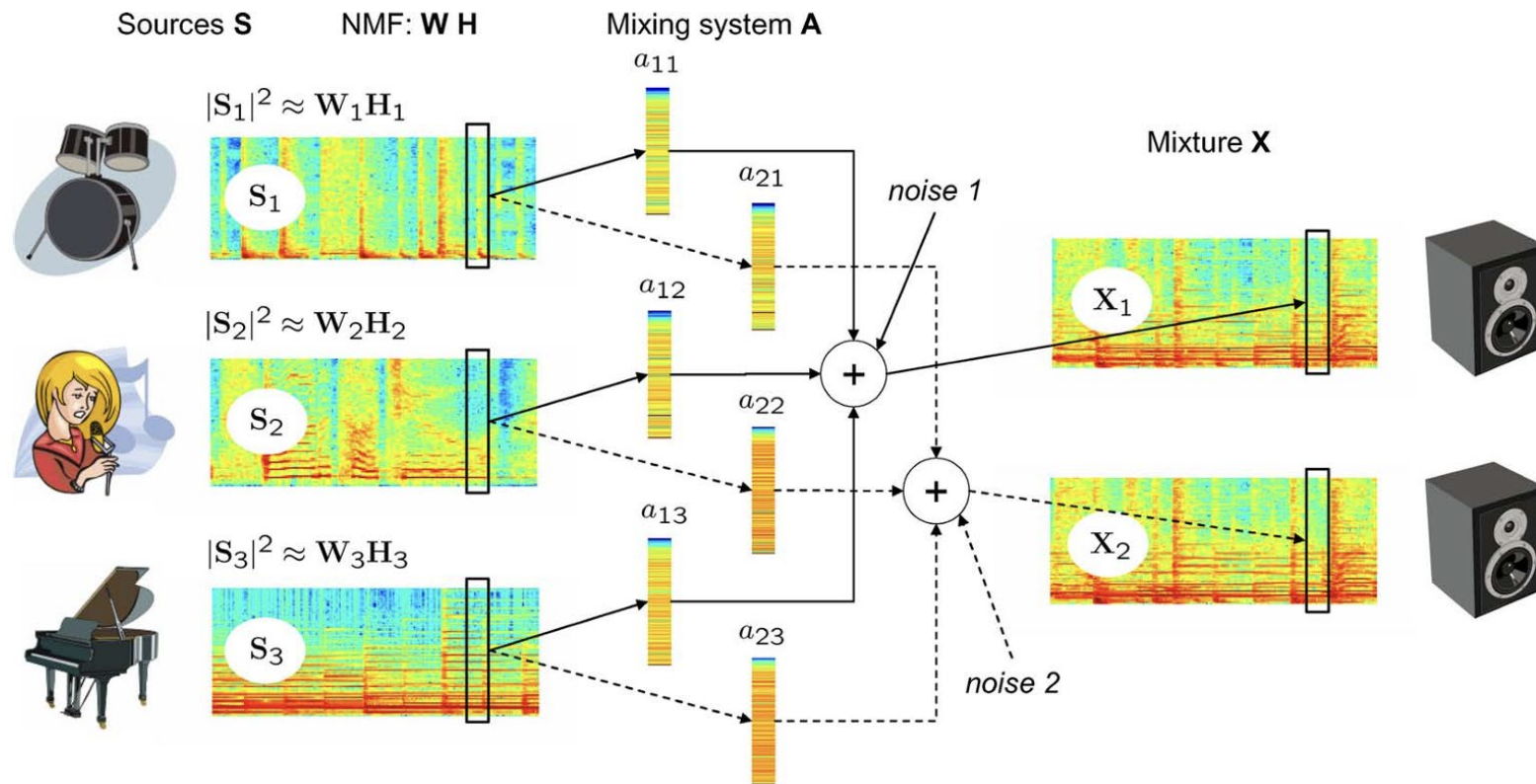
Augmenter  $r$  favorise une meilleure reconstruction du signal, mais cela entre en contradiction avec notre objectif de diviser le signal en deux parties.

# **Échec** sur le problème du mélange de musique et parole

## **Les Limitations de NMF**

1. L'augmentation de la précision de reconstruction est en contradiction avec la tâche de séparation ;
2. Les sources sonores ne respectent pas nécessairement l'hypothèse de distribution gaussienne complexe ;
3. Il ne peut pas traiter des données multicanales.

# MNMF: Factorisation en Matrices Non Négatives Multicanal



## MNMF: Modélisation acoustique

**Hypothèse** : le signal de source est composé de la somme d'une série de variables aléatoires complexes indépendantes.

$$s_{j,fn} = \sum_{k \in \mathcal{K}_j} c_{k,fn} \quad \text{with} \quad c_{k,fn} \sim \mathcal{N}_c(0, w_{fk} h_{kn})$$



$$s_{j,fn} \sim \mathcal{N}_c \left( 0, \sum_{k \in \mathcal{K}_j} w_{fk} h_{kn} \right)$$



$$-\log p(\mathbf{S}_j | \mathbf{W}_j, \mathbf{H}_j) \stackrel{c}{=} \sum_{fn} d_{IS} (|s_{j,fn}|^2 | [\mathbf{W}_j \mathbf{H}_j]_{fn})$$



$$\hat{y}_{i,fn} = \sum_j q_{ij,f} \underbrace{\sum_{k \in \mathcal{K}_j} w_{fk} h_{kn}}_{\tilde{s}_{j,fn}}$$



$$C_2(\boldsymbol{\theta}) = \sum_{ifn} d_{IS} (|y_{i,fn}|^2 | \tilde{y}_{i,fn})$$



En prenant les signaux sources comme objectif d'estimation, en plus de W et H, la matrice Q est également estimée.

