

Étude d'un algorithme d'optimisation pour le fading temps-fréquence

Marina KRÉME^{1,2}, Bruno TORRÉSANI¹,

¹Université de Lorraine, CNRS, Inria, Loria, F-54000 Nancy, France

²Aix Marseille Univ, CNRS, I2M, Marseille, France

ama-marina.kreme@inria.fr, bruno.torresani@univ-amu.fr

Résumé – Nous abordons le problème du filtrage des composantes temps-fréquence localisées dans les signaux. Nous le formulons comme un problème de minimisation d'une forme quadratique appropriée, qui implique un terme d'attache aux données sur la transformée de Fourier à court terme en dehors du support de la composante indésirable et un terme de pénalisation de l'énergie à l'intérieur du support. Nous étudions deux méthodes de résolution : une méthode spectrale et une méthode de type quasi-Newton (BFGS). Ces dernières font intervenir des opérateurs appelés *multiplicateurs de Gabor*. Nous exploitons les techniques de projections aléatoires pour approcher ces opérateurs. Nous étudions et comparons la complexité théorique d'un BFGS classique à un BFGS combiné aux projections aléatoires. Nous comparons le temps de calcul de ces deux méthodes sur plusieurs signaux audio. Nous comparons également leur temps de calcul à celui de la méthode spectrale. Les performances de toutes ces approches sont évaluées et comparées sur plusieurs signaux audio.

Abstract – We address the problem of filtering localized time-frequency components in signals. The problem is formulated as a minimization of an appropriate quadratic form, which involves a data fidelity term on the short-time Fourier transform outside the support of the undesirable component and an energy penalty term inside the support. We study two resolution methods: a spectral method and a quasi-Newton-like method (BFGS). The latter involve operators called *Gabor multipliers*. We exploit random projection techniques to approximate these operators. We study and compare the theoretical complexity of a classical BFGS with a BFGS combined with random projections. We compare the computation time of these two methods on several audio signals. We also compare their computation time to that of the spectral method. The performances of all these approaches are evaluated and compared on several audio signals.

1 Introduction

Les signaux dont le contenu fréquentiel varie dans le temps sont généralement traités en utilisant des représentations temps-fréquence (TF) [3], principalement la transformée de Fourier à court terme et sa version discrète, la transformée de Gabor. Ces transformées TF permettent d'effectuer un filtrage non stationnaire, autrement dit, d'extraire les « composantes » non stationnaires des signaux. Le filtrage de ces composantes est souvent effectué en identifiant les domaines d'intérêt et en inversant la transformée après masquage [12, 11]. Quelle que soit la méthode utilisée, les transformées TF résultantes ne sont généralement pas consistantes (toute matrice à coefficients complexes n'est pas la transformée TF d'un signal), sauf si la consistance est imposée comme une contrainte [8].

Dans ce papier, nous nous intéressons à un filtrage temps-fréquence, que nous définissons comme suit : atténuation des composantes localisées dans le plan TF. Ce problème, abordé dans [6], est appelé *fading temps-fréquence* (TFF). Il est résolu par une approche variationnelle, qui minimise une fonction objectif quadratique bien choisie. Dans [6], nous avons proposé une formulation plus générique de TFF et une approche spectrale pour la résolution. A la base, celle-ci est coûteuse, mais la réduction de dimension par des méthodes de projections aléatoires [4] a permis de contrôler le coût calculatoire. Ceci motive les travaux actuels d'un point de vue computationnel. L'idée est

d'intégrer les projections aléatoires dans une méthode quasi-Newton de type BFGS.

