Détection d'exoplanètes par la méthode des vitesses radiales via l'échantillonnage MCMC d'un modèle Bernoulli-Gaussien étendu

Mégane BOUDINEAU[1], Hervé CARFANTAN[1], Michaël BAZOT[2]

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

[2]Division of Sciences, New York University Abu Dhabi, United Arab Emirates

Megane.Boudineau@irap.omp.eu, Herve.Carfantan@irap.omp.eu, mb6215@nyu.edu

Résumé – La détection d'exoplanètes par la méthode des vitesses radiales peut être modélisée comme un problème inverse non linéaire, pour lequel les astrophysiciens ont développé de nombreuses méthodes de résolution. Dans cet article, nous exploitons les travaux relatifs à la parcimonie sur l'échantillonnage stochastique d'un modèle Bernoulli-Gaussien étendu. Nous montrons les atouts d'une telle méthode, notamment sur deux enjeux majeurs dans la communauté astrophysique : la sélection de l'ordre du modèle et la prise en compte de l'activité stellaire.

Abstract – Exoplanets detection with radial velocities method can be modelled as a nonlinear inverse problem, for which astrophysicians have developped many resolution methods. In this article, we benefit from sparsity-related works on stochastic sampling of an extended Bernoulli-Gaussian model. We show the assets of such a method to two main issues in astrophysics community : model order selection and stellar activity.

1 Introduction

La détection d'exoplanètes (les planètes orbitant autour d'une autre étoile que le Soleil) est un problème d'actualité en astrophysique. La méthode des *vitesses radiales* est une des plus prolifiques et consiste à mesurer le déplacement de l'étoile par rapport au centre de masse du système, sous l'effet de l'attraction gravitationnelle des planètes en orbite [1]. Ceci est effectué par mesure des décalages Doppler successifs, échantillonnés aux instants $\boldsymbol{t}$, dans le spectre de l'étoile hôte. Le signal mesuré des vitesses radiales de l'étoile, $\boldsymbol{y} \in \mathbb{R}^N$, peut alors être modélisé, dans le cadre de la mécanique newtonienne, comme la somme des contributions de plusieurs exoplanètes suivant des orbites képlériennes. Les détecter revient à estimer leur nombre, L, et les paramètres orbitaux correspondants, $\{\boldsymbol{\Theta}_\ell\}_{\ell=1..L}$, afin d'approcher au mieux le signal $\boldsymbol{y}$. Les principales difficultés de ce problème sont liées à l'échantillonnage irrégulier du signal, aux fortes non-linéarités du modèle en les paramètres orbitaux, mais aussi à la présence de bruit corrélé dû à l'activité de l'étoile. Surtout, le caractère inconnu du nombre L d'exoplanètes reste un enjeu majeur. Pour résoudre ce problème, les astrophysiciens ont développé de nombreuses méthodes, se basant par exemple sur des sélections gloutonnes de périodes à l'aide de périodogrammes [2] ou sur le développement d'algorithmes d'échantillonnage MCMC pour approcher la loi *a posteriori* des paramètres orbitaux $\{\boldsymbol{\Theta}_\ell\}_{\ell=1..L}$ à L fixé, sachant les données $\boldsymbol{y}$ [3]. Il est alors nécessaire de mettre en œuvre des méthodes de sélection de modèle pour pouvoir estimer l'ordre L, via par exemple le calcul de la vraisemblance marginale [4], ce qui représente un véritable problème en soi. Par ailleurs, une façon de prendre en compte l'activité stellaire corrélée est sa modélisation via un *processus gaussien* [5].

Cette détection d'exoplanètes est un exemple de problème inverse non linéaire auquel nous nous sommes intéressés dans [6], où l'on cherche à estimer les paramètres $(\nu_\ell, \boldsymbol{\upsilon}_\ell, \boldsymbol{x}_\ell)_\ell$ permettant d'approcher un signal $\boldsymbol{y}$ comme une combinaison linéaire de coefficients $\boldsymbol{x}_\ell \in \mathbb{R}^m$, d'un nombre inconnu L de fonctions $\mathbf{h}(\nu_\ell, \boldsymbol{\upsilon}_\ell)$ dépendant de façon non linéaire de $(\nu_\ell, \boldsymbol{\upsilon}_\ell)$. On distingue parmi les paramètres « non linéaires », le paramètre de *localisation* $\nu_\ell \in \mathcal{I}_\nu$, essentiel car il permet de distinguer deux éléments distincts, et les paramètres de forme $\boldsymbol{\upsilon}_\ell$ qui paramétrisent l'allure de la fonction non linéaire.