## MNMF: Modélisation acoustique

**Hypothèse** : le signal de source est composé de la somme d'une série de variables aléatoires complexes indépendantes.

$$s_{j,fn} = \sum_{k \in \mathcal{K}_j} c_{k,fn} \quad \text{with} \quad c_{k,fn} \sim \mathcal{N}_c(0, w_{fk} h_{kn}) \quad \Rightarrow \quad \hat{y}_{i,fn} = \sum_j a_{ij,f} \underbrace{\sum_{k \in \mathcal{K}_j} w_{fk} h_{kn}}_{\tilde{s}_{j,fn}}$$

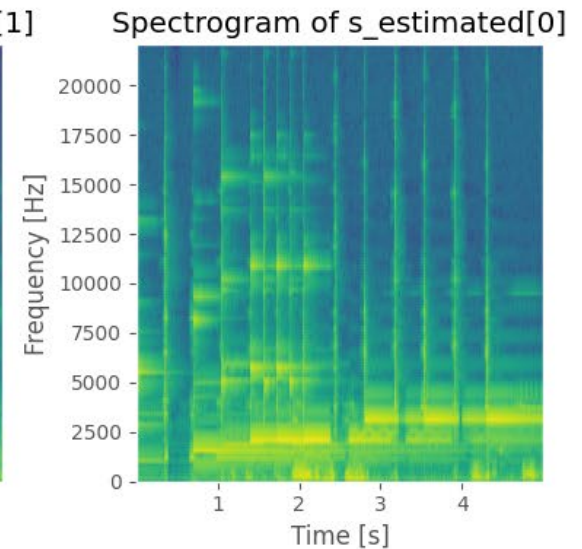
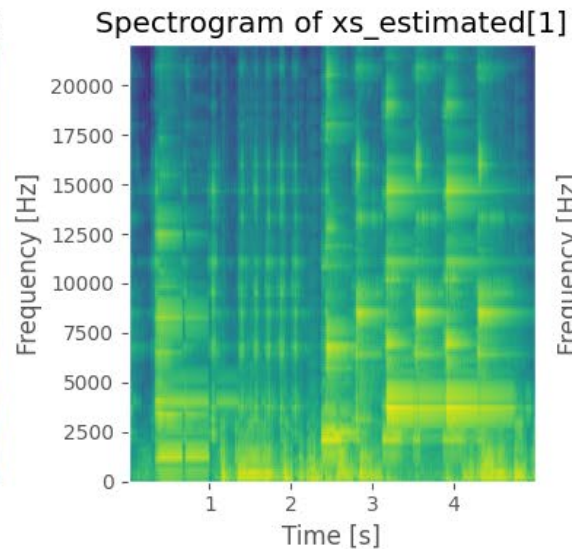
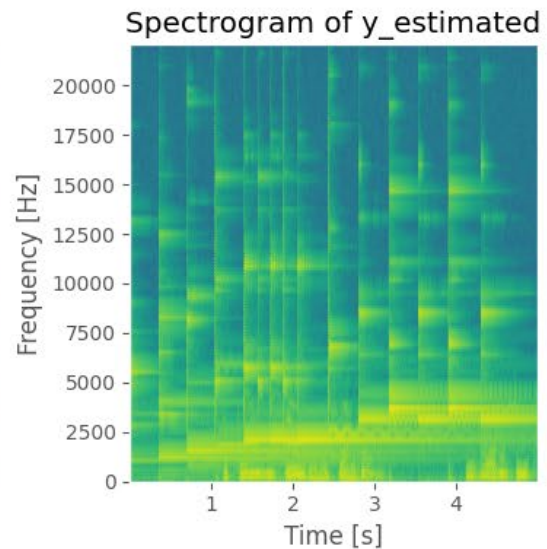
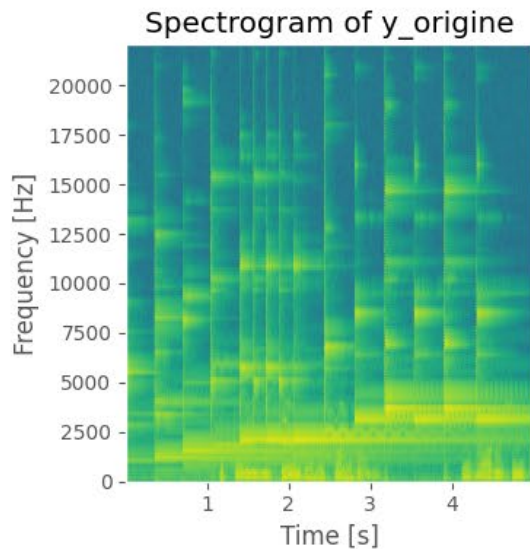
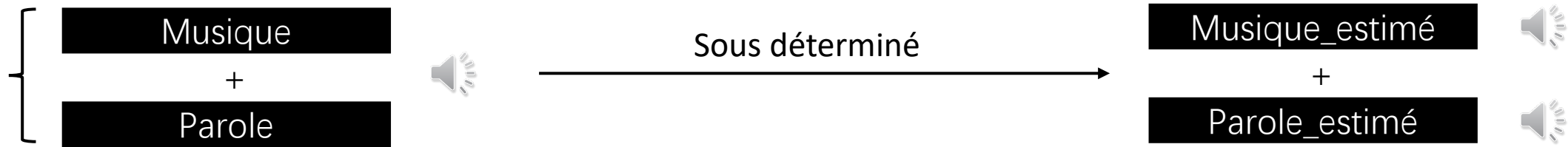
Ainsi, le nombre de bases est séparé du nombre de sources sonores, ce qui nous permet d'exprimer des sources sonores plus complexes et d'utiliser davantage de signaux de microphones.

