

# Modèle gaussien anisotrope à phase informée

## Sommaire

---

<b>8.1</b>	<b>Modèle de Von Mises</b>	<b>130</b>
8.1.1	Modèle de sources	130
8.1.2	Verrous du modèle	132
8.1.3	Relations avec d'autres modèles	132
<b>8.2</b>	<b>Modèle gaussien anisotrope</b>	<b>133</b>
8.2.1	Loi normale complexe	133
8.2.2	Paramètres du modèle	133
8.2.3	Comparaison des deux modèles	135
8.2.4	Estimateur MMSE des sources	135
<b>8.3</b>	<b>Validation expérimentale</b>	<b>140</b>
8.3.1	Influence du paramètre de concentration	140
8.3.2	Séparation de sources musicales	141
<b>8.4</b>	<b>Conclusion</b>	<b>144</b>

---

---

La technique de déroulé linéaire de phase introduite au chapitre 4 a été appliquée à la séparation de sources au chapitre 5. Cette méthode était basée sur la minimisation d'une fonction de coût pénalisant l'écart entre modèle et mélange observé, et l'information sur la phase provenant du modèle sinusoïdal était introduite dans l'algorithme via son initialisation. Ce type d'approche souffre néanmoins de deux écueils. Tout d'abord, l'ensemble des sources estimées ne constitue pas un modèle conservatif (la somme des estimées n'est pas égale au mélange observé). En outre, nous avons constaté dans le cas non-Oracle (où les spectrogrammes ne sont plus connus) une baisse des résultats, qui est due à une contrainte trop forte sur les amplitudes.

Les modèles probabilistes sont des outils adaptés à l'injection de connaissance préalable sur les paramètres que l'on cherche à estimer. Ils permettent donc d'éviter ces deux écueils, puisque l'on peut alors introduire une certaine incertitude sur les paramètres. La plupart des modèles probabilistes de sources, comme FÉVOTTE et al. (2009), sont basés sur une hypothèse d'uniformité de la phase. Cette hypothèse est vérifiée lorsqu'on considère que tous les points TF sont indépendants, et qu'on ne tient pas compte de la structure de la phase PARRY et ESSA (2007). Néanmoins, comme le suggère le modèle de phase introduit dans le chapitre 4, introduire des dépendances entre phases de points TF successifs améliore la qualité des signaux reconstruits. Dans ce chapitre, nous proposons un modèle probabiliste de mélange dans lequel les phases ne sont plus des variables aléatoires uniformes.

Les principales contributions de ce chapitre ont fait l'objet d'une publication à la conférence ICASSP 2017 MAGRON et al. (2017a).

Dans la section 8.1, nous introduisons un nouveau modèle de mélange basé sur la distribution de Von Mises pour représenter la phase. Nous approchons ensuite dans la section 8.2 celui-ci par un modèle gaussien équivalent, plus simple à manipuler. Dans la section 8.3, nous montrons expérimentalement le potentiel de ce modèle pour la séparation de sources musicales. Enfin, nous concluons dans la section 8.4.

## 8.1 Modèle de Von Mises

Nous proposons dans cette partie d'introduire un modèle de mélange basé sur la loi de Von Mises pour représenter les phases.

### 8.1.1 Modèle de sources

On raisonne dans un point TF  $(f, t)$  donné, tous les points étant supposés indépendants. Ainsi, on retire les indices  $f$  et  $t$  dans ce qui suit afin de clarifier les écritures. Le mélange  $X$  est égal à une somme de sources  $Z_k$ . Nous supposons les modules  $V_k$  des sources estimés au préalable (par une NMF par exemple) ou bien connus (ils sont déterministes dans ce modèle). Ainsi, sous forme polaire, on a,  $\forall k$  :

$$Z_k = V_k e^{i\phi_k}. \quad (8.1)$$

Notre idée consiste à considérer que les phases  $\phi_k$  des sources sont des variables aléatoires qui sont relativement "proches" d'une première estimation (obtenue par déroulé linéaire) que nous notons  $\mu_k$ . À la différence de l'approche probabiliste présentée dans le chapitre précédent, notre but ici n'est pas d'estimer au sens MAP la phase, mais d'obtenir un estimateur MMSE des variables  $Z_k$ . Une façon naturelle de modéliser ce comportement serait de considérer que  $\phi_k$  est une gaussienne centrée en  $\mu_k$ . Néanmoins, comme la phase est une variable périodique, il est nécessaire d'utiliser une statistique circulaire. La loi normale périodique MARDIA et JUPP (2000) est une possibilité, et elle a été étudiée notamment dans AGIOMYRGIANNAKIS

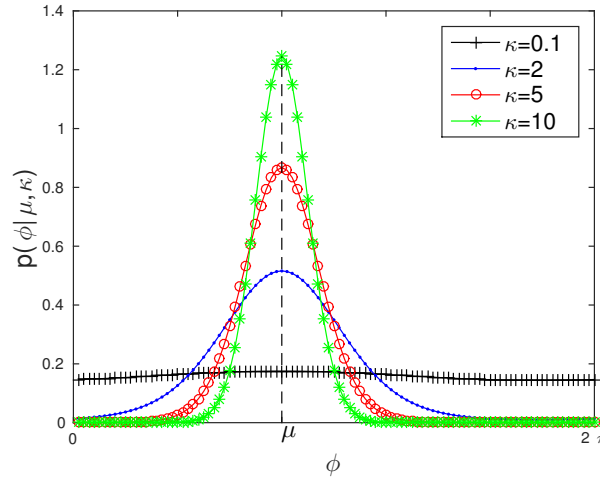


FIGURE 8.1 – Densité de probabilité d’une distribution de Von Mises pour un paramètre de localisation  $\mu$  et plusieurs valeurs du paramètre de concentration  $\kappa$ .

et [STYLIANOU \(2009\)](#) pour modéliser la phase dans une application de rehaussement de la parole. Néanmoins, cette loi ne possède pas de densité analytiquement simple (elle s’écrit comme une somme de série infinie), aussi nous avons choisi d’utiliser son approximation : la loi de Von Mises [MARDIA et ZEMROCH \(1975\)](#). Celle-ci est très populaire dans le domaine des statistiques circulaires, et ce pour trois raisons :

- Sa densité de probabilité est simple à écrire analytiquement ;
- Elle constitue une bonne approximation de la loi normale périodique ;
- C’est la statistique circulaire à entropie maximale (à paramètres de localisation et de concentration donnés), ce qui garantit une certaine régularité et donc de bonnes propriétés mathématiques.

La loi de Von Mises, notée  $\mathcal{VM}(\mu, \kappa)$  dépend de deux paramètres : un paramètre de localisation  $\mu \in [0, 2\pi[$  ( $\mu$  peut être réel, mais on retient en général sa valeur principale), qui joue le rôle de moyenne, et un paramètre de concentration  $\kappa \in ]0, +\infty[$  qui joue le rôle de l’inverse d’une variance. Sa densité, définie sur un intervalle de longueur  $2\pi$ , est donnée par :

$$p(\phi|\mu, \kappa) = \frac{e^{\kappa \cos(\phi-\mu)}}{2\pi I_0(\kappa)}, \quad (8.2)$$

où  $I_n$  désigne la fonction de Bessel modifiée de première espèce d’ordre  $n$  [WATSON \(1995\)](#). En particulier, on remarquera que :

- Lorsque  $\kappa \rightarrow 0$ , la concentration autour du mode  $\mu$  devient nulle : la loi de Von Mises devient alors équivalente à une loi uniforme.
- Lorsque  $\kappa \rightarrow +\infty$ , toute la masse est concentrée sur le mode  $\mu$  de la distribution : la densité tend alors vers un dirac centré en  $\mu$ , ce qui revient à dire que  $\phi$  n’est plus une variable aléatoire et devient déterministe égale à  $\mu$  (modulo  $2\pi$ ).

La figure 8.1 représente la densité de la loi de Von Mises pour plusieurs valeurs de  $\kappa$ .

Le modèle de mélange est donc :

$$X = \sum_k Z_k \text{ avec } Z_k = V_k e^{i\phi_k} \text{ et } \phi_k \sim \mathcal{VM}(\mu_k, \kappa_k). \quad (8.3)$$

---

### 8.1.2 Verrous du modèle

Connaissant les  $V_k$ , et supposant les  $\kappa_k$  fixés (par exemple estimés au préalable sur une base d'apprentissage) et les  $\mu_k$  estimés par déroulé linéaire, la tâche de séparation de sources consiste à obtenir un estimateur des composantes complexes  $Z_k$ . Les estimateurs les plus naturels sont donnés par l'estimateur du maximum de vraisemblance, l'estimateur du maximum à postériori, et l'espérance à postériori de ces variables sachant les observations.

