



HAL
open science

Approximation parcimonieuse avec dictionnaire continu pour l'analyse spectrale de données irrégulièrement échantillonnées

Hervé Carfantan, Mégane Boudineau

► **To cite this version:**

Hervé Carfantan, Mégane Boudineau. Approximation parcimonieuse avec dictionnaire continu pour l'analyse spectrale de données irrégulièrement échantillonnées. GRETSI 2017, GRETSI, Sep 2017, Juan les Pins, France, France. hal-01635129

HAL Id: hal-01635129

<https://hal.science/hal-01635129>

Submitted on 14 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Approximation parcimonieuse avec dictionnaire continu pour l'analyse spectrale de données irrégulièrement échantillonnées

Mégane BOUDINEAU¹, Hervé CARFANTAN¹

¹Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse / CNRS
14 avenue Édouard Belin, 31400 Toulouse, France

Megane.Boudineau@irap.omp.eu, Herve.Carfantan@irap.omp.eu

Résumé – La recherche de fréquences pures dans un signal peut être formalisée comme un problème d'*approximation parcimonieuse*, en discrétisant sur une grille les fréquences recherchées et en formant un dictionnaire à partir des signaux élémentaires associés, mais cela pose des problèmes lorsque les fréquences sont hors de cette grille. Récemment, de nombreux travaux ont cherché à étendre les méthodes d'approximation parcimonieuse classiques (algorithmes gloutons, pénalisation ℓ_1 , modélisation Bernoulli-Gaussienne) à un dictionnaire *continu*, permettant d'éviter l'utilisation d'une telle grille. Cet article vise à comparer des approches d'approximation parcimonieuse à dictionnaires discrets et continus pour le problème d'analyse spectrale de signaux irrégulièrement échantillonnés.

Abstract – Frequencies localisation in a signal can be written as a *sparse approximation* problem: the frequencies are discretised on a grid and a dictionary is built with associated elementary signals. However, issues arise when frequencies are off the grid. Recently, many studies have been conducted in order to extend classical sparse approximation methods (greedy algorithm, ℓ_1 -penalisation, Bernoulli-Gaussian modeling) to a *continuous* dictionary, avoiding the use of such a grid. This article aims at comparing sparse approximation approaches with discrete and continuous dictionaries for spectral analysis of unevenly sampled data.

1 Introduction

Un problème d'*approximation parcimonieuse* consiste à approcher un signal \mathbf{y} comme une combinaison linéaire de K atomes \mathbf{h}_k , formant un *dictionnaire* \mathbf{H} , dont les coefficients x_k sont *a priori* parcimonieux : la plupart des x_k sont nuls. L'estimation de \mathbf{x} peut être effectuée en minimisant sa pseudo-norme ℓ_0 (le nombre de x_k non nuls) avec $\mathbf{y} \approx \mathbf{H}\mathbf{x}$. Ce problème étant reconnu comme NP-difficile, différentes alternatives ont été proposées. Les *algorithmes gloutons*, popularisés par [1], consistent à sélectionner itérativement les atomes \mathbf{h}_k les plus corrélés au résidu. La *relaxation* de la pénalisation ℓ_0 par une pénalisation ℓ_1 convexe permet de garantir une solution parcimonieuse, unique et calculable. Enfin, l'introduction d'un *modèle a priori Bernoulli-Gaussien* [2] permet de modéliser la parcimonie dans un cadre probabiliste et d'estimer \mathbf{x} à partir d'échantillons générés par des algorithmes MCMC.

En parallèle, un problème classique en sciences de l'univers est la recherche de fréquences ν_ℓ et d'amplitudes x_ℓ d'un nombre inconnu L d'oscillations sinusoïdales responsables de la réalisation d'un signal $\mathbf{y} = \sum_{\ell=-L}^L \mathbf{h}(\nu_\ell)x_\ell + \epsilon$, avec ϵ un bruit additif et $\mathbf{h}(\nu) = e^{j2\pi\nu t}$, non linéaire en ν . Les valeurs y_n du signal peuvent être acquises *irrégulièrement* et présenter des données manquantes de façon *périodique*, cas courant dans les observations astronomiques soumises par exemple aux alternances jour/nuit. Si l'irrégularité des instants d'échantillonnage t_n annule la limite de Nyquist, les trous périodiques dans l'observation introduisent quant à eux de forts lobes secondaires dans la fenêtre spectrale $W(\nu) = \sum_n e^{j2\pi\nu t_n}$. Ces derniers

rendent la recherche de fréquences particulièrement difficile, notamment lorsqu'elles sont multiples : comme illustré en Figure 1, les fréquences du signal ne s'identifient plus avec les maxima du spectre de Fourier.