$$C_2(\theta) = \sum_{ifn} d_{IS}(|y_{i,fn}|^2 | \tilde{y}_{i,fn})$$

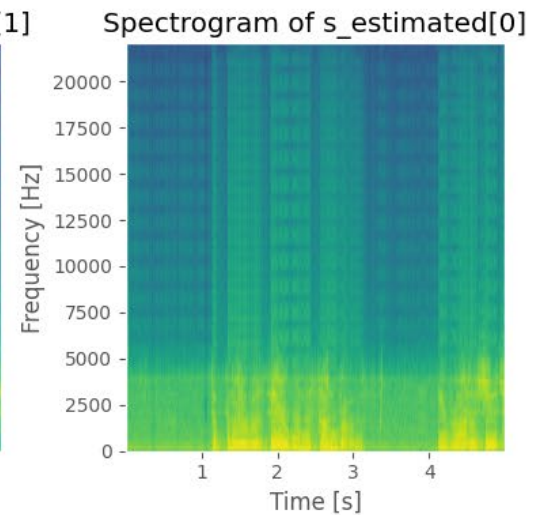
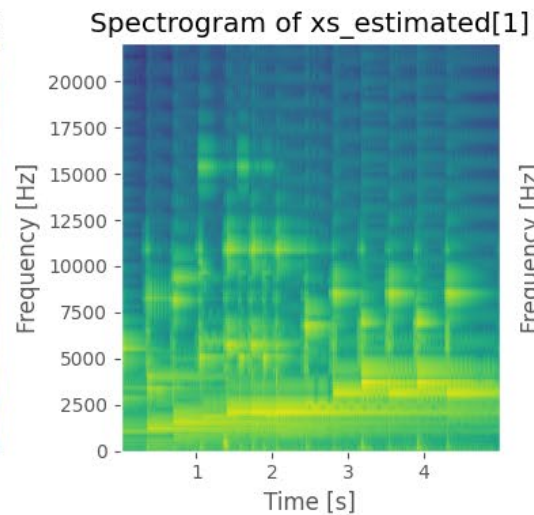
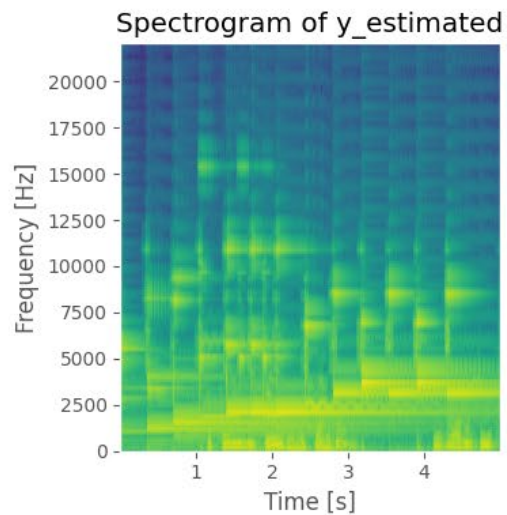
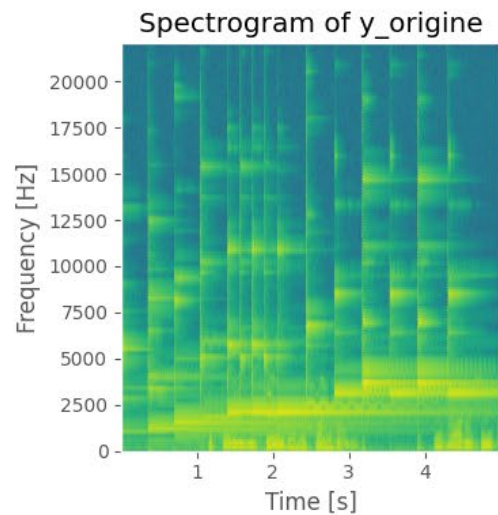
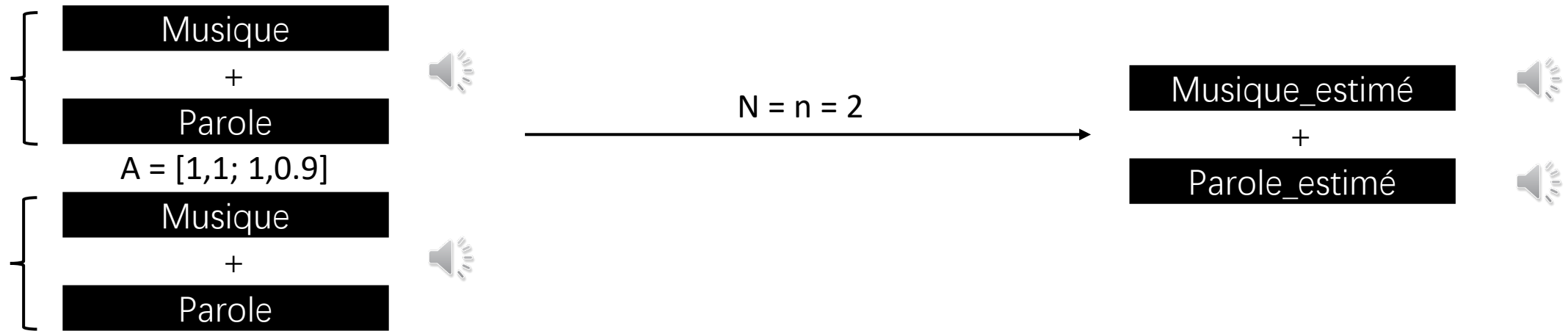
$$-\log p(\mathbf{S}_j | \mathbf{W}_j, \mathbf{H}_j) \stackrel{c}{=} \sum_{fn} d_{IS}(|s_{j,fn}|^2 | [\mathbf{W}_j \mathbf{H}_j]_{fn})$$