Le problème est qu'en pratique, le calcul de ces estimateurs requiert nécessairement le calcul de la vraisemblance (la densité à postériori dépendant de la vraisemblance). Cela impose donc d'être capable d'estimer la loi de  $X$ , ainsi que la loi des  $Z_k$ . Or, malgré nos efforts calculatoires, nous ne sommes pas parvenus à les obtenir sous une forme analytique simple. En effet, bien que la densité de  $\phi_k$  soit aisée à écrire, la densité de  $X$  s'écrit sous la forme d'une intégrale d'ordre  $K - 1$  à l'intérieur de laquelle apparaissent des sommes de séries de Bessel.

L'estimation de l'espérance à postériori des variables  $Z_k$  pourrait être effectuée par des techniques de simulation telles que les méthodes de Monte Carlo par chaînes de Markov (MCMC) [ANDRIEU et al. \(2003\)](#); [ROBERT et CASELLA \(2013\)](#). L'idée principale est d'approcher l'espérance par la moyenne empirique calculée à partir d'échantillons suivant la loi dont on cherche l'espérance (méthodes de Monte Carlo). Comme on n'a pas forcément accès à cette loi (ici la loi à postériori des sources), on génère des échantillons selon un mécanisme de chaîne de Markov qui est construite de sorte à "ressembler" à cette loi (c'est le principe de l'algorithme de Metropolis-Hasting). Néanmoins, là encore, les algorithmes MCMC nécessitent certaines informations comme la loi à priori sur les variables. Or, celle-ci n'est pas exprimable analytiquement sous forme simple.

Ainsi, le modèle (8.3), malgré son apparente simplicité, semble trop complexe pour que les méthodes traditionnelles permettent d'obtenir un estimateur des sources. Par ailleurs, bien que des développements sur les méthodes MCMC pourraient permettre d'obtenir une solution, celles-ci sont basées sur la simulation et donc sur une procédure itérative, nécessitant un grand nombre d'itérations.

Nous proposons donc de l'approcher par un modèle simplifié, qui permette d'effectuer un certain nombre de calculs, tout en conservant la contrainte de phase qui motive cette approche.

### 8.1.3 Relations avec d'autres modèles

De tels modèles de phase pour tenir compte d'un modèle sinusoïdal ont été utilisés dans un cadre bayésien pour l'estimation de parole en milieu bruité, notamment dans [GERKMANN \(2014\)](#); [KULMER et MOWLAEE \(2015\)](#); [MOWLAEE et KULMER \(2015\)](#). La loi de Von Mises gagne donc depuis peu une certaine popularité en audio. Néanmoins, dans [MOWLAEE et KULMER \(2015\)](#), la phase est estimée au sens du Maximum à Postériori, alors que nous nous intéressons ici à l'obtention d'un estimateur MMSE des sources. Cette approche est conduite dans [GERKMANN \(2014\)](#), mais ce modèle est limité à un mélange parole plus bruit : aussi, les calculs deviennent plus aisés et il est possible d'obtenir une écriture de l'estimateur MMSE du signal de parole sous la forme d'une intégrale relativement simple à approcher numériquement. Dans notre cas (mélange de  $K$  sources), ce type d'approche ne fonctionne plus, d'où notre idée d'une approximation par modèle simplifié.

## 8.2 Modèle gaussien anisotrope

### 8.2.1 Loi normale complexe

L'idée que nous avons retenue consiste à approcher le modèle (8.3) par un modèle qui soit plus aisé à manipuler. Pour cela, nous avons considéré un modèle gaussien complexe non isotrope. En effet, les distributions gaussiennes sont omniprésentes en probabilités, avec de très nombreux développements et résultats théoriques rendant leur utilisation agréable. Elles apparaissent naturellement, comme c'est formalisé dans le théorème central limite notamment, comme une famille de référence pouvant approcher bon nombre de lois. L'idée qui consiste à approcher une somme de variables indépendantes à phases aléatoires par une somme de gaussiennes équivalentes n'est d'ailleurs pas neuve, comme en témoigne BECKMANN (1962). On propose donc le modèle suivant :

$$X = \sum_k X_k \text{ avec } X_k \sim \mathcal{N}(m_k, \gamma_k, c_k), \quad (8.4)$$

où les  $X_k$  sont indépendants, et  $m_k \in \mathbb{C}$ ,  $\gamma_k \in \mathbb{R}_+$  et  $c_k \in \mathbb{C}$  désignent respectivement la moyenne, la variance et le terme de *relation* de  $X_k$ , définis comme suit :

$$\begin{aligned} m_k &= \mathbb{E}(X_k), \\ \gamma_k &= \mathbb{E}(|X_k - m_k|^2), \\ c_k &= \mathbb{E}((X_k - m_k)^2). \end{aligned}$$

Nous définissons également la matrice de covariance  $\Gamma_k$  du vecteur  $\underline{X}_k = \begin{pmatrix} X_k \\ \bar{X}_k \end{pmatrix}$  :

$$\Gamma_k = \begin{pmatrix} \gamma_k & c_k \\ \bar{c}_k & \gamma_k \end{pmatrix}, \quad (8.5)$$

où  $\bar{\cdot}$  désigne le conjugué complexe. Nous rappelons également l'expression de la densité de probabilité d'une telle loi :

$$p(X_k | m_k, \gamma_k, c_k) = \frac{1}{\pi \sqrt{\det(\Gamma_k)}} e^{-\frac{1}{2}(\underline{X}_k - \underline{m}_k)^H \Gamma_k^{-1} (\underline{X}_k - \underline{m}_k)}, \quad (8.6)$$

où  $\det$  est le déterminant et  $\cdot^H$  désigne la transposée Hermitienne. En particulier, si  $m_k = 0$  et  $c_k = 0$ , on parle de loi normale complexe circulaire-symétrique. Si en outre la variance  $\gamma_k$  est égale à 1, il s'agit de la loi normale complexe standard (l'équivalent complexe de la loi normale centrée réduite).

La propriété d'additivité de la famille des lois normales implique alors :

$$X \sim \mathcal{N}(m_X, \gamma_X, c_X) \text{ avec } m_X = \sum_k m_k, \gamma_X = \sum_k \gamma_k \text{ et } c_X = \sum_k c_k, \quad (8.7)$$

la matrice de covariance de  $X$  étant égale à  $\Gamma_X = \sum_k \Gamma_k$ .

### 8.2.2 Paramètres du modèle

Afin de garantir la cohérence entre notre modèle initial et son approximation gaussienne, il convient que les moments des gaussiennes  $X_k$  soient identiques à ceux des sources  $Z_k$  dans

le modèle précédent de Von Mises :

$$\begin{aligned} m_k &= \mathbb{E}(X_k) = \mathbb{E}(Z_k), \\ \gamma_k &= \mathbb{E}(|X_k - m_k|^2) = \mathbb{E}(|Z_k - m_k|^2), \\ c_k &= \mathbb{E}((X_k - m_k)^2) = \mathbb{E}((Z_k - m_k)^2). \end{aligned}$$

La conservation de ces moments est donc ce qui permet de conserver, dans ce nouveau modèle, une propriété de non-uniformité des phases.

**Moyenne** Commençons par déterminer la moyenne  $m_k$  de chaque composante :

$$m_k = \mathbb{E}(X_k) = \mathbb{E}(Z_k) = V_k \mathbb{E}(e^{i\phi_k}). \quad (8.8)$$

Lorsque une variable est circulaire, on nomme *moment circulaire d'ordre n* la quantité  $\mathbb{E}(e^{in\phi_k})$ . Dans le cas de la loi de Von Mises, on ne sait pas déterminer la loi de  $e^{i\phi_k}$  mais on sait en revanche calculer ses moments circulaires **MARDIA et ZEMROCH (1975)** :

$$\forall n \in \mathbb{Z}, \mathbb{E}(e^{in\phi_k}) = \frac{I_{|n|}(\kappa_k)}{I_0(\kappa_k)} e^{in\mu_k}. \quad (8.9)$$

Ainsi, la moyenne de  $X_k$  est donnée par :

$$m_k = V_k \frac{I_1(\kappa_k)}{I_0(\kappa_k)} e^{i\mu_k} = \lambda_k \tilde{X}_k, \quad (8.10)$$

où l'on a posé  $\lambda_k = \frac{I_1(\kappa_k)}{I_0(\kappa_k)}$ , et  $\tilde{X}_k = V_k e^{i\mu_k}$  est la source estimée par déroulé linéaire appliqué isolément sur les sources.