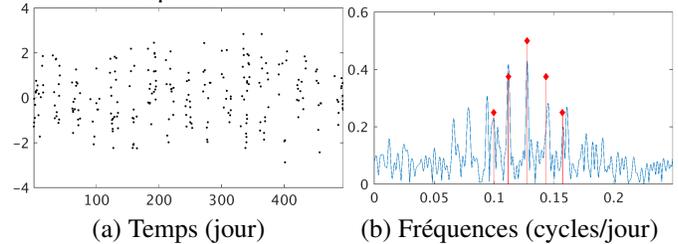


FIG. 1 – (a) Signal irrégulièrement échantillonné. (b) Spectre de Fourier (-) et les 5 fréquences réelles (\diamond).

Ce problème d'analyse spectrale peut se reformuler en terme d'approximation parcimonieuse : en discrétisant les fréquences ν sur une grille $\Omega_\nu = [\nu_{-K}, \dots, \nu_K]$, avec $K \gg L$, on réécrit le signal $\mathbf{y} = \sum_{k=-K}^K x_k \mathbf{h}_k + \epsilon = \mathbf{H}\mathbf{x} + \epsilon$, où les atomes \mathbf{h}_k s'écrivent $\mathbf{h}(\nu_k)$. Les méthodes d'approximation parcimonieuse sur dictionnaire discret ont donc été largement utilisées pour résoudre ce problème. Cependant, la discrétisation des fréquences entraîne nécessairement une perte de précision, puisque les fréquences estimées appartiennent à la grille Ω_ν . Pour limiter ce problème, une discrétisation fine de la grille peut être envisagée, mais cela augmente la cohérence du dictionnaire et la complexité du problème d'approximation parcimonieuse. Un traitement *a posteriori* à partir d'estimation sur une grille relâchée peut alors être envisagé, comme le pro-

posent [3] dans la méthode SparSpec et [4] avec une minimisation locale autour des fréquences estimées.

Récemment, des alternatives ont été proposées pour exploiter un *dictionnaire continu*. Un remplacement de la norme ℓ_1 par la *norme atomique* et une reformulation adéquate du problème d'optimisation permet à [5] de le résoudre sur un véritable dictionnaire continu. Une alternative, proposée par [6] et adaptée par [7], consiste à utiliser une interpolation du dictionnaire permettant une estimation continue des paramètres ν au prix d'une erreur d'approximation. Enfin, une extension du modèle Bernoulli-Gaussien permettant d'estimer les fréquences hors grille a été proposée par [2, 8].

L'objet de cet article est la comparaison des méthodes d'approximation parcimonieuse sur dictionnaire discret avec post-traitement et sur dictionnaire continu pour de l'analyse spectrale en échantillonnage irrégulier. En particulier, la méthode de post-traitement SparSpec [3] est comparée avec les méthodes à base d'interpolation de dictionnaire (PoIBOMP et CCBP) développées par [7], avec le modèle Bernoulli-Gaussien étendu (BGE) [8] et avec LOOMP, un algorithme gloutin intégrant une étape d'optimisation locale. L'utilisation de la norme atomique proposée par [5] étant pour le moment limitée par l'irrégularité de nos données, elle ne sera pas étudiée ici. Nous commençons donc par présenter succinctement ces diverses méthodes, puis nous présentons les données sur lesquelles les tests sont effectués et enfin nous analysons les résultats obtenus.