En prenant les signaux sources comme objectif d'estimation, en plus de W et H, la matrice QQQ est également estimée.

# MNMF: Simulation sur des mélange de parole et musique



# MNMF: Simulation sur des mélange de parole et musique





# MNMF: Experience sur des mélange de musique

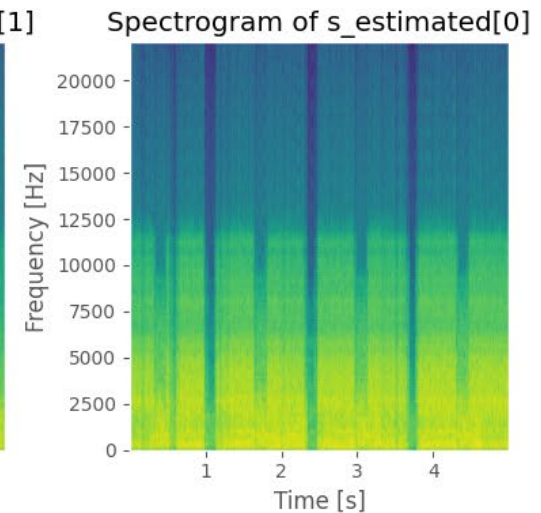
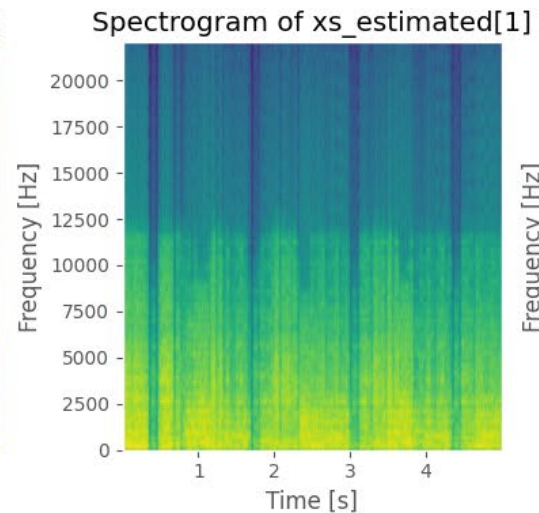
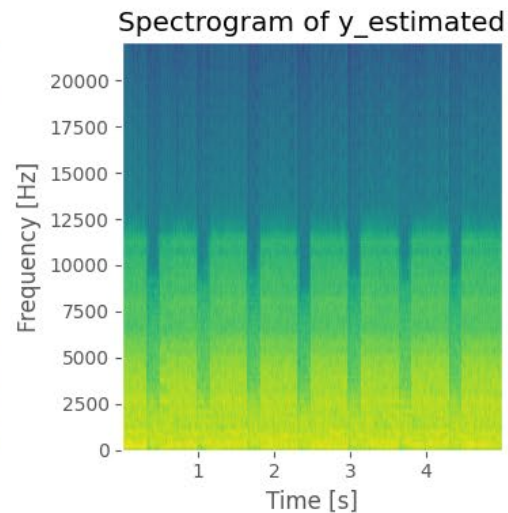
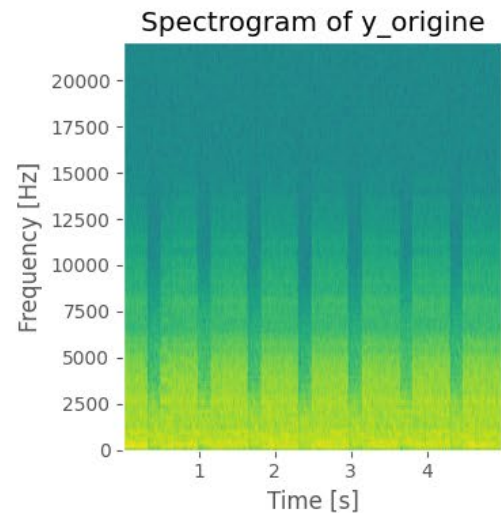
Musique composée de  
guitare et de  
synthétiseur