2 Éléments d'analyse de Gabor

2.1 Notations et transformée de Gabor

Soit \mathbf{x} un signal numérique de longueur L . On définit $\mathbb{Z}_L = \{0, 1, \dots, L-1\}$. Soient $a, b \in \{1, \dots, L\}$ deux diviseurs de L , désignant les pas d'échantillonnage temporel et fréquentiel. On note $M = L/b$, $N = L/a$. Soient la grille TF définie par $\Lambda = \{(mb, na), m \in \mathbb{Z}_M, n \in \mathbb{Z}_N\}$ et $\mathbf{g} \in \mathbb{R}^L$ une fenêtre d'analyse. L'atome de Gabor \mathbf{g}_{mn} au point TF $(mb, na) \in \Lambda$ est défini par : $\mathbf{g}_{mn}[l] = \mathbf{g}[l - na]e^{2i\pi mbl/L}, \forall l \in \mathbb{Z}_L$. La transformée de Gabor $\mathcal{V}_{\mathbf{g}}\mathbf{x} \in \mathbb{C}^{M \times N}$ de \mathbf{x} est définie par :

$$\mathcal{V}_{\mathbf{g}}\mathbf{x}[m, n] = \langle \mathbf{x}, \mathbf{g}_{mn} \rangle = \sum_{l=0}^{L-1} \mathbf{x}[l] \mathbf{g}[l - na] e^{-2i\pi mbl/L}.$$

Étant donné l'adjoint $\mathcal{V}_{\mathbf{g}}^*$ de $\mathcal{V}_{\mathbf{g}}$, aussi appelé *opérateur de synthèse*, l'*opérateur de repère* $\mathcal{S} = \mathcal{V}_{\mathbf{g}}^* \mathcal{V}_{\mathbf{g}}$ est borné, auto-adjoint et semi-défini positif. Si $\mathcal{S} > 0$, \mathcal{S} est inversible, ce qui permet de reconstruire tout $\mathbf{x} \in \mathbb{C}^L$ à partir des coefficients de Gabor.

Les repères dits de Parseval, pour lesquels $\mathcal{S} = I$, i.e. $\mathcal{V}_{\mathbf{g}}$ est un inverse à gauche de $\mathcal{V}_{\mathbf{g}}$, présentent un intérêt particulier. Pour simplifier, nous nous concentrons sur ce cas ici.

2.2 Multiplicateurs de Gabor

Les *multiplicateurs de Gabor* sont des opérateurs linéaires qui effectuent un filtrage variant dans le temps par multiplication ponctuelle dans le domaine de Gabor.

Définition 2.1. *Le multiplicateur de Gabor associé à (\mathbf{g}, Λ) avec le masque $\mathbf{m} \in \mathbb{C}^{M \times N}$ est défini par $\mathfrak{M}_{\mathbf{m}} = \mathcal{V}_{\mathbf{g}}^* \mathbf{m} \mathcal{V}_{\mathbf{g}}$*

$$\mathfrak{M}_{\mathbf{m}} \mathbf{x} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathbf{m}[m, n] \langle \mathbf{x}, \mathbf{g}_{mn} \rangle \mathbf{g}_{mn} .$$

Les multiplicateurs de Gabor ont été étudiés de manière très approfondie (voir [1, 2] et les références qui y figurent). L'approche spectrale de [6] repose sur la propriété suivante :

Propriété 2.1. *Si \mathbf{m} est à valeur réelle, alors $\mathfrak{M}_{\mathbf{m}}$ est auto-adjoint. Il existe alors une base orthonormée de \mathbb{C}^L formée par les vecteurs propres de $\mathfrak{M}_{\mathbf{m}}$.*

3 Fading temps-fréquence (TFF)

3.1 Formulation

Soient $\mathbf{x}^o \in \mathbb{R}^L$ le signal d'intérêt et $\mathbf{y}^o \in \mathbb{R}^L$ un signal de perturbation. Nous supposons que \mathbf{y}^o est essentiellement localisé dans une région spécifique Ω du plan TF Λ . Nous supposons que la région Ω est symétrique par rapport à l'axe des fréquences nulles puisque les signaux sont à valeurs réelles. Nous désignons par $\bar{\Omega} = \Lambda \setminus \Omega$ le complémentaire de Ω . Le signal observé \mathbf{z} est de la forme :