2 Présentation des différentes méthodes

Nous définissons les notations suivantes : Δ est l'écart entre deux points (ν_k, ν_{k+1}) de la grille Ω_ν , et \mathcal{I}_k est l'intervalle $[\nu_k - \Delta/2; \nu_k + \Delta/2]$ entourant $\nu_k \in \Omega_\nu$. Comme les données \mathbf{y} sont à valeurs réelles, les $x_k \in \mathbb{C}$ associés à la modélisation en exponentielles complexes $\mathbf{H}\mathbf{x}$ sont à symétrie hermitienne : $x_{-k} = x_k^*$. Une modélisation équivalente, spécifique à l'analyse spectrale, est utilisée ici : $\mathbf{H}\mathbf{x} \Leftrightarrow \mathbf{G}\mathbf{s}$, où le dictionnaire \mathbf{G} vient de la discrétisation de $\mathbf{g}(\nu) = [\cos(2\pi\nu\mathbf{t}), \sin(2\pi\nu\mathbf{t})]$ sur la grille des fréquences positives, notée $\Omega_{\nu+}$, et $\mathbf{s} = [\mathbf{s}_k]_{k \in \mathbb{N}}$ avec $\mathbf{s}_k \in \mathbb{R}^2$. Cette modélisation fait appel à la *parcimonie structurée*, dans le sens où les amplitudes sont généralement nulles ou non simultanément. Elle sera utilisée par les méthodes BGE et LOOMP. Par soucis de concision, le cas $k = 0$ ne sera pas traité ici.

Nous présentons ici les grandes lignes des méthodes d'approximation parcimonieuse faisant appel à du post-traitement et à un dictionnaire continu. Pour plus de détails sur ces méthodes, le lecteur est invité à se rapprocher des articles cités.

2.1 Méthode SparSpec

La méthode SparSpec [3] résout, pour l'analyse spectrale en échantillonnage irrégulier, le problème de minimisation avec une pénalisation ℓ_1 : $\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$, en gardant les mêmes notations que précédemment. Les fréquences solutions de ce problème appartiennent donc à la grille Ω_ν ; si une

fréquence $\nu \notin \Omega_\nu$ est présente dans le signal, la solution l'encadrera généralement avec un doublon $(\nu_k, \nu_{k+1}) \in \Omega_\nu^2$. De plus, la norme ℓ_1 apporte un biais à l'estimation des amplitudes \hat{x}_k . Il est proposé dans [3] une réestimation *a posteriori* des amplitudes et un remplacement des doublons par leur barycentre ; cette dernière estimation n'appartient donc pas à la grille.

2.2 Algorithme LOOMP

Nous présentons ici une extension de l'algorithme classique d'*Orthogonal Matching Pursuit* (OMP), nommée LOOMP (Local OMP), inspirée de l'utilisation de descente de gradient locale envisagée par [4]. Le modèle (\mathbf{G}, \mathbf{s}) de parcimonie structurée est employé. Considérons l'algorithme suivant :

LOOMP Début : résidu $\mathbf{r} = \mathbf{y}$, atomes sélectionnés $\hat{\mathbf{G}} = \emptyset$. Tant que le critère d'arrêt n'est pas atteint, on répète :

1. Calcul d'une grille locale $\hat{\Omega}_\nu = [\hat{\nu}_1, \dots, \hat{\nu}_K]$:
pour tout k , $(\hat{\nu}_k, \hat{\mathbf{s}}_k) = \arg \min_{\nu \in \mathcal{I}_k, \mathbf{s}} \|\mathbf{r} - \mathbf{g}(\nu)\mathbf{s}\|^2$.
2. Sélection : $\ell = \arg \min_k \|\mathbf{r} - \mathbf{g}(\hat{\nu}_k)\hat{\mathbf{s}}_k\|^2$.
3. Mise à jour : $\hat{\mathbf{G}} = [\hat{\mathbf{G}}, \mathbf{g}(\hat{\nu}_\ell)]$, $\mathbf{r} = \mathbf{y} - \hat{\mathbf{G}}\hat{\mathbf{s}}$, où $\hat{\mathbf{s}} = (\hat{\mathbf{G}}^T \hat{\mathbf{G}})^{-1} \hat{\mathbf{G}}^T \mathbf{y}$.

Si les 2 dernières étapes correspondent à un OMP classique, la première consiste à calculer une grille locale $\hat{\Omega}_\nu$ constituée des fréquences qui minimisent le critère dans chaque intervalle \mathcal{I}_k . Le critère en fonction de ν n'étant pas convexe, les fréquences estimées correspondent potentiellement à un minimum local.