$N = 2, n = 1$

guitare

+

synthétiseur



# MNMF: Experience sur des mélange de musique

Musique composée de  
guitare et de  
synthétiseur



$N = 2, n = 2$

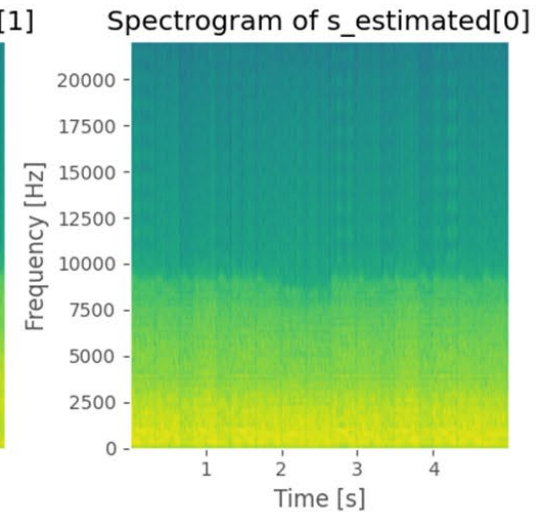
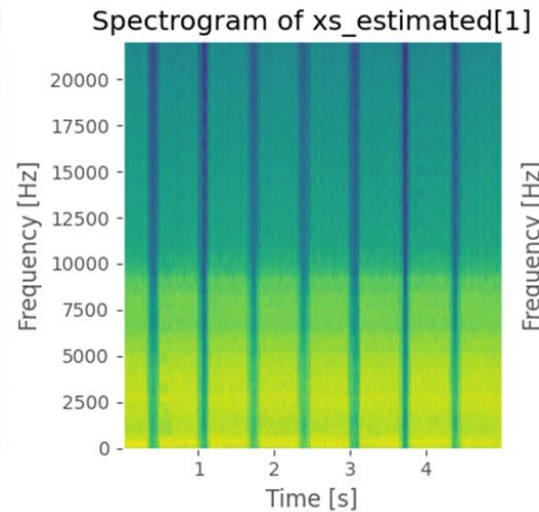
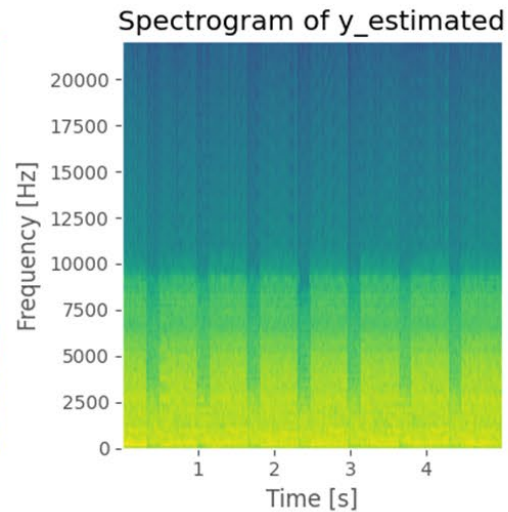
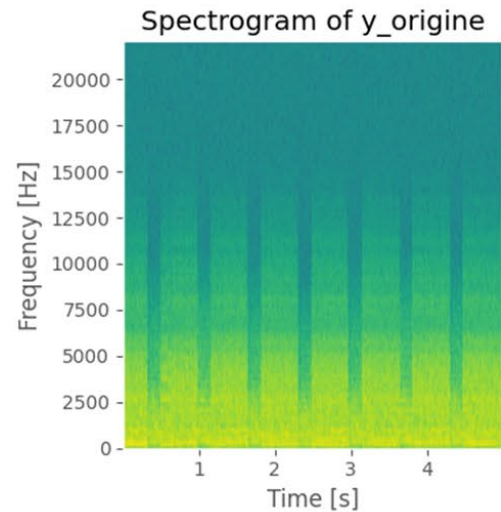