Nous proposons dans cet article de tirer parti des travaux en lien avec la parcimonie pour ce problème de détection d'exoplanètes. Plus particulièrement, nous nous plaçons dans un cadre Bayésien pour exploiter un modèle Bernoulli-Gaussien étendu pour prendre en compte les non-linéarités du modèle, pour lequel nous avons proposé dans [7] un algorithme d'échantillonnage stochastique efficace reposant sur la marginalisation des amplitudes. Après avoir rappelé les principaux concepts de ce modèle dans un cadre général, une adaptation aux spécificités de la détection d'exoplanètes est présentée, notamment pour l'utilisation d'un processus gaussien pour modéliser le bruit stellaire. Nous finissons par présenter des résultats concluants obtenus sur des données simulées pour montrer que notre méthode représente une véritable alternative vis-à-vis des méthodes utilisées à ce jour dans la communauté astrophysique.

2 Modèle Bernoulli-Gaussien étendu et échantillonnage stochastique

Pour résoudre les problèmes inverses non linéaires présentés précédemment, nous avons proposé dans [6] de tirer parti de la parcimonie via une reformulation du problème en Approximation Parcimonieuse Non Linéaire (APNL), où l'axe des localisations $\mathcal{I}_\nu$ est découpé un grand nombre J d'intervalles $\{\mathcal{I}_j\}_{j=1..J}$, et où l'on cherche à approcher le signal $\boldsymbol{y}$ de la façon suivante :

$$\boldsymbol{y} = \sum_{j=1}^{J} \mathbf{h}(\nu_j, \boldsymbol{\upsilon}_j)\boldsymbol{x}_j + \boldsymbol{\epsilon} \tag{1}$$

où $\nu_j \in \mathcal{I}_j$, $\boldsymbol{\upsilon}_j \in \mathcal{I}_{\boldsymbol{\upsilon}}$ et où peu de $\boldsymbol{x}_j$ sont non nuls : il s'agit de l'*a priori* de parcimonie. L'ordre du modèle L sera alors estimé comme le nombre de $\boldsymbol{x}_j \neq \mathbf{0}$.

Pour modéliser exactement la parcimonie, on introduit des variables binaires $\{q_j\}_{j=1..J}$ telles que $q_j = 0$ implique la nullité des amplitudes $\boldsymbol{x}_j$. Dans un cadre bayésien, on peut pour cela définir un modèle hiérarchique Bernoulli-Gaussien (BG) sur les couples $(q_j, \boldsymbol{x}_j)$:

$$q_j \sim \mathcal{B}(\lambda) \quad \text{et} \quad \boldsymbol{x}_j \,|\, q_j \sim \mathcal{N}(\mathbf{0}, q_j\sigma_x^2\mathbf{I}_m) \tag{2}$$

Le modèle est alors étendu à la présence des paramètres non linéaires $(\nu_j, \boldsymbol{\upsilon}_j)$ en leur donnant des lois *a priori* indépendantes des $(q_j, \boldsymbol{x}_j)$, comme par exemple des lois uniformes sur leur intervalle de définition : $\nu_j \sim \mathcal{U}(\mathcal{I}_j)$ et $\boldsymbol{\upsilon}_j \sim \mathcal{U}(\mathcal{I}_{\boldsymbol{\upsilon}})$. On parlera de modèle BG étendu (BGE).

En supposant un bruit gaussien centré de matrice de covariance $\boldsymbol{\Sigma} = \sigma_\epsilon^2\mathbf{I}_N$, on étudie alors la loi *a posteriori* dans un cadre non supervisé $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\upsilon}, \boldsymbol{x}, \lambda, \sigma_x^2, \sigma_\epsilon^2 \,|\, \boldsymbol{y})$. Nous avons proposé dans [7] un algorithme d'échantillonnage stochastique efficace pour échantillonner cette loi *a posteriori*. Alliant l'algorithme de Gibbs-Hybride proposé par [8] pour l'échantillonnage des paramètres non linéaires et le principe du *Partially-Collapsed Gibbs sampler* (PCGS) de [9] pour celui des hyperparamètres de variance, il bénéficie de la marginalisation des amplitudes $\boldsymbol{x}$ pour l'échantillonnage des paramètres d'intérêt $(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\upsilon})$; nous avons montré que cette marginalisation permettait une convergence plus rapide de l'échantillonneur par rapport à celui de la loi jointe. Un coût de calcul raisonnable est garanti grâce à des améliorations calculatoires inspirées de [10]. L'algorithme est récapitulé dans Tab. 1.