2.3 Modèle Bernoulli-Gaussien étendu

Afin de traiter l'*a priori* de parcimonie de façon probabiliste, une séquence de Bernoulli $\mathbf{q} = (q_k)_k$ est introduite, telle que $q_k = 0$ implique $\mathbf{s}_k = 0$. On peut alors donner une loi *a priori* Bernoulli-Gaussienne aux couples (q_k, \mathbf{s}_k) :

$$q_k \sim \mathcal{B}(\lambda) \text{ et } \begin{cases} \mathbf{s}_k | q_k = 1 & \sim \mathcal{N}(0, \sigma_s^2 \mathbf{I}_2), \\ \mathbf{s}_k | q_k = 0 & \sim \delta(\mathbf{s}_k). \end{cases} \quad (1)$$

où $\lambda = \Pr(q_k = 1)$ règle le degré de parcimonie, $\mathcal{N}(0, \sigma_s^2 \mathbf{I}_2)$ est une loi gaussienne centrée de variance σ_s^2 et $\delta(\mathbf{s}_k)$ la distribution de Dirac en $\mathbf{0}$. L'utilisation de la séquence de Bernoulli \mathbf{q} facilite l'usage de la modélisation avec parcimonie structurée (\mathbf{G}, \mathbf{s}) , puisque $q_k = 0$ implique alors que les deux amplitudes associées soient nulles.

Afin de pouvoir estimer des fréquences hors grille, [2] a proposé de relâcher les fréquences ν_k autour du point de la grille, si bien que $\nu_k \in \mathcal{I}_k$. On associe à ν_k un *a priori* uniforme sur \mathcal{I}_k . Le modèle devient $\sum_k \mathbf{g}(\nu_k) \mathbf{s}_k$ et les paramètres à estimer sont donc $(q_k, \mathbf{s}_k, \nu_k)_k$. Notons que $q_k = 1$ correspond maintenant à la présence d'une fréquence dans l'intervalle \mathcal{I}_k .

En supposant un bruit ϵ gaussien centré indépendant et identiquement distribué de variance σ_ϵ^2 et l'indépendance *a priori* entre les paramètres $(q, \mathbf{s}, \nu)_i$ et $(q, \mathbf{s}, \nu)_k$ pour tout (i, k) , la loi de probabilité *a posteriori* $p(\mathbf{q}, \mathbf{s}, \boldsymbol{\nu} | \mathbf{y})$ peut alors être échantillonnée grâce à un algorithme de type Gibbs hybride. En effet, la loi *a posteriori* conditionnelle en (q_k, \mathbf{s}_k) est toujours

Bernoulli-Gaussienne et ces derniers s'échantillonnent directement, alors qu'une étape de Métropolis-Hastings additionnelle est nécessaire pour ν_k à cause de la non-linéarité du modèle en ce paramètre. Afin d'éviter l'écueil de rester bloqué dans des modes locaux constaté par [2], nous avons proposé de marginaliser la loi *a posteriori* suivant les amplitudes \mathbf{s} et ainsi d'accélérer la convergence de l'échantillonnage stochastique [8].

Le coût de calcul induit par cette marginalisation est atténué grâce à la mise en œuvre de factorisations et de mises à jour de Cholesky. Le *Partially Collapsed Hybrid-Gibbs sampler* (PCGS) obtenu peut être utilisé de façon non-supervisée pour échantillonner la loi $p(\mathbf{q}, \boldsymbol{\nu}, \boldsymbol{\theta} | \mathbf{y})$, où $\boldsymbol{\theta} = [\sigma_s^2, \sigma_e^2, \lambda]$ regroupe les hyperparamètres, qui n'ont plus besoin d'être réglés en amont.

Une fois la convergence de l'algorithme atteinte, l'espérance *a posteriori* des paramètres $(\mathbf{q}, \boldsymbol{\nu}, \boldsymbol{\theta})$ peut être estimée en calculant la moyenne empirique des échantillons. Un paramètre de Bernoulli estimé $\hat{q}_k \in [0; 1]$ peut s'interpréter comme la probabilité d'existence d'une fréquence ν_k dans l'intervalle \mathcal{I}_k . Les échantillons de cette dernière, via leur écart-type, fournissent aussi une barre d'erreur sur son estimation.