guitare

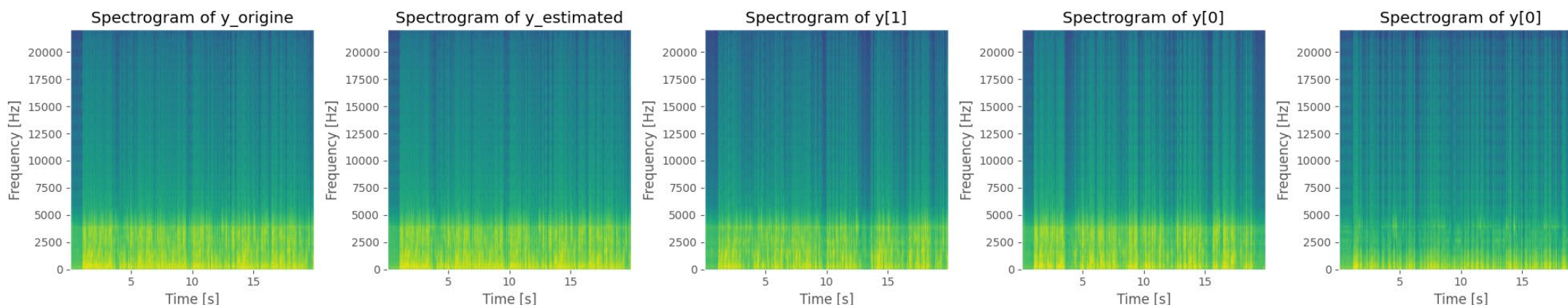
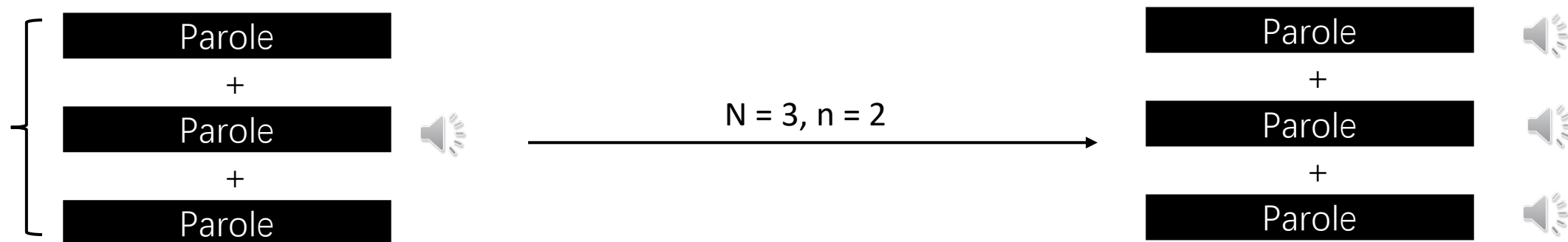


+

synthétiseur

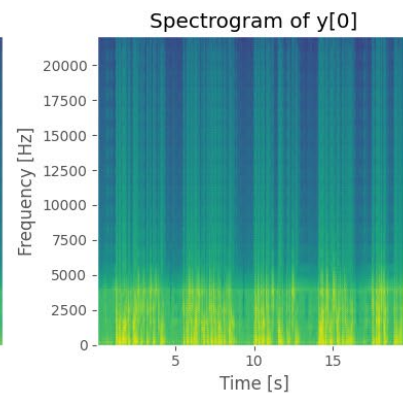
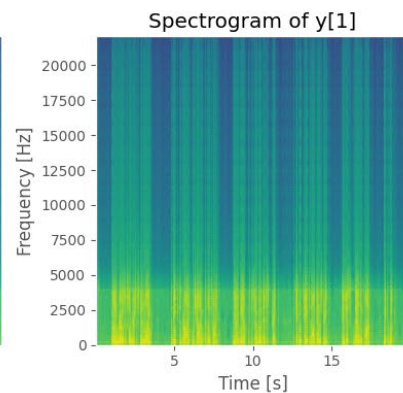
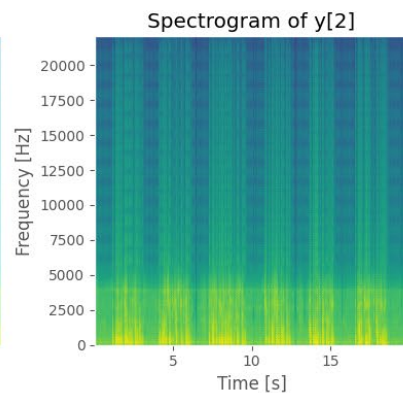
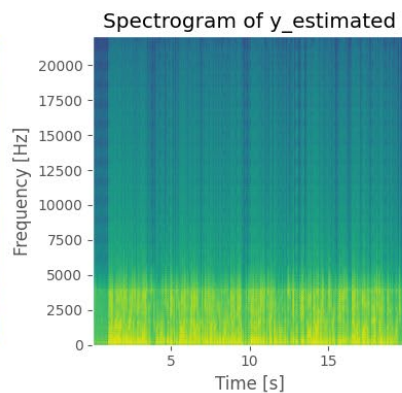
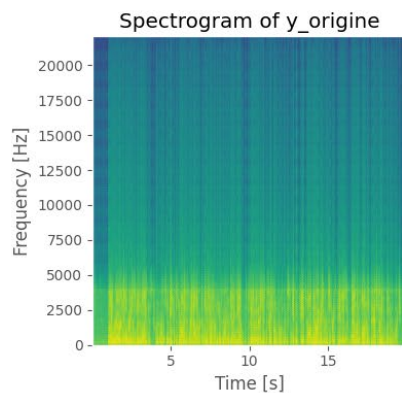
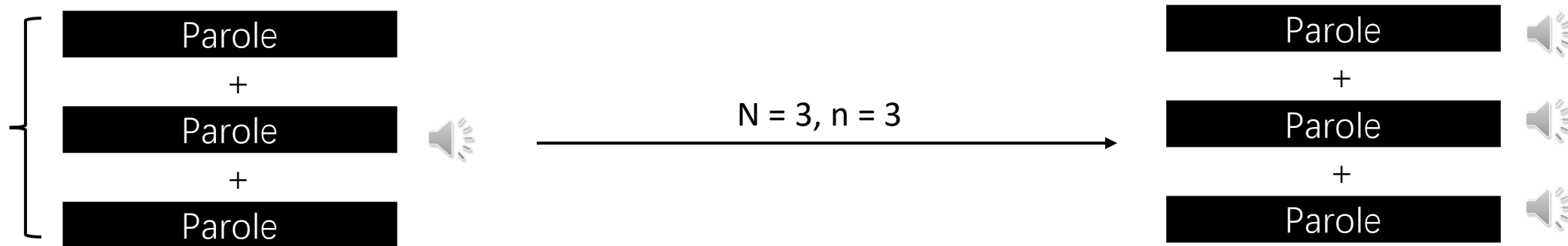