$$\mathbf{z} = \mathbf{x}^o + \mathbf{y}^o . \quad (1)$$

Définition 3.1 (Fading temps-fréquence). *Soit $\mathbf{z} \in \mathbb{R}^L$ le modèle défini en (1). Soient $\mathcal{V}_{\mathbf{g}}$ la transformation de Gabor et Ω une région du plan temps-fréquence. Le problème du fading temps-fréquence s'écrit comme suit : pour $\lambda > 0$,*

$$\mathbf{x}_{\lambda} = \arg \min_{\mathbf{x} \in \mathbb{R}^L} f_{\lambda}(\mathbf{x}) , \quad (2)$$

$$f_{\lambda}(\mathbf{x}) = \|\mathcal{V}_{\mathbf{g}} \mathbf{x} - \mathcal{V}_{\mathbf{g}} \mathbf{z}\|_{\Omega}^2 + \lambda \|\mathcal{V}_{\mathbf{g}} \mathbf{x}\|_{\bar{\Omega}}^2 . \quad (3)$$

Le problème (2) revient à trouver une estimation \mathbf{x}_{λ} de \mathbf{x}^o qui correspond le mieux à \mathbf{z} en dehors de Ω et dont l'énergie à l'intérieur de Ω est contrôlée. La fonction objectif f_{λ} dépend du paramètre de régularisation $\lambda > 0$. Le premier terme de cette fonction est un terme d'attache aux données qui fait correspondre la transformée d'un signal estimé à celui de l'observation en dehors de Ω . Le second terme contrôle l'énergie à garder dans la région Ω , et le paramètre de régularisation contrôle le compromis entre les deux termes.

Pour simplifier, nous désignons par \mathfrak{M}_{Ω} et $\mathfrak{M}_{\bar{\Omega}}$ les multiplicateurs associés aux fonctions indicatrices $\mathbb{1}_{\Omega}$ et $\mathbb{1}_{\bar{\Omega}}$.

3.2 Solution analytique, approche spectrale

La fonction objectif définie dans (2) est une forme quadratique. Sa minimisation aboutit donc à un système linéaire. Ceci

conduit à la solution suivante :

$$\mathbf{x}_{\lambda} = \mathbf{z} - \left(\mathbf{I} - (1 - \lambda) \mathfrak{M}_{\Omega} \right)^{-1} \lambda \mathfrak{M}_{\Omega} \mathbf{z} , \quad (4)$$

l'inversibilité de $\mathbf{I} - (1 - \lambda) \mathfrak{M}_{\Omega}$ étant assurée par le fait que le spectre de \mathfrak{M}_{Ω} est inclus dans $[0, 1]$ (voir [6]).

Comme mentionné dans la Propriété 2.1, le masque étant binaire, et donc réel, le multiplicateur de Gabor \mathfrak{M}_{Ω} peut être diagonalisé et donc factorisé sous la forme $\mathfrak{M}_{\Omega} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{U}^*$ où \mathbf{U} est unitaire et $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\sigma})$ est diagonale. Ici $\boldsymbol{\sigma} \in \mathbb{R}^L$ est le vecteur des valeurs propres, qui sont triées par ordre décroissant $1 \geq \sigma[1] \geq \dots \geq \sigma[L] \geq 0$.

Nous pouvons alors écrire

$$\mathbf{x}_{\lambda} = \mathbf{z} - \mathbf{U} \boldsymbol{\Gamma} \mathbf{U}^* \mathbf{z} \quad (5)$$

où $\boldsymbol{\Gamma} = \text{diag}(\boldsymbol{\gamma})$, avec $\boldsymbol{\gamma} = (\gamma[1], \dots, \gamma[L])$ et

$$\gamma[i] = \frac{\lambda \sigma[i]}{1 - (1 - \lambda) \sigma[i]} .$$

On montre que la suite $\boldsymbol{\gamma} = (\gamma[1], \dots, \gamma[L])$ est décroissante et incluse dans $[0, 1]$, que nous utiliserons pour la troncature.