2.4 Interpolation de dictionnaire

Le principe de l'interpolation de dictionnaire présenté par [6] est le suivant : pour $\nu \in \mathcal{I}_k$, i.e. $\nu = \nu_k + \delta$ avec $|\delta| < \Delta/2$, la fonction $\mathbf{h}(\nu)$ est approchée par une combinaison linéaire de M nouveaux atomes $\tilde{\mathbf{h}}_{k,m}$ dépendant uniquement de ν_k : $x \mathbf{h}(\nu_k + \delta) \approx \sum_m \tilde{x}_{k,m}(x, \delta) \tilde{\mathbf{h}}_{k,m} = \tilde{\mathbf{h}}_k \tilde{\mathbf{x}}_k(x, \delta)$. Deux types d'interpolation ont été proposés par [6] : développements de Taylor et interpolation polaire, caractérisés par une erreur d'approximation plus faible que celle commise avec un dictionnaire discret. L'identifiabilité $\tilde{\mathbf{x}}_k = f(x, \delta) \Leftrightarrow (x, \delta) = f^{-1}(\tilde{\mathbf{x}}_k)$ est primordiale : connaître $\tilde{\mathbf{x}}_k$ permet alors de calculer (x, δ) sans ambiguïté et de retrouver la vraie fréquence $\nu = \nu_k + \delta$ du modèle.

Le problème d'approximation parcimonieuse est alors reformulé sur un nouveau dictionnaire $\tilde{\mathbf{H}}$, constitué des atomes $\tilde{\mathbf{h}}_k$: on cherche $\tilde{\mathbf{x}}$ tels que $\mathbf{y} \approx \tilde{\mathbf{H}}\tilde{\mathbf{x}}$. Les méthodes usuelles (algorithmes gloutons ou pénalisation ℓ_1 comme le propose [6] avec le *Continuous Basis Pursuit* (CBP)) peuvent alors être utilisées, à condition de rajouter des contraintes aux problèmes d'optimisation afin de garantir l'identifiabilité $(x, \delta) = f^{-1}(\tilde{\mathbf{x}}_k)$: en particulier, le δ associé à $\tilde{\mathbf{x}}_k$ doit respecter $|\delta| < \Delta/2$. La convexité du problème, fragilisée par l'ajout de ces contraintes, peut être généralement garantie dans le cas où les amplitudes du modèle sont réelles et positives.

Comme notre problème d'analyse spectrale présente des amplitudes complexes, nous testerons ici l'extension aux complexes du CBP de [6] proposée par [7], nommé *Complex CBP* (CCBP), qui utilise l'interpolation polaire, et dont le principe est de remplacer un atome du dictionnaire augmenté $\tilde{\mathbf{h}}_k$ par ses alias positif et négatif, partie réelle et partie imaginaire. Par ailleurs, une autre alternative à l'algorithme OMP sur dictionnaire continu \mathbf{y} est proposée : *Polar Interpolate Band-exclusive*

(OMP (PoIBOMP)), qui effectue une interpolation polaire après sélection de la fréquence $\nu^{(i)} = \nu_k \in \Omega_\nu$, afin de trouver la meilleure fréquence dans \mathcal{I}_k . Notons que notre signal \mathbf{y} étant réel, nous avons légèrement modifié CCBP et PoIBOMP afin de garantir la symétrie hermitienne des amplitudes estimées.

3 Tests numériques

Nous testons les différentes méthodes présentées dans la section précédente en les appliquant aux données de [8] : 249 échantillons temporels irréguliers d'une somme de 5 sinusoïdes de fréquences comprises entre 0.1 et 0.16 cycles par jour (c/d), sur une durée de $T = 500$ jours (voir Figure 1). Les fréquences sont recherchées entre 0 et 0.25 c/d. Le signal \mathbf{y} est simulé à partir des fréquences et amplitudes théoriques avec un bruit additif gaussien centré de rapport signal sur bruit de 5 dB.