# MNMF: Simulation sur des mélange de parole



Une resultat comprenant une parole relativement claire, un mélange de deux paroles et une portion de bruit.

# MNMF: Simulation sur des mélange de parole

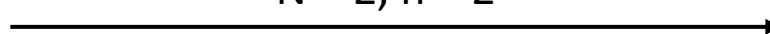


# **MNMF**: expérience utilisant la micro

Un enregistrement d'une musique diffusée par un ordinateur et le son d'une guitare.



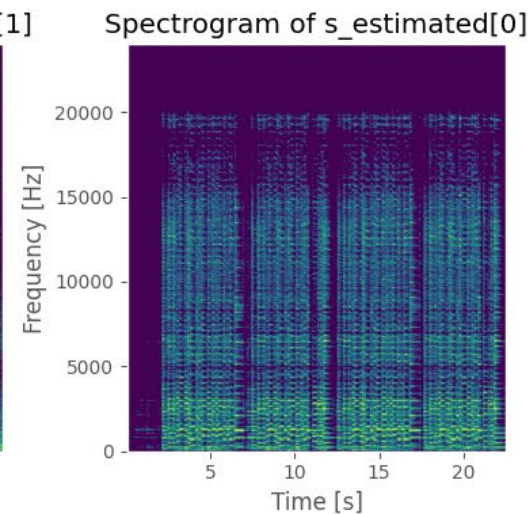
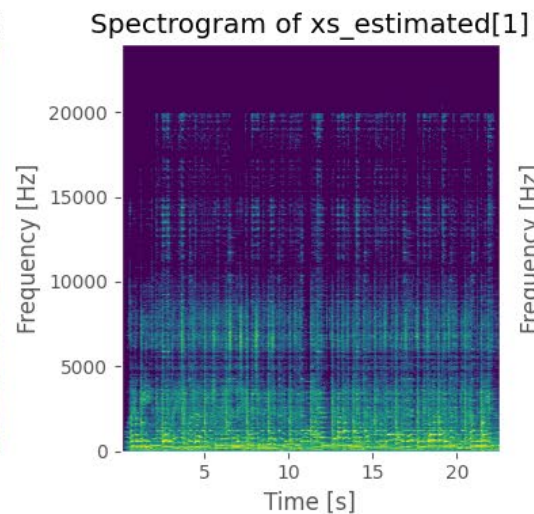
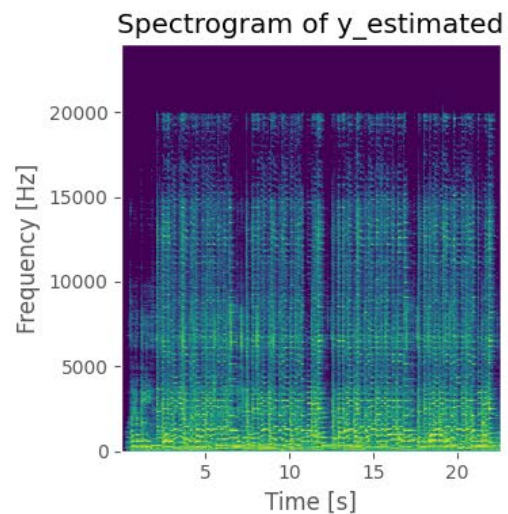
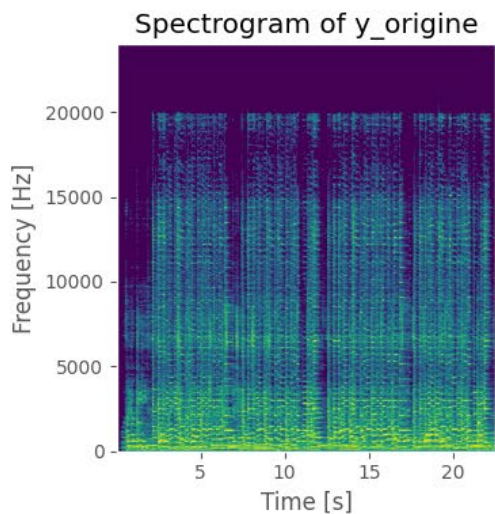
$N = 2, n = 2$