À partir d'échantillons ($\boldsymbol{q}^{(t)}$, $\boldsymbol{\nu}^{(t)}$, $\boldsymbol{\upsilon}^{(t)}$, $\boldsymbol{x}^{(t)}$, $\lambda^{(t)}$, $\sigma_x^{2(t)}$, $\sigma_\epsilon^{2(t)}$) de la loi *a posteriori*, des estimateurs peuvent alors être définis et calculés. Notamment, la moyenne des variables de Bernoulli, $\widehat{q}_j = \sum_{(t)} q_j^{(t)} \in [0;1]$ peut être interprétée comme une probabilité de détection localisée dans l'intervalle $\mathcal{I}_j$. La décision de détection se fait en binarisant $\widehat{\boldsymbol{q}}$ avec un seuil α pour obtenir la séquence binaire $\widehat{\boldsymbol{q}}_\alpha \in \{0,1\}^J$. La sélection de modèle se fait alors via la somme de la séquence binaire estimée, ce qui permet d'estimer l'ordre du modèle. Les paramètres non linéaires $(\nu_j, \boldsymbol{\upsilon}_j)$ sont estimés comme la moyenne des échantillons conditionnellement à la séquence binaire $\widehat{\boldsymbol{q}}_\alpha$.

TAB. 1 – Échantillonneur de type PCGS [7], $\boldsymbol{\theta} = (\lambda, \sigma_x^2, \sigma_\epsilon^2)$.

Itération t :

1. pour tout $j = 1 \ldots J$:
   (a) tirer $q_j^{(t)} \,|\, \boldsymbol{q}_{-j}, \boldsymbol{\nu}, \boldsymbol{\upsilon}, \boldsymbol{\theta}, \boldsymbol{y}$ $\sim$ loi de Bernoulli
   (b) tirer $\nu_j^{(t)}, \boldsymbol{\upsilon}_j^{(t)} \,|\, \boldsymbol{q}, \boldsymbol{\nu}_{-j}, \boldsymbol{\upsilon}_{-j}, \boldsymbol{\theta}, \boldsymbol{y}$ $\sim$ étape MH
2. tirer $\lambda^{(t)} \,|\, \boldsymbol{q}, \boldsymbol{y}$ $\sim$ loi Beta
3. tirer $\boldsymbol{x}^{(t)} \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\upsilon}, \boldsymbol{\theta}, \boldsymbol{y} \sim$ loi gaussienne
4. tirer $\sigma_x^{2(t)} \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\upsilon}, \boldsymbol{y}$ $\sim$ loi inverse-Gamma
5. tirer $\sigma_\epsilon^{2(t)} \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\upsilon}, \boldsymbol{y}$ $\sim$ loi inverse-Gamma

3 Détection d'exoplanètes et adaptations du modèle BGE

Le problème de détection d'exoplanètes par la méthodes des vitesses radiales est donc un cas particulier des problèmes inverses décrits précédemment. Plus précisément, la fonction non linéaire peut s'écrire de la façon suivante :

$$\mathbf{h}(P, e, \phi) = [e + \cos(\boldsymbol{\rho}(P, e, \phi)) \quad \sin(\boldsymbol{\rho}(P, e, \phi))] \tag{3}$$

où l'anomalie vraie $\boldsymbol{\rho}(P, e, \phi)$ en les instants $\boldsymbol{t}$ se calcule en fonction de l'anomalie excentrique, elle-même solution de l'équation de Kepler, qui nécessite une résolution numérique (voir [6] pour plus de précisions). Les amplitudes $\boldsymbol{x}$ sont ici vectorielles ($\boldsymbol{x}_j \in \mathbb{R}^2$) et les paramètres non linéaires sont composés de la période de rotation P de l'exoplanète autour de son étoile, de l'excentricité $e \in [0;1[$ de sa trajectoire, et d'une phase $\phi \in [0;1[$ reliée au temps de passage au périastre. C'est la période P qui jouera le rôle de paramètre de localisation, puisqu'elle permet de distinguer deux exoplanètes distinctes. Pour une excentricité nulle (trajectoire circulaire de la planète), ce modèle se ramène à celui de l'analyse spectrale en échantillonnage irrégulier, sur lequel nous avons testé le modèle BGE dans [7].