**Variance** On cherche à présent à déterminer les termes de la matrice de covariance  $\Gamma_k$  de  $X_k$ , ce qui implique de déterminer les paramètres de variance et de relation. On a :

$$\begin{aligned} \gamma_k &= \mathbb{E}(|Z_k - m_k|^2) \\ &= \mathbb{E}(|Z_k|^2) - |m_k|^2 \\ &= V_k^2 - V_k^2 \frac{I_1(\kappa_k)^2}{I_0(\kappa_k)^2}, \end{aligned}$$

d'après le théorème de König-Huygens. Avec la notation  $\lambda_k$ , on a donc :

$$\gamma_k = (1 - \lambda_k^2) V_k^2. \quad (8.11)$$

**Terme de relation** De façon similaire, on calcule le terme de relation :

$$\begin{aligned} c_k &= \mathbb{E}((Z_k - m_k)^2) \\ &= \mathbb{E}(Z_k^2) - m_k^2 \\ &= V_k^2 \mathbb{E}(e^{i2\phi_k}) - V_k^2 \frac{I_1(\kappa_k)^2}{I_0(\kappa_k)^2} e^{i2\mu_k} \\ &= V_k^2 \frac{I_2(\kappa_k)}{I_0(\kappa_k)} e^{i2\mu_k} - V_k^2 \frac{I_1(\kappa_k)^2}{I_0(\kappa_k)^2} e^{i2\mu_k} \\ &= \rho_k \tilde{X}_k^2. \end{aligned}$$

où l'on a introduit la notation suivante :

$$\rho_k = \frac{I_2(\kappa_k)I_0(\kappa_k) - I_1(\kappa_k)^2}{I_0(\kappa_k)^2}. \quad (8.12)$$

**Mélange** L'intérêt du modèle gaussien est que l'on peut sommer les distributions précédemment identifiées en vertu de l'équation (8.7). On a donc :

$$m_X = \sum_k \lambda_k \tilde{X}_k, \quad (8.13)$$

$$\gamma_X = \sum_k (1 - \lambda_k^2) V_k^2, \quad (8.14)$$

$$c_X = \sum_k \rho_k \tilde{X}_k^2. \quad (8.15)$$

### 8.2.3 Comparaison des deux modèles

Afin de comparer les deux modèles (Von Mises et gaussien anisotrope), nous générons 10000 échantillons de chaque variable aléatoire  $Z_k$  et  $X_k$ , et ce pour différentes valeurs du paramètre de concentration. Nous fixons le module à 1 et le paramètre de localisation de phase à  $\pi/3$ . Nous représentons les histogrammes 2D de ces distributions sur la figure 8.2. On observe tout d'abord que les échantillons générés dans le modèle de Von Mises sont tous situés sur le cercle de rayon 1 : c'est logique puisque dans ce modèle, l'amplitude ne varie pas. L'approximation par modèle gaussien introduit une incertitude sur cette amplitude, qui est liée aussi bien au module qu'à la phase, mais également au paramètre de concentration. Plus celui-ci est élevé, plus la variation d'amplitude devient faible. On remarque enfin que plus le paramètre de concentration est élevé, plus les deux distributions ont tendance à se ressembler.

### 8.2.4 Estimateur MMSE des sources

#### Obtention de l'estimateur

Maintenant que l'on a accès à tous les éléments du modèle, on peut fournir un estimateur MMSE des sources, donné par l'espérance a posteriori de celles-ci. Dans le cadre d'un modèle gaussien, il existe une formule simple qui donne cette espérance (voir par exemple BISHOP (2006)) :

$$\hat{X}_k = \underline{m}_k + \Gamma_k \Gamma_X^{-1} (\underline{X} - \underline{m}_X). \quad (8.16)$$

L'expression ci-dessus donne l'estimateur  $\hat{X}_k$  ainsi que son conjugué  $\bar{\hat{X}}_k$ . Ainsi, pour éviter cette redondance d'information, on ne considère que la première ligne de (8.16) :

$$\hat{X}_k = m_k + (\gamma_k \quad c_k) \Gamma_X^{-1} \begin{pmatrix} X - m_X \\ \bar{X} - \bar{m}_X \end{pmatrix} \quad (8.17)$$

On constate d'ores et déjà que cet estimateur est conservatif. En effet, comme  $\sum_k \Gamma_k = \Gamma_X$ , on a :

$$\sum_k \hat{X}_k = \sum_k m_k + (\underline{X} - \underline{m}_X) = \underline{X}. \quad (8.18)$$

Cette propriété garantit que quelle que soit la valeur des  $\kappa_k$ , on est assuré que la somme des estimateurs restera égale au mélange, ce qui ne serait pas le cas si on avait considéré les estimateurs obtenus par déroulé linéaire isolé  $\tilde{X}_k$ .

#### Interprétation

L'estimateur (8.16) réalise en fait une interpolation entre le filtrage de Wiener (qui n'utilise que la phase du mélange), et l'estimateur construit en utilisant  $\mu_k$  uniquement (et pas la phase du mélange). Considérons ici que  $\forall k$ , les  $\kappa_k$  sont constants et égaux à une valeur  $\kappa$ . On

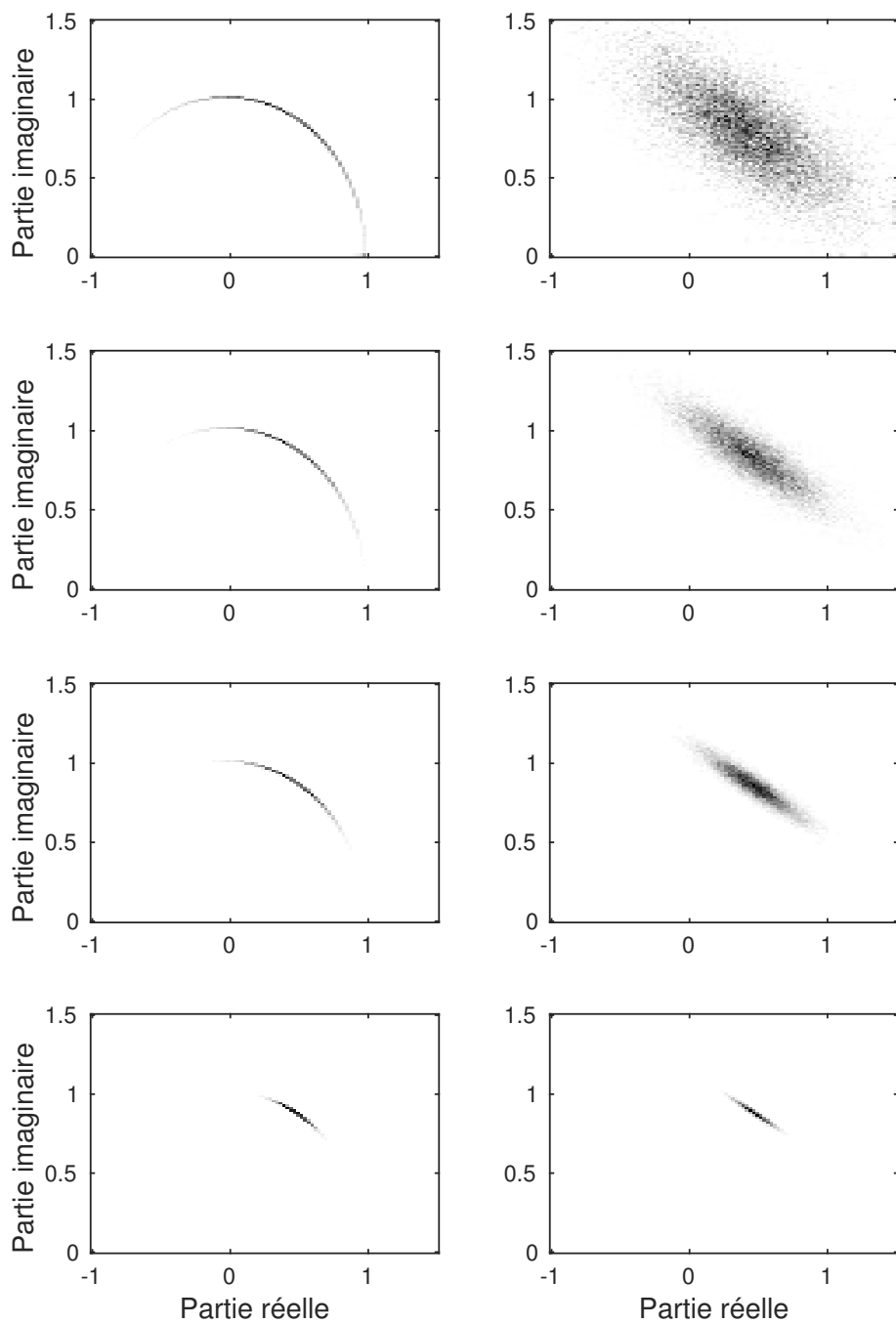
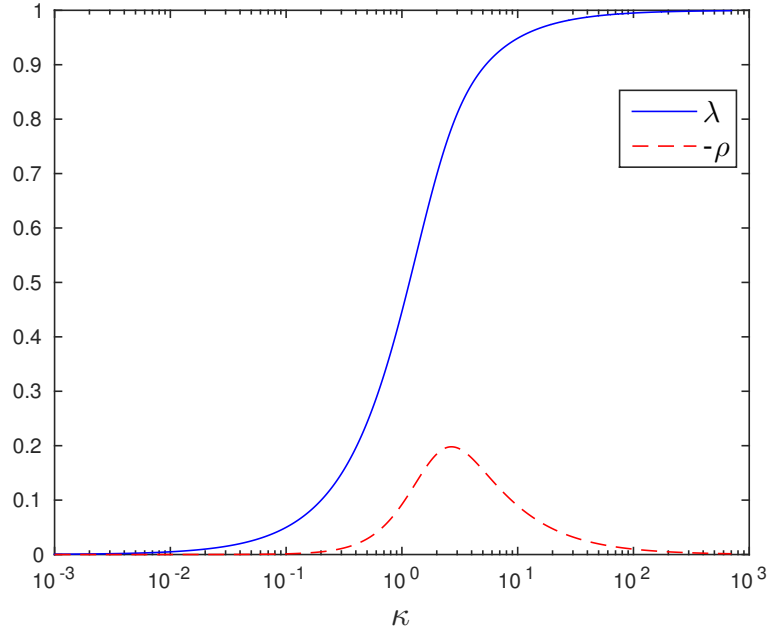