Source 1

+

Source 2



# MNMF: experience utilisant la micro

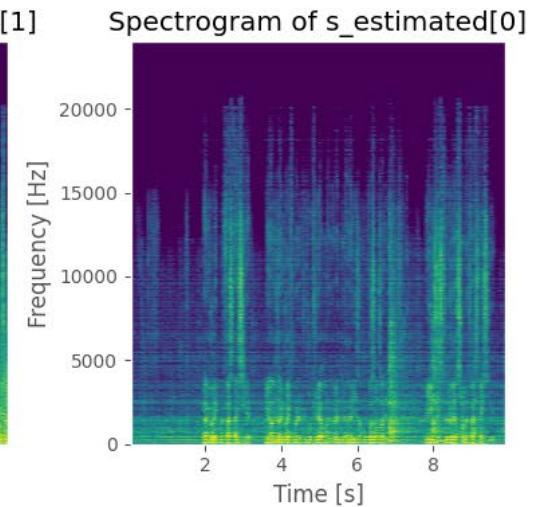
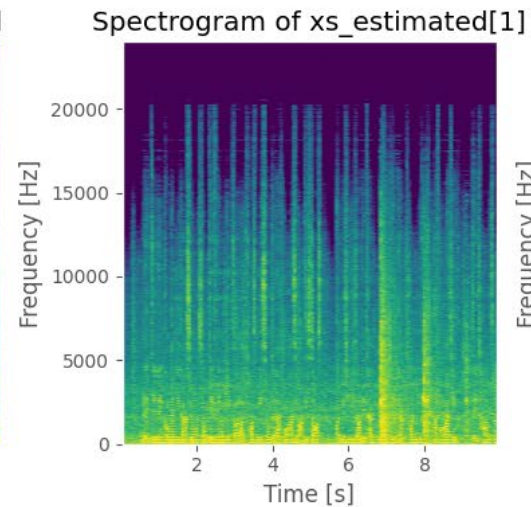
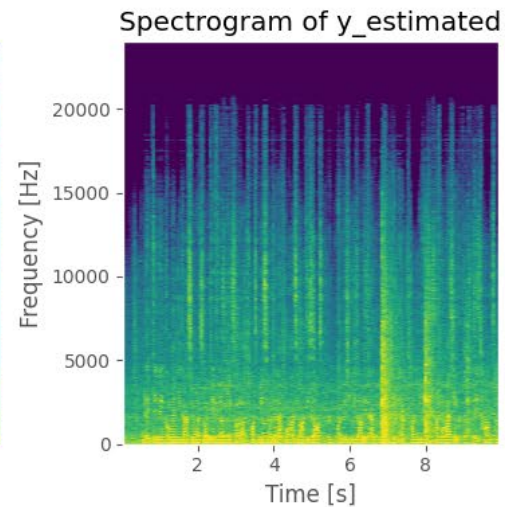
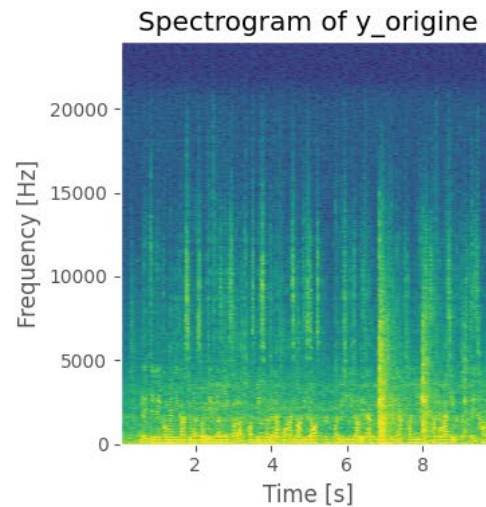
Un enregistrement à deux microphones avec deux sources..

$N = 2, n = 2$

Source 1

+

Source 2



## **MNMF: la résumé**

---

### **Avantages :**

1. Performances excellentes lorsque  $n = N$  ;
2. Bien que la séparation ne soit pas totalement satisfaisante, la reconstruction est très réussie.

### **Limitation:**

1. Performances pas très bien lorsque  $n < N$ , en particulier dans le cas des paroles ;
2. Lorsque  $n < N$ , tendance à fournir une sortie mélangeant plusieurs canaux et une autre sortie presque sans information.



**l'hypothèse autorisant un canal à être composée de plusieurs distributions gaussiennes complexes rend possible qu'un canal séparé contienne des informations provenant de plusieurs sources.**

## Références

---

- [1] Multichannel Nonnegative Matrix Factorization in Convolutional Mixtures for Audio Source Separation | IEEE Journals & Magazine | IEEE Xplore'. Accessed: Nov. 29, 2024. [Online]. Available: <https://ieeexplore.ieee.org/document/5229304>
- [2] P. Smaragdis and J. C. Brown, 'Non-negative matrix factorization for polyphonic music transcription', in *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (IEEE Cat. No.03TH8684)*, Oct. 2003, pp. 177–180. doi: [10.1109/ASPAA.2003.1285860](https://doi.org/10.1109/ASPAA.2003.1285860).
- [3] 非负矩阵分解 (2) : 算法推导与实现 - LeeLin。 - 博客园'. Accessed: Nov. 29, 2024. [Online]. Available: <https://www.cnblogs.com/xingshansi/p/6670214.html>
- [4] AshishAbrahamSamuel/Music-Source-Separation-using-NMF'. Accessed: Nov. 29, 2024. [Online]. Available: <https://github.com/AshishAbrahamSamuel/Music-Source-Separation-using-NMF/tree/main>



**Merci**