La solution approchée lorsqu'on tronque aux K plus grandes valeurs propres ($K < L$) est donnée par :

$$\mathbf{x}_{\lambda}^{(K)} = \mathbf{z} - \mathbf{U}^{(K)} \boldsymbol{\Gamma}^{(K)} (\mathbf{U}^{(K)})^* \mathbf{z} \quad (6)$$

où $\mathbf{U}^{(K)}$ est la sous-matrice $L \times K$ constituée des premières K colonnes de \mathbf{U} , et $\boldsymbol{\Gamma}^{(K)}$ est la sous-matrice diagonale constituée des K premières valeurs de $\boldsymbol{\gamma}$. La solution de TFF obtenue par l'approche spectrale (voir (6)) sera notée TFF-s. L'erreur de troncature peut être bornée comme dans [5] par :

$$\|\mathbf{x}_{\lambda} - \mathbf{x}_{\lambda}^{(K)}\|_2 \leq \gamma[K+1] \|\mathbf{z}\|_2 . \quad (7)$$

4 TFF et projections aléatoires

4.1 Projections aléatoires

La solution explicite obtenue en (5) s'appuie fortement sur les décompositions en valeurs propres (EVD) des multiplicateurs de Gabor. Cependant, la plupart des algorithmes EVD ont une complexité de $\mathcal{O}(L^3)$, qui devient prohibitive en grande dimension. Cette complexité peut être réduite lorsque le rang numérique K est bien inférieur à la dimension. Si K est connu, une telle approximation de rang faible est calculée en $\mathcal{O}(KL^2)$ par des algorithmes bien connus basés sur l'algorithme de Lanczos tels que la méthode *Implicitly Restarted Arnoldi* [7].

Dans [4], une méthode basée sur des projections aléatoires est proposée afin d'accélérer le calcul de la décomposition en valeurs singulières (SVD) : soit \mathbf{A} une matrice de taille $M \times N$ que l'on souhaite approcher comme produit de deux matrices \mathbf{B} et \mathbf{C} de tailles respectives $M \times K$ et $K \times N$ avec $K \ll N$ et $\mathbf{A} \approx \mathbf{B}\mathbf{C}$. L'approche de [4] repose sur deux étapes :

1. **étape 1** : trouver une base approchée de l'espace image de \mathbf{A} , c'est-à-dire une matrice \mathbf{Q} à K colonnes orthonormées telle que $\mathbf{A} \approx \mathbf{Q}\mathbf{Q}^* \mathbf{A}$.

2. **étape 2** : étant donnée une telle matrice Q , si $K \ll N$ l'utiliser pour calculer la décomposition en valeurs propres. Concrètement, on pose $B = Q^*A$ et on calcule la SVD de $B = \tilde{U}\Sigma V^*$. Pour obtenir la SVD de A , on pose $U = Q\tilde{U}$ et on obtient $A \approx U\Sigma V^*$

Dans [5, 6], nous avons exploité cette méthode dans le cas d'une EVD. Nous la combinons avec une méthode quasi-Newton du type BFGS, la résultante sera notée TFF-qnr.

4.2 TFF-quasi Newton randomisée (TFF-qnr)

La fonction objectif f_λ peut s'écrire :

$$f_\lambda(\mathbf{x}) = \|\mathbf{x} - \mathbf{z}\|_2^2 - \langle \mathfrak{M}_\Omega(\mathbf{x} - \mathbf{z}), \mathbf{x} - \mathbf{z} \rangle + \lambda \langle \mathfrak{M}'\mathbf{x}, \mathbf{x} \rangle.$$