FIGURE 8.2 – Histogrammes en 2D d'échantillons générés par le modèle de Von Mises (gauche) et le modèle équivalent gaussien (droite), pour  $V = 1$  et  $\mu = \pi/3$ . Le paramètre de concentration  $\kappa$  prend, de haut en bas, les valeurs 5, 10, 20 et 100.

désignera alors par  $\lambda$  et  $\rho$  les paramètres correspondants, et on illustre le comportement de ceux-ci en fonction de  $\kappa$  sur la figure 8.3.




 FIGURE 8.3 – Evolution des paramètres  $\lambda$  et  $\rho$  en fonction de  $\kappa$ .

Supposons que le paramètre de concentration  $\kappa$  tende vers 0. On montre alors, en utilisant les propriétés des fonctions de Bessel [WATSON \(1995\)](#) (et comme c'est suggéré sur la figure 8.3) que  $\lambda \rightarrow 0$  et  $\rho \rightarrow 0$ . On a alors  $m_k \rightarrow 0$ ,  $c_k \rightarrow 0$  et  $\gamma_k \rightarrow V_k^2$ . Ainsi, on constate que l'estimateur (8.16) tend vers le filtrage de Wiener :

$$\hat{X}_k \xrightarrow{\kappa \rightarrow 0} \begin{pmatrix} V_k^2 & 0 \\ \sum_l V_l^2 & 1 \\ 0 & \sum_l V_l^2 \end{pmatrix} \begin{pmatrix} X \\ \bar{X} \end{pmatrix} = \frac{V_k^2}{\sum_l V_l^2} X, \quad (8.19)$$

où l'on reconnaît le gain de Wiener traditionnel  $G_k = \frac{V_k^2}{\sum_l V_l^2}$ . Ce résultat est logique, puisqu'une valeur nulle du paramètre de concentration dans la distribution de Von Mises correspond, comme on l'a vu, à une distribution uniforme. Cela revient à dire que l'on a aucune information a priori sur la phase, et on obtient finalement un modèle gaussien isotrope classique [FÉVOTTE et al. \(2009\)](#) pour lequel on sait que l'estimateur MMSE des sources est donné par le filtrage de Wiener traditionnel.

Les deux estimateurs précédents (filtrage de Wiener et déroulé linéaire appliqué isolément sur les sources) ont montré leurs limites, comme nous avons eu plusieurs fois l'occasion de le rappeler. Le filtrage de Wiener introduit en effet des interférences lorsque les signaux sont recouvrants dans le domaine TF, et le déroulé linéaire seul, appliqué isolément sur les sources, a tendance à créer des artefacts du fait des erreurs d'estimation sur les fréquences instantanées (voir chapitre 4). L'estimateur (8.17) que nous proposons est un compromis entre ces deux approches, dont la performance dépendra du réglage du paramètre  $\kappa$ .

De façon intéressante, ce type d'estimateur qui combine optimalement filtrage de Wiener et modèle de phase par déroulé linéaire a déjà été introduit dans [KRAWCZYK et GERKMANN \(2015\)](#) dans un but de débruitage de la parole, néanmoins sans utiliser un modèle de Von Mises.

---

## Algorithme de séparation de sources

Comme on considère que tous les points TF sont indépendants, on pourrait appliquer l'estimateur 8.17 matriciellement et ainsi reconstruire directement toutes les sources. Néanmoins, similairement à ce que nous avons constaté dans le chapitre 5, il est préférable d'adopter une approche séquentielle sur les trames temporelles : en effet, mieux vaut estimer  $\mu_k(f, t)$  par déroulé linéaire à partir de  $\phi_k(f, t - 1) = \angle \hat{X}_k(f, t - 1)$ , qui est un estimateur de la phase dans la trame précédente où l'erreur sur  $\mu_k$  est "corrigée" par la phase du mélange, plutôt que directement à partir de  $\mu_k(f, t - 1)$ . Autrement dit, on choisit de calculer séquentiellement les estimateurs afin d'éviter au maximum de propager les erreurs d'estimation qui proviennent du déroulé linéaire.

L'algorithme 10 résume cette procédure de séparation de sources dans un cas informé (les amplitudes sont préalablement estimées). On rappelle que  $\Omega_k$  désigne l'ensemble des trames d'attaque pour la source  $k$ .

### Un cas particulier : le modèle parole plus bruit

Comme nous l'avons rappelé plus haut, les modèles de mélanges de sources utilisant la loi de Von Mises ont déjà été étudiés en audio, notamment dans le cadre du débruitage de la parole GERKMANN (2014); KULMER et MOWLAEE (2015). On peut voir ce type de mélanges comme un cas particulier de notre modèle avec  $K = 2$  sources, en supposant néanmoins que les paramètres de concentration ne sont pas les mêmes pour les deux sources. En effet, on a  $X = \tilde{S} + B$ ,  $\tilde{S}$  correspondant au signal de parole, que l'on peut écrire sous la forme  $V e^{i\phi}$ ,  $\phi$  suivant une loi de Von Mises de paramètres  $\kappa$  et  $\mu$  (phase obtenue par déroulé linéaire).  $B$  est généralement considéré comme étant un bruit blanc gaussien.

Ainsi, on peut utiliser notre approximation par mélange de gaussiennes anisotropes.  $\tilde{S}$  est donc approchée par  $S$ , qui est une variable gaussienne dont les paramètres sont décrits par les équations données dans la section 8.2.2. Le bruit suit également ce modèle, à condition que l'on choisisse un paramètre de concentration nul (on a alors une gaussienne complexe circulaire-symétrique).

À la différence de GERKMANN (2014) qui obtient un estimateur MMSE directement à partir du modèle de Von Mises, le notre est obtenu à partir d'une approximation gaussienne. Néanmoins, il ne requiert pas le calcul numérique d'une intégrale, aussi on peut supposer qu'il soit plus efficace sur le plan du temps de calcul. L'application de ce modèle au débruitage de la parole pourra donc être envisagée à l'avenir.

**Algorithme 10** Séparation de sources par application d'un estimateur MMSE combinant filtrage de Wiener et déroulé linéaire.

**Entrées :**

Mélange  $X \in \mathbb{C}^{F \times T}$ ,

Spectrogrammes  $V_k \in \mathbb{R}_+^{F \times T}$ ,  $\forall k \in \llbracket 1, K \rrbracket$ ,

Ensembles des trames d'attaque  $\Omega_k$ ,  $\forall k \in \llbracket 1, K \rrbracket$ ,

Phases d'attaque  $\phi_k^o(f, t)$ ,  $\forall t \in \Omega_k$ ,

Paramètre de concentration  $\kappa_k \in \mathbb{R}_+^{F \times T}$ .