Pour définir le problème d'APNL, il est nécessaire de découper l'axe des localisations, ici $\mathcal{I}_\nu = [P^{\min}, P^{\max}]$ en J intervalles disjoints. Les planètes sont généralement recherchées entre 0.1/1 jour et quelques milliers de jours. Nous avons montré dans [6] qu'un découpage régulier en logarithme semble le plus approprié pour éviter la présence de deux exoplanètes dont les périodes se retrouvent dans le même intervalle ; c'est au moins le cas pour les planètes du système solaire. Le modèle BGE peut alors naturellement s'appliquer à ce problème d'exoplanètes ; les paramètres non linéaires prennent une loi *a priori* uniforme sur leur intervalle de définition, soit $P_j \sim \mathcal{U}(\mathcal{I}_j)$ et $(e_j, \phi_j) \sim \mathcal{U}([0;1[^2)$. Néanmoins, l'hypothèse sur le bruit nécessaire pour employer l'échantillonneur du Tab. 1 ne permet pas de prendre en compte un bruit corrélé généralement induit par l'activité stellaire. En effet, l'apparition de tâches stellaires éphémères à la surface de l'étoile peut engendrer une contribution en vitesse radiale pseudo-périodique pouvant parfois imiter la contribution d'une exoplanète. Pour filtrer ce type

d'activité, elle est parfois modélisée comme un processus gaussien [5, 4], ce qui revient à prendre un bruit $\boldsymbol{\epsilon}$ gaussien centré dont la matrice de covariance $\boldsymbol{\Sigma}$ est paramétrée suivant des paramètres $\boldsymbol{\theta}_{\text{act}}$ relatifs à l'activité de l'étoile. Le modèle pseudo-périodique utilisé par [4] est de la forme suivante :

$$\begin{aligned}\boldsymbol{\Sigma}_{m,n}(\boldsymbol{\theta}_{\text{act}}) &= (\sigma_n^2 + \sigma_{\text{jit}}^2)\delta_{m,n} \\ &+ \theta_{\text{act},1}^2 \exp\left(-\frac{(t_m - t_n)^2}{2\,\theta_{\text{act},2}^2} - \frac{\sin^2\left(\frac{\pi(t_m - t_n)}{\theta_{\text{act},3}}\right)}{2\,\theta_{\text{act},4}^2}\right) \end{aligned} \quad (4)$$

où $\boldsymbol{\theta}_{\text{act}} = (\sigma_{\text{jit}}^2, \theta_{\text{act},1}, \theta_{\text{act},2}, \theta_{\text{act},3}, \theta_{\text{act},4})$ et σ_n^2 sont les variances connues sur les mesures. Ici, les paramètres $(\theta_{\text{act},3}, \theta_{\text{act},4})$ sont liés respectivement à la période et à la durée des variations périodiques induites par l'activité stellaire. Soit les paramètres $\boldsymbol{\theta}_{\text{act}}$ sont connus (estimés via des indicateurs indépendants, voir par ex. [5]), soit ils doivent être estimés conjointement à la détection des exoplanètes.

La prise en compte d'un tel processus gaussien au sein du modèle BGE ne présente pas de problème théorique particulier, puisque ce dernier repose sur l'hypothèse d'un bruit gaussien. Dans [6], nous avons adapté les calculs des lois nécessaires à l'algorithme de la Tab. 1 à la présence d'une matrice $\boldsymbol{\Sigma}$ pleine. Par ailleurs, dans le cas où les paramètres $\boldsymbol{\theta}_{\text{act}}$ sont à estimer conjointement aux paramètres orbitaux, nous proposons dans [6] de remplacer l'étape 5 du Tab. 1 par une étape de Métropolis-Hastings (MH) permettant d'échantillonner ces paramètres suivant leur loi *a posteriori* conditionnelle marginalisée par rapport à $\boldsymbol{x}$, $p(\boldsymbol{\theta}_{\text{act}} \,|\, \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{\nu}, \upsilon, \lambda, \sigma_x^2)$.

4 Tests sur des données simulées

Nous présentons ici les résultats de notre méthode sur des signaux simulés avec activité stellaire issus de [4]. Plus précisément, nous exploitons les données des systèmes 2 et 3 de [4] représentées en Figure 1.