Si $\mathbf{x} \in \mathbb{R}^L$, alors la matrice associée à \mathfrak{M}_Ω est de taille L^2 . \mathfrak{M}_Ω est Hermitienne, et des techniques de projections aléatoires fournissent une matrice Q de taille LK (avec $K \ll L$) telle que

$$\mathfrak{M}_\Omega \simeq QQ^*\mathfrak{M}_\Omega QQ^* = Q\mathfrak{M}'Q^*,$$

où l'on a posé $\mathfrak{M}' = Q^*\mathfrak{M}_\Omega Q$. La fonction objectif approchée est alors définie par :

$$f^{\text{app}}(\mathbf{x}) = \|\mathbf{x} - \mathbf{z}\|_2^2 - \langle \mathfrak{M}'Q^*(\mathbf{x} - \mathbf{z}), Q^*(\mathbf{x} - \mathbf{z}) \rangle + \lambda \langle \mathfrak{M}'Q^*\mathbf{x}, Q^*\mathbf{x} \rangle,$$

et son gradient s'écrit

$$\nabla_{\mathbf{x}} f^{\text{app}}(\mathbf{x}) = 2 \left[\mathbf{x} - \mathbf{z} + (\lambda - 1)Q\mathfrak{M}'Q^*\mathbf{x} + Q\mathfrak{M}'Q^*\mathbf{z} \right].$$

Complexité de calcul de $\nabla_{\mathbf{x}} f(\mathbf{x})$. Elle dépend de celle de $\mathfrak{M}_\Omega \mathbf{x}$. Le calcul d'une transformée de Gabor repose sur N transformées de Fourier rapides (*Fast Fourier Transform* ou FFT en anglais) de taille M , soit $\mathcal{O}(MN \log_2 M)$. De même, la complexité de l'inverse d'une transformée de Gabor vaut $\mathcal{O}(MN \log_2 M)$. La complexité de la multiplication terme à terme par le masque vaut $\mathcal{O}(MN)$. La complexité totale est donc de $\mathcal{O}(2MN \log_2 M + MN)$ opérations.

Complexité de calcul de $\nabla_{\mathbf{x}} f^{\text{app}}(\mathbf{x})$. Q est de taille LK avec $K \ll L$. Si $\mathbf{x} \in \mathbb{R}^L$, alors $Q^*\mathbf{x} \in \mathbb{R}^K$, ce qui fait un coût calculatoire de $\mathcal{O}(KL)$. Aussi, $\mathfrak{M}'Q^*\mathbf{x} \in \mathbb{R}^K$ et $Q\mathfrak{M}'Q^*\mathbf{x} \in \mathbb{R}^L$, ce qui conduit à des complexités respectives de $\mathcal{O}(K^2)$ et $\mathcal{O}(KL)$, d'où une complexité totale $\mathcal{O}(2KL + K^2)$.

Si $MN > L$, alors $2MN \log_2 M + MN > 2L \log_2 M + L$. Et si $K \ll L$, alors $2KL + K^2$ sera très petit. Et donc, en fonction du nombre d'itérations qui sera fixé, $2KL + K^2$ sera bien plus petit que $2L \log_2 M + L$. Par conséquent, nous pensons qu'avec l'optimisation numérique, TFF-qnr convergera rapidement, car en grande dimension la méthode spectrale est coûteuse même avec des projections aléatoires.

5 Résultats numériques

Les simulations ont été effectuées sous MATLAB et sur un système macOS équipé d'un processeur Apple M1.

Les signaux audio considérés ont été échantillonnés à 8 kHz, ont une longueur $L = 16384$ échantillons et sont disponibles

en ligne¹. Le signal observé $\mathbf{z} = \mathbf{x}^o + \mathbf{y}^o$ est construit comme étant la somme d'un signal d'intérêt \mathbf{x}^o et d'un signal de perturbation \mathbf{y}^o . Nous considérons 9 mélanges de signaux obtenus à partir de toutes les paires de signaux d'intérêts (bruit du moteur d'une voiture, d'un avion et d'un train) et de signaux de perturbations (bips, chant d'oiseau, claquements de doigts). Les signaux d'intérêts ont un spectrogramme large bande, les signaux de perturbation sont bien localisés en temps-fréquence.

Les paramètres de la transformée de Gabor sont : $b = 32$ et $a = 64$. Nous avons utilisé une fenêtre gaussienne de longueur 256. On obtient une matrice TF de taille $MN = 513 \times 256$. La génération de la région Ω est faite comme dans [6]. Le spectrogramme et le masque de l'exemple oiseau + voiture est présenté sur la Figure 1.