**pour**  $t = 1$  à  $T - 1$  **faire**

**pour**  $k = 1$  à  $K$  **faire**

**si**  $t \in \Omega_k$  **alors**

            Phase d'attaque :  $\forall f$ ,  $\mu_k(f, t) = \phi_k^o(f, t)$ .

**sinon**

$\mu_k(f, t)$  est obtenue par déroulé linéaire (*cf.* Algorithme 2).

**fin si**

**fin pour**

Pour toutes les sources  $k$  et tous les canaux fréquentiels  $f$  :

**À priori**  $\tilde{X}_k(f, t) = V_k(f, t)e^{i\mu_k(f, t)}$ .

**Paramètres**  $\lambda_k(f, t) = \frac{I_1(\kappa_k(f, t))}{I_0(\kappa_k(f, t))}$  et  $\rho_k(f, t) = \frac{I_2(\kappa_k(f, t))I_0(\kappa_k(f, t)) - I_1(\kappa_k(f, t))^2}{I_0(\kappa_k(f, t))^2}$ .

**Moments**

$m_k(f, t) = \lambda_k(f, t)\tilde{X}_k(f, t)$ ,  $m_X(f, t) = \sum_k m_k(f, t)$ ,

$\gamma_k(f, t) = (1 - \lambda_k(f, t)^2)V_k^2$ ,  $\gamma_X(f, t) = \sum_k \gamma_k(f, t)$ ,

$c_k(f, t) = \rho_k(f, t)\tilde{X}_k(f, t)^2$ ,  $c_X(f, t) = \sum_k c_k(f, t)$ .

**Estimateur MMSE**

$$\hat{X}_k(f, t) = m_k(f, t) + \begin{pmatrix} \gamma_k(f, t) & c_k(f, t) \end{pmatrix} \begin{pmatrix} \gamma_X(f, t) & c_X(f, t) \\ \bar{c}_X(f, t) & \gamma_X(f, t) \end{pmatrix}^{-1} \begin{pmatrix} X(f, t) - m_X(f, t) \\ \bar{X}(f, t) - \bar{m}_X(f, t) \end{pmatrix}$$

**fin pour**

**Sortie :**

$\forall k \in \llbracket 1, K \rrbracket$ ,  $\hat{X}_k \in \mathbb{C}^{F \times T}$ .

---

## 8.3 Validation expérimentale

Dans cette partie, nous conduisons une série d'expériences pour montrer le potentiel de l'estimateur (8.17). Les données utilisées proviennent de la base DSD100 [ONO et al. \(2015\)](#), constituée de 100 morceaux de musique polyphonique composés de  $K = 4$  sources. Celle-ci est divisée en deux moitiés (une pour l'apprentissage du paramètre de concentration optimal et une autre pour l'expérience de séparation de sources). Les signaux sont échantillonnés à  $F_s = 44100$  Hz et la TFCT est calculée avec une fenêtre de Hann de longueur 4096 échantillons et 75 % de recouvrement. Les amplitudes  $V_k$  sont soit supposées connues (cas Oracle) soit estimées par NMF sur les sources séparées (cas semi-Oracle). On initialise les phases dans les trames d'attaque en leur donnant celle du mélange. La qualité de la séparation est mesurée par le SDR, le SIR et le SAR calculés avec la boîte à outils BSS Eval [VINCENT et al. \(2006\)](#).

### 8.3.1 Influence du paramètre de concentration

Tout d'abord, nous considérons 50 morceaux issus de la base de données (base de développement) et appliquons l'algorithme 10 pour estimer les sources séparées à partir des mélanges en faisant varier le paramètre de concentration  $\kappa$ . Nous examinons deux cas de figure : soit le paramètre de concentration  $\kappa$  est le même pour tous les points TF, soit il est variable, auquel cas on utilise un paramètre de concentration  $\tilde{\kappa}_k(f, t)$  tel que :

$$\tilde{\kappa}_k(f, t) = \kappa(1 - G_k(f, t)), \quad (8.20)$$

où  $G_k(f, t)$  est le gain de Wiener. En effet, ce choix est motivé par le fait que plus le gain de Wiener est important, plus une source est prédominante dans un point TF considéré. Dans ce cas, on peut s'attendre à ce que la phase du mélange soit très proche de la phase de la source dominante, auquel cas la phase de  $\tilde{X}_k$  dans l'estimateur des sources devient moins importante que celle du mélange  $X$ .

Les résultats dans le cas Oracle sont présentés sur la figure 8.4. On constate tout d'abord que lorsque  $\kappa$  devient proche de 0, les performances de l'estimateur deviennent similaires à celles obtenues par filtrage de Wiener. Cela confirme notre interprétation de la section 8.2.4 et montre que l'estimateur (8.17) est une généralisation du filtrage de Wiener. Nous observons ce qui semble être un maximum de SDR et de SAR autour de  $\kappa = 1.6$ , qui permet d'augmenter les performances par rapport à Wiener d'environ 2 dB pour ces indicateurs, et d'environ 3 dB pour le SIR. Le SIR ne semble pas maximisé pour cette valeur : en effet, augmenter le paramètre de concentration continue d'augmenter le SIR. C'est logique puisque plus ce paramètre est élevé, moins la phase du mélange devient importante, ce qui implique une réduction des interférences entre les sources. Néanmoins, on constate que les SDR et SAR décroissent après ce pic, ce qui là encore est expliqué par le fait qu'une concentration trop forte autour de  $\mu_k$  par rapport à la donnée de phase du mélange peut avoir tendance à créer des artéfacts (dûs aux erreurs d'estimation et aux écarts au modèle lorsqu'on applique le déroulé linéaire).

On constate par ailleurs que les paramètres de concentration variables (8.20) ne donnent pas globalement de meilleurs résultats qu'un paramètre constant. On pourrait cependant chercher une nouvelle définition de ces paramètres qui potentiellement mènerait à de meilleurs résultats, car le choix (8.20) était arbitraire.

Les résultats dans le scénario semi-Oracle, présentés sur la figure 8.5, sont similaires, et montrent qu'il est préférable, lorsque les amplitudes ne sont plus exactement égales à la vérité terrain, d'utiliser un paramètre de concentration constant  $\kappa = 1$  plutôt qu'un paramètre variable.