Dans le cas de la recherche de fréquences, on peut fixer la taille des intervalles de la grille de recherche relativement à la résolution du périodogramme, toujours de l'ordre de $1/T$ en échantillonnage irrégulier. Pour notre recherche de fréquence entre 0 et 0.25 c/d, nous avons choisi une grille $\Omega_{\nu+}$ de $K = 100$ points, soit légèrement moins que le nombre maximal autorisé par la résolution, ce qui donne un écart entre les fréquences de $\Delta = 2.5 \cdot 10^{-3}$ c/d. La grille Ω_ν avec fréquences négatives comporte donc deux fois plus de points, mais l'intervalle Δ entre deux fréquences reste le même.

Nous présentons les résultats des estimations pour :

- SparSpec : la méthode est testée en faisant décroître λ jusqu'à ce que la norme du résidu ne puisse plus se distinguer de celle du bruit¹. On trouve alors 13 fréquences plutôt que 5. Pour l'analyse des résultats, nous prenons les 5 fréquences ayant les amplitudes les plus élevées.
- LOOMP : l'algorithme glouton avec grille locale est testé avec le même test d'arrêt que précédemment¹.
- BGE : le PCGS est lancé sur 10^4 itérations et les estimations sont calculées sur la dernière moitié des échantillons. On conserve les fréquences présentant une probabilité de présence supérieure à 95% (résultats de [8]).
- IBOMP, CCBP : les méthodes utilisant l'interpolation de dictionnaire sont testées à partir d'une adaptation des codes mis à disposition (voir [7]). En particulier, notons que le nombre de fréquences à estimer y est supposé connu.

En Figure 2, on peut observer les fréquences (en abscisse) et amplitudes (en ordonnée) estimées dans un intervalle de 2Δ centré autour de chaque fréquence théorique (représentée par la croix rouge) pour chaque méthode. Si la fréquence estimée n'appartient pas à cet intervalle, elle n'apparaît pas, ce qui arrive en particulier pour PoIBOMP, CCBP et SparSpec. Les méthodes BGE et LOOMP parviennent quant à elle à estimer correctement chaque fréquence, la méthode BGE permettant en plus de connaître l'écart-type sur les fréquences estimées, représenté par un trait vert horizontal.

¹Le bruit étant distribué suivant une loi $\mathcal{N}(\mathbf{0}, \sigma_e^2 \mathbf{I}_N)$, on sait calculer τ tel que $\text{Proba} \left(\frac{\|\mathbf{r}\|^2}{\sigma_e^2} < \tau \right) = 95\%$. On s'arrête quand $\frac{\|\mathbf{r}\|^2}{\sigma_e^2} < \tau$.

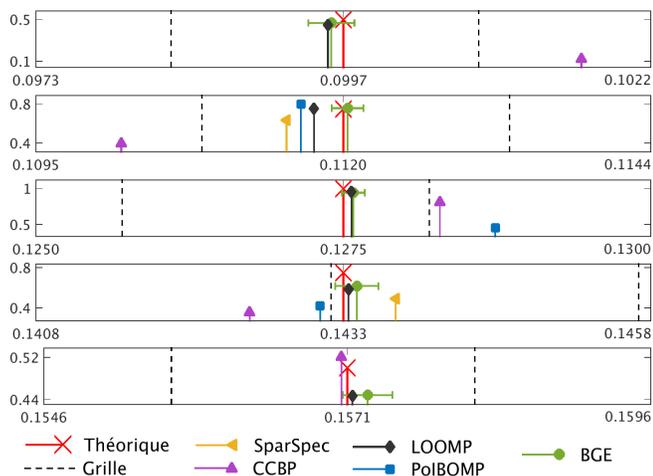


FIG. 2 – Fréquences et amplitudes estimées.

Ces résultats sont confirmés par le Tableau 1. En effet, l’erreur moyenne de l’estimation des fréquences (MSE), ici calculé en pourcentage des fréquences théoriques, est plus faible pour les méthodes BGE et LOOMP que pour SparSpec, PoIBOMP et CCBP. Par ailleurs, LOOMP (4 s) est bien plus rapide¹ que BGE (350 s).

TAB. 1 – Résultats : erreur moyenne sur l’estimation des fréquences (MSE) et temps de calcul.

	SparSpec	LOOMP	BGE	IBOMP	CCBP
MSE %	2.3	0.089	0.075	2.1	0.94
Temps	0.02 s	4.79 s	351 s	0.05 s	442 s

4 Analyse des résultats et conclusion

Les résultats confirment les attentes à propos de SparSpec : le traitement *a posteriori* des fréquences estimées sur grille fonctionne pour une grille plus fine (soit ici $K > 200$). N’ayant pas de garantie théorique pour l’estimation des fréquences hors grille par barycentre, cette approche donne de moins bons résultats qu’avec un dictionnaire continu.