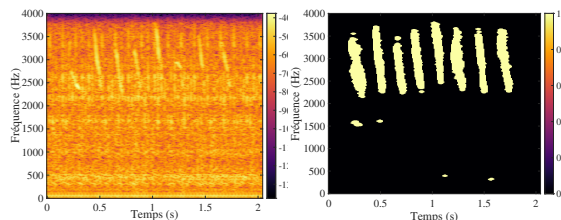


FIGURE 1 – bruit d'un moteur + chant d'oiseau (gauche) et masque binaire (droite).

Pour cet exemple, TFF-qnr et TFF-qnr sont mis en œuvre en utilisant la fonction *minFunc* [9]. Nous fixons la tolérance à 10^{-5} et le nombre d'itérations maximal à 3000. Tous les autres paramètres par défaut de cette fonction sont conservés. L'initialisation de la solution recherchée est faite aléatoirement. On fait varier λ de 10^{-11} à 10^{10} . La valeur optimale de λ (pour toutes les méthodes) est obtenue en maximisant le rapport source à distorsion (SDR) [10] exprimé en décibel. Nous obtenons $\lambda = 10^{-3}$. Le temps de calcul est d'environ 1551 secondes pour TFF-qnr contre environ 1102 secondes pour TFF-qnr. Ce qui semble confirmer l'étude de la complexité théorique faite plus haut.

Pour chaque paire de données, nous évaluons à la fois les performances de reconstruction (Tableau 1) et les temps de calcul (Figure 2). Les résultats présentés dans le Tableau 1 correspondent à une moyenne sur 100 exécutions. Les caractères gras sont utilisés pour souligner les meilleures valeurs parmi TFF-qnr, TFF-qnr et TFF-s. Les bonnes performances sont souvent réalisées par TFF-s. Les performances de TFF-qnr et TFF-qnr sont quasiment similaires et restent relativement proche de TFF-s. Le manque de précision de TFF-qnr par rapport à TFF-s peut s'expliquer par l'accumulation d'erreurs liées à l'approximation par projection à chaque itération.

Les temps de calcul pour TFF-qnr et TFF-s sont représentés en Figure 2. Chaque barre correspond à un mélange. On voit que TFF-s est généralement plus rapide que TFF-qnr sauf pour le mélange 8. Cela pourrait s'expliquer par le fait que la décomposition spectrale (étape la plus coûteuse) est recalculée pour chaque valeur de λ ; ce qui n'est pas le cas pour TFF-s.

1. <https://gitlab.lis-lab.fr/skmad-suite/tff2020>

	voiture	train	avion	
TFF-qn	22.32	21.01	26.87	Bips
TFF-qnr	22.46	21.10	26.02	
TFF-s	24.40	21.10	27.22	
TFF-qn	15.65	12.88	15.73	Claquement
TFF-qnr	15.95	13.02	15.88	
TFF-s	14.20	12.73	16.12	
TFF-qn	16.92	23.38	23.54	Oiseau
TFF-qnr	16.80	22.22	23.50	
TFF-s	17.22	23.45	24.50	

TABLE 1 – SDR en décibel (dB) pour chaque méthode et chaque configuration.

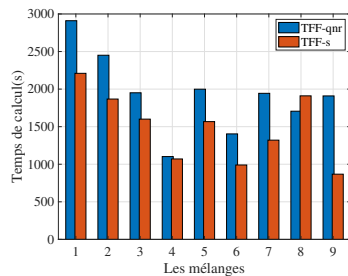


FIGURE 2 – Temps de calcul de TFF-s par rapport à TFF-qnr avec une barre par mélange

Un exemple de reconstruction est présenté sur la Figure 3 (paire oiseau + voiture). Visuellement, on peut observer qu'il reste plus de trace de l'oiseau (bande de fréquence autour de 1.5 secondes) sur les spectrogrammes obtenus par TFF-qn et TFF-qnr comparé à celui fourni par TFF-s.