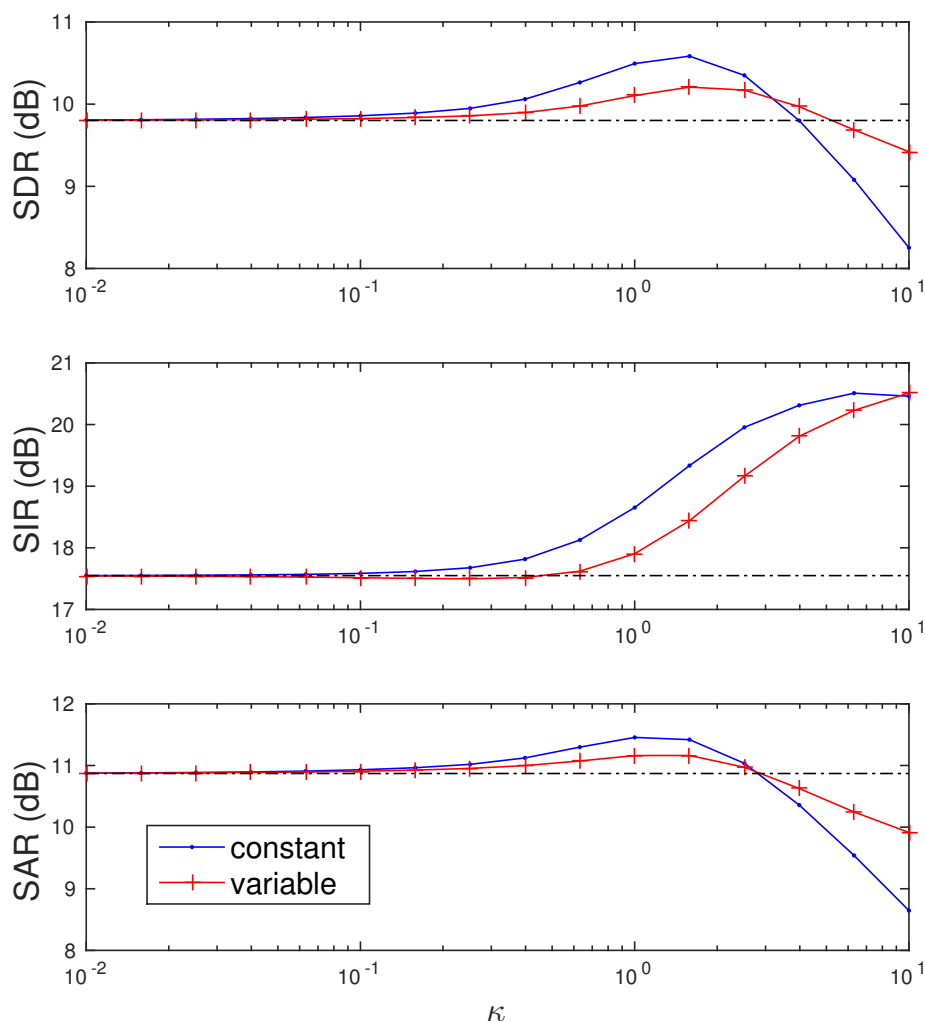


FIGURE 8.4 – Influence du paramètre de concentration  $\kappa$  sur la qualité de la séparation de sources dans l’algorithme 10 (courbes en traits pleins) et comparaison au filtrage de Wiener (courbe en pointillés) dans le cas Oracle. Le paramètre de concentration peut être constant pour toutes les sources et les points TF ou bien variable, selon l’équation (8.20).

Cette première expérience montre l’intérêt de notre estimateur. Pour la suite de nos expériences, nous utiliserons dans le cas Oracle la valeur  $\kappa = 1.6$ , qui semble être un bon compromis entre les différents indicateurs. Dans le cas semi-Oracle, on prendra  $\kappa = 1$ .

### 8.3.2 Séparation de sources musicales

Nous considérons à présent la deuxième moitié de la base de données (50 morceaux constituant la base de tests) et les valeurs du paramètre de concentration apprises précédemment. Nous appliquons le même protocole que dans l’expérience précédente, et comparons les résultats de notre approche (notée **MMSE**) avec plusieurs autres méthodes :

- le filtrage de Wiener [FÉVOTTE et al. \(2009\)](#) qui sera noté **Wiener** ;
- le filtrage de Wiener consistant [LE ROUX et VINCENT \(2013\)](#), noté **W-Cons**. Confor-

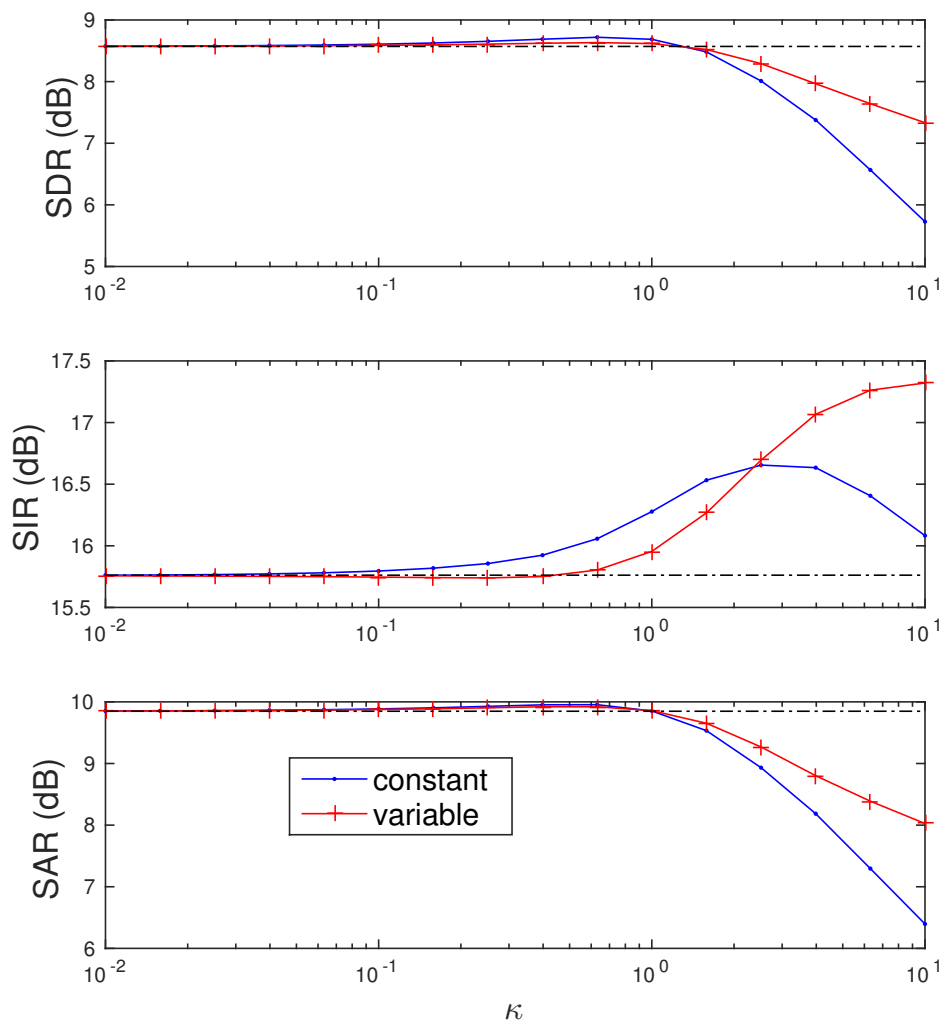


FIGURE 8.5 – Influence du paramètre de concentration  $\kappa$  sur la qualité de la séparation de sources dans l’algorithme 10 (courbes en traits pleins) et comparaison au filtrage de Wiener (courbe en pointillés) dans le cas semi-Oracle. Le paramètre de concentration peut être constant pour toutes les sources et les points TF ou bien variable, selon l’équation (8.20).

mément aux résultats obtenus au chapitre 5 dans la section 5.3.1, nous choisissons un paramètre de consistance égal à 4 pour cet algorithme.

- l’estimateur n’utilisant que le déroulé de phase  $\tilde{X}_k$ , c’est-à-dire sans tenir compte de la phase du mélange. Il sera noté **Unwrap**.

Les résultats sont présentés sur la figure 8.6. Nous constatons qu’aussi bien dans le cas Oracle que semi-Oracle, l’estimateur que nous proposons fournit des résultats légèrement meilleurs que le filtrage de Wiener, et légèrement moins bons que le filtrage de Wiener consistant en SDR, SIR et SAR. La méthode **Unwrap** ne conduit pas à des résultats satisfaisants, notamment au niveau du SDR et du SAR. On peut interpréter ceci comme au chapitre 5 : l’application du déroulé sans tenir compte de la phase du mélange a tendance à faire se propager l’erreur de déroulé linéaire sans le corriger au fur et à mesure, ce qui crée des artéfacts qui sont perceptibles. Le SIR, en revanche, n’est pas trop impacté, puisque cet estimateur

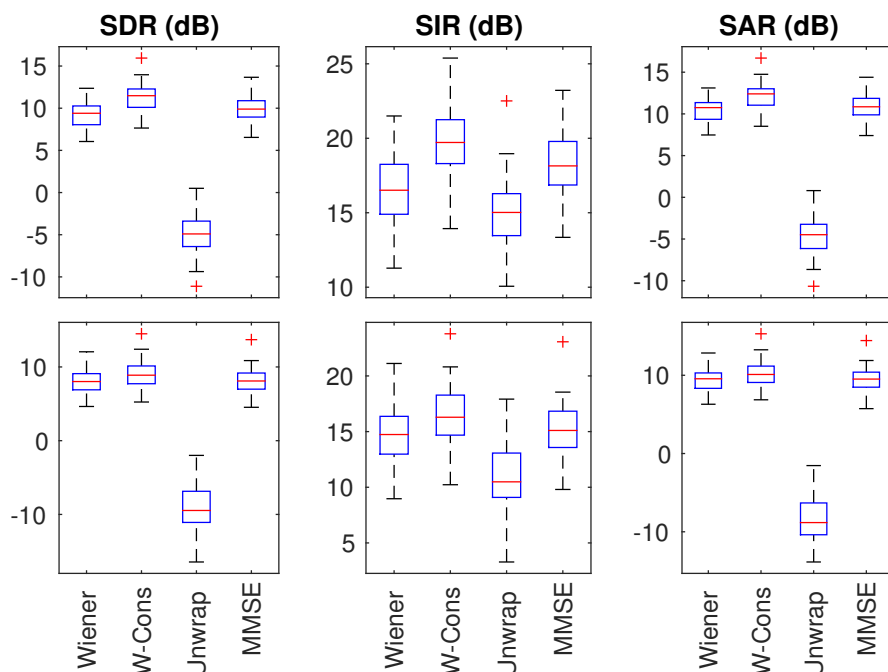


FIGURE 8.6 – Performance de la séparation de sources (SDR, SIR et SAR en dB) pour diverses méthodes sur la base DSD100. Résultats Oracle (en haut) et sur spectrogrammes approchés (en bas).