Les méthodes PoIBOMP et LOOMP sont deux possibilités d’algorithme OMP sur grille continue, LOOMP utilisant une descente de gradient et PoIBOMP une interpolation pour prendre en compte les fréquences hors grille. Si LOOMP montre de très bons résultats sur notre cas test, PoIBOMP montre de sévères lacunes pour une grille lâche avec $K = 100$. En effet, l’interpolation polaire de PoIBOMP permet certes de trouver des fréquences hors grille, mais entraîne une erreur d’approximation non négligeable, alors que LOOMP n’en commet pas : la grille étant basée sur la résolution spectrale, le critère est relativement doux en ν et les optimisations (locales) de LOOMP ont toutes les chances de trouver le minimum global.

Cette influence de l’erreur d’approximation induite par l’interpolation polaire se retrouve dans les résultats de CCBP ; si l’on retrouve bien une fréquence estimée par CCBP autour de chaque fréquence théorique en Figure 2, celles-ci ne se trouvent

presque jamais dans le bon intervalle. Cela arrive même pour des grilles plus fines et se constate aussi pour PoIBOMP. On peut alors supposer que l’erreur d’approximation est peut-être accompagnée d’un problème d’identification ou de convexité dû au fait que les amplitudes sont complexes.

Enfin, deux méthodes se distinguent : LOOMP et BGE. Si LOOMP est bien plus rapide que BGE, cet algorithme glouton reste sous-optimal, dans le sens où les atomes sont choisis itérativement pour ne plus être remis en cause. Il est donc sujet à de fausses détections dues aux lobes secondaires de la fenêtre spectrale, dont dépendra fortement la qualité de l’estimation des fréquences par LOOMP. Par ailleurs, l’approche par modèle BGE permet d’obtenir une probabilité de détection et une barre d’erreur sur l’estimation des fréquences, et ne nécessite pas le réglage d’hyperparamètres dans le cadre non-supervisé.

Au regard des tests effectués, les méthodes d’approximation parcimonieuse sur dictionnaire continu sont prometteuses pour l’analyse spectrale en échantillonnage irrégulier. Une recherche plus approfondie sur l’interpolation polaire en complexe permettrait de mieux comprendre les problèmes rencontrés et ainsi de tester les méthodes PoIBOMP, CCBP mais aussi LOOMP et BGE sur des problèmes avec une fenêtre spectrale plus compliquée. Enfin, la possible utilisation de la norme atomique pour ce problème reste à étudier.

Références

- [1] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Trans. SP*, 41(12), 1993.
- [2] S. Bourguignon and H. Carfantan. Spectral analysis of irregularly sampled data using a Bernoulli-Gaussian model with free frequencies. In *ICASSP 2006 Proceedings*.
- [3] S. Bourguignon, H. Carfantan, and T. Böhm. SparSpec : a new method for fitting multiple sinusoids with irregularly sampled data. *Astronomy & Astrophysics*, 462(1), 2007.
- [4] L. Jacques and C. De Vleeschouwer. A geometrical study of matching pursuit parametrization. *IEEE Trans. SP*, 56(7), 2008.
- [5] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht. Compressed sensing off the grid. *IEEE Trans. IT*, 59(11), 2013.
- [6] C. Ekanadham, D. Tranchina, and E. P. Simoncelli. Recovery of sparse translation-invariant signals with continuous basis pursuit. *IEEE Trans. SP*, 59(10), 2011.
- [7] K. Fyhn, M. F. Duarte, and S. H. Jensen. Compressive Parameter Estimation for Sparse Translation-Invariant Signals Using Polar Interpolation. *IEEE Trans. SP*, 63(4), 2015.
- [8] M. Boudineau, H. Carfantan, S. Bourguignon, and M. Bazot. Sampling schemes and parameter estimation for non-linear Bernoulli-Gaussian sparse models. In *IEEE SSP Workshop*, 2016.

¹Les méthodes sont codées en Matlab, sauf SparSpec, codée en C.