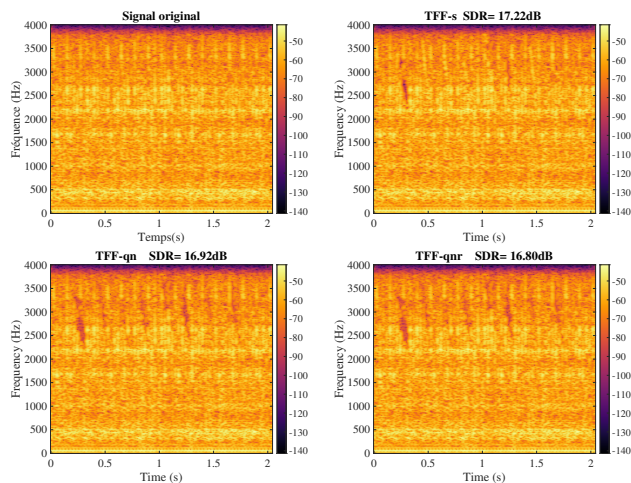


FIGURE 3 – oiseau+voiture : spectrogrammes du signal cible et de la reconstruction par chaque méthode

6 Conclusion

Nous avons abordé le problème de l'estimation d'un signal cible, sans hypothèse sur son contenu, lorsqu'il est perturbé par un signal additif qui est bien localisé dans une région Ω

du plan TF. Nous avons étudié la complexité théorique d'une méthode de type quasi-Newton (TFF-qn) versus la combinaison de TFF-qn avec des projection aléatoires (TFF-qnr). Les résultats numériques montrent que TFF-qnr est significativement plus rapide que TFF-qn mais toujours plus lent que l'algorithme TFF présenté dans [6].

Références

- [1] M. Dörfler. *Gabor analysis for a class of signals called music*. PhD thesis, University of Vienna, Austria, 2002.
- [2] H. G. Feichtinger and K. Nowak. A first survey of Gabor multipliers. In H. G. Feichtinger and T. Strohmer, editors, *Advances in Gabor Analysis*, pages 99–128. Birkhäuser, 2002.
- [3] K. Gröchening. *Foundations of Time-Frequency Analysis*. Birkhäuser, Boston (MA), 2011.
- [4] N. Halko, P.-G. Martinsson, and J. A. Tropp. Finding structure with randomness : Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM J. Math. Anal.*, 53(2) :217–288, 2011.
- [5] M. A. Krémé, V. Emiya, C. Chaux, and B. Torrèsani. Filtering out time-frequency areas using Gabor multipliers. In *Proc. Int. Conf. Acoust. Speech Signal Process.*, pages 5790–5794, 2020.
- [6] M. A. Krémé, V. Emiya, C. Chaux, and B. Torrèsani. Time-frequency fading algorithms based on Gabor multipliers. *IEEE Journal of Selected Topics in Signal Processing*, 15(1) :65–77, jan 2021.
- [7] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of the National Bureau of Standards*, 45(4) :255, oct 1950.
- [8] J. L. Roux and E. Vincent. Consistent Wiener filtering for audio source separation. *IEEE Signal Processing Letters*, 20(3) :217–220, Mar. 2013.
- [9] M. Schmidt. minFunc : unconstrained differentiable multivariate optimization in Matlab. www.cs.ubc.ca/~schmidtm/Software/minFunc.html, 2005.
- [10] E. Vincent, R. Gribonval, and C. Févotte. Performance measurement in blind audio source separation. *IEEE Transactions on Audio, Speech and Language Processing*, 14(4) :1462–1469, July 2006.
- [11] D. Wang. Time-frequency masking for speech separation and its potential for hearing aid design. *Trends in Amplification*, 12(4) :332–353, oct 2008.
- [12] O. Yilmaz and S. Rickard. Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on Signal Processing*, 52(7) :1830–1847, jul 2004.