	Scénario Oracle			Scénario semi-Oracle		
	SDR	SIR	SAR	SDR	SIR	SAR
<b>Iter</b>	<b>10.0</b>	<b>20.5</b>	<b>10.4</b>	5.3	<b>15.0</b>	5.9
<b>MMSE</b>	9.0	16.7	9.9	<b>7.4</b>	13.9	<b>8.6</b>

TABLEAU 8.1 – Performance de reconstruction (SDR en dB) pour divers jeux de données.

agissant sur les sources séparées, il crée moins d'interférences entre sources que les autres méthodes, qui utilisent la phase du mélange. Du point de vue perceptif, nous n'avons pas entendu de différence significative entre les pistes `vocals`, `other` et `drum` estimées par filtrage de Wiener consistant et par notre approche, néanmoins nous avons noté une nette diminution des artéfacts dans la piste `bass` par notre méthode. Enfin, soulignons que notre approche est nettement moins gourmande que **W-Cons** en temps de calcul (d'un facteur 7).

Nous effectuons enfin une comparaison entre la méthode présentée au chapitre 5 (méthode **Iter**) et celle introduite ici (**MMSE**). Les résultats moyennés sur les 50 morceaux de la base de données sont présentés dans le tableau 8.1. Nous constatons que dans le cas Oracle, l'approche itérative donne les meilleurs résultats pour tous les indicateurs. Néanmoins, lorsque les amplitudes ne sont plus égales à la vérité terrain, l'estimateur **MMSE** conduit à un SDR et un SAR plus élevés, bien que **Iter** donne toujours un meilleur SIR. Cela s'explique par le fait que l'estimateur **MMSE** tient compte d'une incertitude sur les amplitudes, alors que dans l'approche itérative, celles-ci sont fixées à chaque itération à la valeur objectif, qui n'est plus exacte dans le cas semi-Oracle. Les deux approches utilisent un temps de calcul similaire.

En fin de compte, on peut résumer ces conclusions comme suit :

- Si on est soucieux d'implémenter une méthode très rapide, quitte à perdre en qualité,

---

on utilisera le filtrage de Wiener ;

- Si les spectrogrammes qui ont été préalablement estimés sont considérés comme étant d'excellente qualité, on pourra appliquer la procédure **Iter** ;
- Si on s'éloigne de cette valeur Oracle, il est plus prudent d'appliquer l'estimateur **MMSE** introduit dans ce chapitre.

Enfin, le filtrage de Wiener consistant donne de bons résultats, mais reste limitant du point de vue du temps de calcul.

## 8.4 Conclusion

Le modèle introduit dans ce chapitre se base sur une hypothèse de phase non-uniforme, contrairement à ce qui est fréquemment supposé dans la littérature. Du modèle gaussien découle un estimateur MMSE des sources qui allie bonne qualité de séparation et rapidité de calcul. Nous avons expérimentalement démontré son potentiel pour une application de séparation de sources. Nous avons proposé que le paramètre de localisation de la phase soit estimé par déroulé linéaire, mais n'importe quel  $\mu_k$  pourrait en réalité être utilisé dans le modèle.

Bien que l'on puisse considérer que l'approche itérative introduite au chapitre 5 mène à des résultats légèrement meilleurs dans le cas Oracle, l'avantage de ce modèle est qu'il pose les bases d'une méthode complète de séparation de sources. En effet, les amplitudes étaient ici supposées estimées par avance, et nous n'avons pas directement modélisé l'incertitude sur ces estimées (elle a été introduite par le passage au modèle gaussien anisotrope). On pourrait alors envisager de modéliser les  $V_k$  comme des variables aléatoires suivant une loi à support non-négatif (loi de Poisson, de Rayleigh, Gamma...), et structurer les paramètres de dispersion de celle-ci par un modèle de type NMF. Des applications en séparation aveugle de sources pourront alors être envisagées.



# Chapitre 9

## Modélisation robuste de données non-négatives

### Sommaire

---

<b>9.1 Distributions Positives <math>\alpha</math>-stables</b> . . . . .	<b>147</b>
9.1.1 Lois stables . . . . .	147
9.1.2 Variables à support non-négatif . . . . .	147
9.1.3 Distribution de Lévy . . . . .	148
<b>9.2 Modèle de Lévy NMF</b> . . . . .	<b>150</b>
9.2.1 Modèle de mélange . . . . .	150
9.2.2 Estimation du modèle . . . . .	150
9.2.3 Algorithmes de Lévy NMF . . . . .	159
9.2.4 Interprétation des mises à jour . . . . .	159
<b>9.3 Estimateur des sources</b> . . . . .	<b>160</b>
9.3.1 Somme de 2 variables . . . . .	160
9.3.2 Somme de $K$ variables . . . . .	164
<b>9.4 Expériences</b> . . . . .	<b>164</b>
9.4.1 Algorithmes de Lévy NMF . . . . .	165
9.4.2 Représentation de bruits impulsionnels . . . . .	166
9.4.3 Applications . . . . .	167
<b>9.5 Conclusion</b> . . . . .	<b>172</b>

---

---

Dans la deuxième partie de ce manuscrit, nous nous sommes intéressés à des modèles de signaux qui permettent d'obtenir des contraintes de phase dans un but de reconstruction de celle-ci. Au chapitre précédent, nous avons proposé un modèle probabiliste qui prend en compte les à priori que l'on peut obtenir sur la phase grâce à de tels modèles. Néanmoins, nous avons jusqu'à présent supposé que les amplitudes des sources étaient connues (cas Oracle) ou bien estimées par avance. En tout cas, elles étaient déterministes. Dans le but de mettre au point un modèle de sources complet pour une application de séparation de sources réaliste, nous allons chercher à modéliser les amplitudes des sources. En outre, nous avons constaté expérimentalement (aux chapitres 5 et 8 en particulier) que la performance des techniques de reconstruction de phase dépend de la qualité des amplitudes estimées préalablement.

C'est pourquoi, dans ce chapitre, nous nous intéressons à la modélisation de données non-négatives. En particulier, nous mettons en évidence le fait que les lois usuelles (telles que la loi de Poisson [VIRTANEN et al. \(2008\)](#); [CEMGIL \(2009\)](#)) employées pour modéliser des données non-négatives, notamment en traitement du signal audio, ne sont pas à *queue lourde* : cela implique que de telles distributions représentent mal des valeurs éloignées de leurs modes et qui sont considérées comme "abbérantes", alors que de telles valeurs sont pourtant fréquentes en audio. Les lois  $\alpha$ -stables [NOLAN \(2015\)](#), en revanche, présentent cette propriété de robustesse aux valeurs abbérantes tout en conservant la propriété d'additivité des lois de Poisson. Ces distributions ont été utilisées avec succès dans des applications de traitement du signal robuste, notamment en audio [SAMORADNITSKY et TAQQU \(1994\)](#); [GODSILL et KURUOGLU \(1999\)](#); [BASSIOU et al. \(2013\)](#). Nous introduisons un sous-ensemble des lois stables : les distributions Positives  $\alpha$ -stables (P $\alpha$ S), qui sont à support non-négatif, et peuvent donc modéliser des données non-négatives telles que des amplitudes de TFCT de signaux audio.

Après avoir introduit cette famille de lois, nous en étudions un cas particulier : la distribution de Lévy. Il s'agit du seul cas de loi P $\alpha$ S pour lequel nous pouvons écrire analytiquement la densité de probabilité. Nous proposons un modèle de mélange de données non-négatives appelé *Lévy NMF*, qui est appliqué à la séparation de sources non-négatives. Ceci ne porte que sur un sous-problème particulier de notre objectif final, qui consiste à séparer des sources complexes (en utilisant notamment un modèle de données non-négatives), mais nous avons trouvé intéressant d'étudier plus en détail ce modèle, car il s'agit d'une thématique de recherche active. La séparation de sources non-négatives est en effet un problème fondamental dans de nombreux domaines tels que l'imagerie par résonance magnétique [SAJDA et al. \(2004\)](#), la reconnaissance de visage [GUILLAMET et VITRIA \(2002\)](#) ou la fouille de données textuelles [PAUCA et al. \(2004\)](#). Nous proposons des applications en débruitage de spectrogrammes audio et en séparation de sources en spectroscopie de fluorescence [LIU et al. \(2013\)](#).