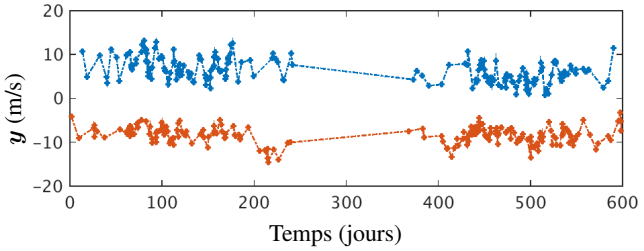

FIG. 1 – Signaux de vitesses radiales simulés [4] : système 2 en bleu et système 3 en rouge.

Ces signaux de $N = 200$ échantillons sont chacun constitués des contributions képlériennes de $L = 2$ exoplanètes ainsi que d'une contribution d'activité stellaire ; l'ensemble des paramètres à retrouver sont donnés dans [4]. Pour chaque planète, les auteurs ont attribué un degré de difficulté. Pour tester notre méthode, nous avons effectué les réglages suivants, pour chacun des deux signaux :

- Les périodes sont recherchées entre $P^{\min} = 1$ jour et $P^{\max} = 10000$ jours sur une grille de $J = 100$ intervalles.
- L'algorithme en Tab. 1 (étapes 1-4) est employé, en fixant les paramètres du processus gaussien $\boldsymbol{\theta}_{\text{act}}$, comme proposé dans [4].
- 10 chaînes avec différentes initialisations des paramètres et de la graine aléatoire sont lancées sur 10000 itérations.
- Les estimations sont réalisées sur les 5000 dernières itérations.

Nous présentons en Figure 2 les séquences de Bernoulli moyennées $\widehat{\boldsymbol{q}}$ sur les 5000 dernières itérations, cumulées sur les 10 chaînes indépendantes ; ces moyennes étant assez similaires pour chacune des chaînes, on peut déduire la bonne convergence de notre échantillonneur. Pour le système 3, les 2 exoplanètes, pourtant supposées difficiles dans [4], sont détectées avec une forte probabilité (supérieure à 95%), sans fausse détection. Le cas du système 2 est un peu plus délicat. Tandis que la première exoplanète, supposée facile, est retrouvée sans difficulté (probabilité de détection de presque 100%), la seconde, supposée difficile, présente une probabilité de détection beaucoup plus faible (inférieure à 50% sur toutes les chaînes) et une troisième exoplanète semble ressortir à 3 intervalles de distance, toujours avec une faible probabilité.

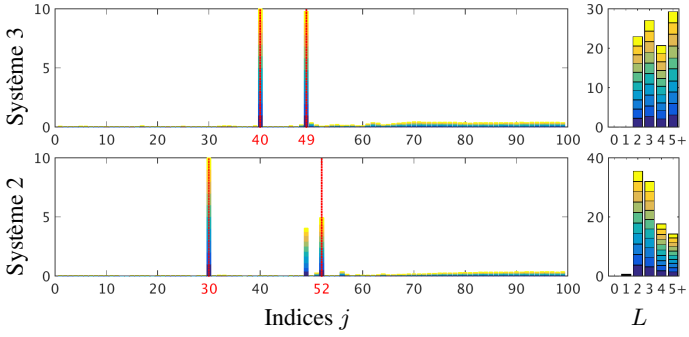

FIG. 2 – Résultats cumulés sur 10 chaînes indépendantes (une couleur = une chaîne). À gauche, séquences de Bernoulli moyennées $\widehat{\boldsymbol{q}}$ (indices rouges : périodes théoriques). À droite, histogrammes de $L^{(t)}$.

La Figure 3(a) montre pour le système 2, l'évolution de la séquence de Bernoulli sur les 5000 dernières itérations, pour les indices $\Omega = \{49, 51, 52\}$ proches de celui de la seconde exoplanète. On observe un phénomène d'*anti-corrélation* : si l'on définit la séquence $q_\Omega^{(t)}$ égale à 1 si un seul des $q_j^{(t)} = 1$ pour $j \in \Omega$, alors $q_\Omega^{(t)} = 1$ pour 92.1% des itérations. Les histogrammes des échantillons des périodes et des excentricités aux indices Ω sont représentés respectivement en Fig. 3(b) et Fig. 3(c). Ils permettent de mettre en exergue deux localisations bien distinctes pour les périodes, et par conséquent un caractère bimodal de la loi *a posteriori*. De telles observations nous permettent de conclure avec une probabilité de 92.1% à la présence d'une seule exoplanète pour les indices Ω, mais dont la localisation est incertaine, dans l'un ou l'autre des deux modes disjoints.