Un des résultats forts de ce chapitre, la généralisation du filtrage de Wiener aux variables P $\alpha$ S, a fait l'objet d'une publication dans un rapport technique déposé dans la base de données de Télécom ParisTech [MAGRON et al. \(2016c\)](#). Le modèle de Lévy NMF ainsi que la partie expérimentale ont fait l'objet d'un article soumis dans la revue *IEEE Signal Processing Letters* [MAGRON et al. \(2017c\)](#). Ce travail ayant été mené conjointement avec Antoine Liutkus de l'INRIA Nancy, nous profitons de ce préambule pour le remercier.

Dans la section 9.1, nous introduisons les distributions P $\alpha$ S. Dans la section 9.2, nous présentons un modèle de mélange de données non-négatives, dont nous estimons les paramètres par différentes méthodes. Dans la section 9.3, nous fournissons une justification théorique à l'utilisation d'un filtrage de Wiener généralisé pour estimer les sources de mélanges P $\alpha$ S. Une validation expérimentale est proposée dans la section 9.4 avec des exemples d'application en audio et en spectroscopie de fluorescence. Enfin, nous concluons dans la section 9.5.

## 9.1 Distributions Positives $\alpha$ -stables

### 9.1.1 Lois stables

Les distributions  $\alpha$ -stables [NOLAN \(2015\)](#), notées  $\mathcal{S}(\alpha, \sigma, \mu, \beta)$ , sont définies comme l'ensemble des distributions sur la variable aléatoire  $X$  à valeurs dans  $\mathbb{R}$  dont la fonction caractéristique est :

$$\varphi_X(t) = \mathbb{E}(e^{itX}) = \begin{cases} e^{it\mu - \sigma^\alpha |t|^{\alpha(1-i\beta\Phi sg(t))}} & \text{si } \alpha \neq 1, \\ e^{it\mu - \sigma |t|^{1+i\beta\frac{2}{\pi} \log(|t|)sg(t)}} & \text{si } \alpha = 1, \end{cases} \quad (9.1)$$

où  $sg(t)$  désigne le signe de  $t \in \mathbb{R}$ , et  $\Phi = \tan(\frac{\pi\alpha}{2})$ . Elles dépendent de 4 paramètres :

- Un exposant caractéristique ou paramètre de forme  $\alpha \in ]0, 2]$  qui détermine la forme de la queue de la distribution (pour des valeurs faibles de  $\alpha$ , la distribution est dite à queue lourde) ;
- Un paramètre d'échelle  $\sigma \in ]0, +\infty[$  qui mesure la dispersion de la distribution autour de son mode ;
- Un paramètre de localisation  $\mu \in \mathbb{R}$  ;
- Un paramètre d'asymétrie  $\beta \in [-1, 1]$ .

La propriété de stabilité se traduit par le fait que si  $K$  variables  $X_k$  suivent une loi  $\alpha$ -stable et sont indépendantes, alors la somme de ces variables suit également une loi  $\alpha$ -stable. En particulier, si  $X_k \sim \mathcal{S}(\alpha, \sigma_k, \mu_k, \beta_k)$ , alors :

$$X = \sum_k X_k \sim \mathcal{S}(\alpha, \sigma, \mu, \beta) \text{ avec } \begin{cases} \sigma^\alpha = \sum_k \sigma_k^\alpha, \\ \mu = \sum_k \mu_k, \\ \beta = \frac{\sum_k \beta_k \sigma_k^\alpha}{\sum_k \sigma_k^\alpha}. \end{cases} \quad (9.2)$$

Cette propriété est utilisée pour manipuler les grandeurs  $X$  et obtenir un certain nombre de résultats. Elle donne des informations sur la densité du mélange, ce qui permet l'application de méthodes de type ML ou MAP pour l'estimation des paramètres.

Les distributions dont le paramètre d'asymétrie  $\beta$  est nul sont dites Symétriques  $\alpha$ -stables (S $\alpha$ S) [LIUTKUS et BADEAU \(2015\)](#), notée  $\mathcal{S}\alpha\mathcal{S}(\alpha, \sigma, \mu) = \mathcal{S}(\alpha, \sigma, \mu, 0)$ . Ces distributions sont à support dans  $\mathbb{R}$  et peuvent être étendues à  $\mathbb{C}$  en raisonnant sur les parties réelles et imaginaires des variables complexes considérées (lois S $\alpha$ S isotropes).

### 9.1.2 Variables à support non-négatif

Les distributions S $\alpha$ S ne sont pas adaptées à la modélisation de données non-négatives puisqu'elles sont à support dans  $\mathbb{R}$  (ou  $\mathbb{C}$ ). Nous proposons d'introduire un autre cas particulier de distributions  $\alpha$ -stables qui soient à support non-négatif : les distributions Positives  $\alpha$ -stables (P $\alpha$ S).

Lorsque  $\beta = 1$  et  $\alpha < 1$ , on peut montrer [NOLAN \(2015\)](#) que la distribution a pour support  $[\mu, +\infty[$ . Pour  $\mu = 0$ , on obtient un ensemble de distributions qui sont à support non-négatif. Une distribution P $\alpha$ S est donc telle que  $\mathcal{P}\alpha\mathcal{S}(\sigma) = \mathcal{S}(\alpha, \sigma, 0, 1)$  avec  $\alpha < 1$ .

---

Soit  $X_k \sim \mathcal{P}\alpha\mathcal{S}(\sigma_k)$  pour  $k \in \llbracket 1, K \rrbracket$ , les  $X_k$  étant indépendants. On a, par stabilité :

$$X = \sum_k X_k \sim \mathcal{P}\alpha\mathcal{S}(\sigma) \text{ avec } \sigma^\alpha = \sum_k \sigma_k^\alpha. \quad (9.3)$$

En règle générale, nous ne pouvons pas écrire analytiquement la densité de probabilité des distributions  $\alpha$ -stables. Ainsi, pour estimer les paramètres de modèles de mélanges comme (9.3), il faut mettre en oeuvre des méthodes de type MCMC. C'est ce qui a été effectué dans [SIMSEKLI et al. \(2015\)](#) pour les distributions  $\mathcal{S}\alpha\mathcal{S}$ , et pourrait par la suite être étendu au cas de distributions  $\mathcal{P}\alpha\mathcal{S}$ .

### 9.1.3 Distribution de Lévy

Il existe trois cas de figure pour lesquels on a une expression analytique de la densité pour des distributions  $\alpha$ -stables :

- $\alpha = 2$  et  $\beta = 0$  : C'est la loi normale, très populaire notamment en audio [FÉVOTTE et al. \(2009\)](#) ;
- $\alpha = 1$  et  $\beta = 0$  : C'est la loi de Cauchy, utilisée dans [LIUTKUS et al. \(2015\)](#) ;
- $\alpha = 1/2$  et  $\beta = 1$  : C'est la loi de Lévy.

Les deux premiers cas correspondent à des distributions  $\mathcal{S}\alpha\mathcal{S}$  et ne nous intéressent donc pas ici puisqu'on cherche à modéliser des données non-négatives. Par contre, la loi de Lévy présente l'avantage d'être une distribution  $\mathcal{P}\alpha\mathcal{S}$  puisque  $\beta = 1$  et  $\alpha < 1$ . Son paramètre de localisation peut être fixé arbitrairement, et comme on cherche à modéliser des données non-négatives, on choisit  $\mu = 0$ . Par abus de langage, on désignera par loi de Lévy cette loi dans le cas particulier où le paramètre de localisation est nul. Sa densité de probabilité, illustrée sur la figure 9.1 est donnée par :

$$p(x|\sigma) = \begin{cases} \sqrt{\frac{\sigma}{2\pi}} \frac{1}{x^{3/2}} e^{-\frac{\sigma}{2x}} & \text{si } x > 0 \\ 0 & \text{sinon.} \end{cases} \quad (9.4)$$

Outre le fait d'avoir une densité de probabilité que l'on peut écrire analytiquement, la distribution de Lévy possède une queue lourde, comme illustré sur la figure 9.2. Cela signifie que des valeurs éloignées du mode de la distribution restent tout de même probables. Cette propriété se traduit par la robustesse d'un modèle de Lévy par rapport aux valeurs aberrantes : en effet, celles-ci modifient moins l'estimation des paramètres de la loi que pour des distributions à queue non lourde.