Enfin, la probabilité *a posteriori* de l'ordre du modèle

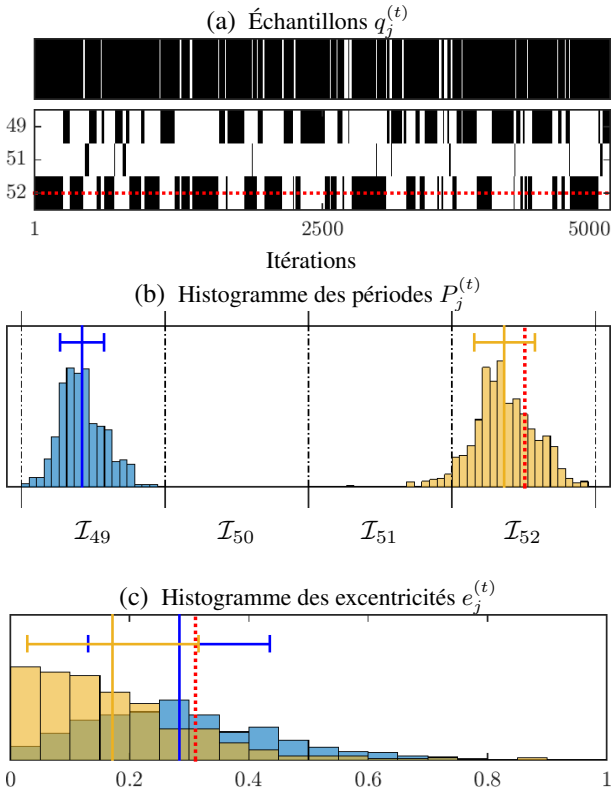

FIG. 3 – (a) Évolution de $q_j^{(t)}$ pour $j \in \Omega$ sur les 5000 dernières itérations (en noir $q_j^{(t)} = 1$). Au dessus, évolution de $q_\Omega^{(t)}$. (b) Histogrammes des périodes $P_{j\in\Omega}^{(t)}$, en regroupant par couleur les deux modes (bleu et jaune), dont les moyennes sont représentées par des traits verticaux avec barres d'erreur à $\pm 1\sigma$. Le trait rouge vertical indique la valeur théorique. (c) Histogrammes des excentricités $e_{j\in\Omega}^{(t)}$, où les couleurs sont relatives au mode associé.

$p(L \,|\, \boldsymbol{y})$ peut être estimée sans difficulté à partir des échantillons des séquences de Bernoulli, comme l'histogramme de l'ordre du modèle $L^{(t)} = \sum_{j=1}^{J} q_j^{(t)}$. En Fig. 2, nous présentons ces histogrammes cumulés sur les 10 chaînes. Pour les deux systèmes, $p(L = 2 \,|\, \boldsymbol{y})$ est du même ordre de grandeur voir plus faible que pour L supérieur, tandis que les 2 exoplanètes sont clairement détectées avec une étude plus approfondie des échantillons des autres variables (Fig. 2 et 3).

5 Conclusions et perspectives

Le Tab. 2 récapitule les résultats d'estimation obtenus à partir de nos échantillonneurs. Sur ces données, notre méthode retrouve sans difficulté les deux planètes pour chaque système (avec par exemple un seuil de $\alpha = 0.85$). Les périodes estimées sont assez proches des valeurs théoriques et des barres d'erreur peuvent être calculées comme l'écart-type des échantillons.

Le modèle BGE et l'échantillonneur associé représentent donc une alternative innovante aux méthodes existantes pour la détection d'exoplanètes, et apporte un complément important pour l'estimation du nombre d'exoplanètes notamment, puis-

TAB. 2 – Résultats d'estimation et de détection pour les deux systèmes. Périodes (en jours) estimées avec barre d'erreur à 1σ et probabilités de détection.

Sys.	Difficulté	P^{th}	$\widehat{P}$	Proba. (%) de détection	
2	facile	15.96	15.96 ± 0.016	100	
	difficile	120.5	90.24 ± 1.55 118.95 ± 3.30	44.07 55.93	92.1
3	difficile	40.4	40.15 ± 0.183	100	
	difficile	91.9	92.06 ± 1.10	98.02	

qu'il est estimé directement à partir de la séquence de Bernoulli moyennée $\widehat{\boldsymbol{q}}$. Bien que la valeur de $p(L \,|\, \boldsymbol{y})$ semble d'un intérêt particulier pour les astrophysiciens, elle se révèle bien moins informative qu'une étude plus approfondie des échantillons fournis par le modèle BGE. La présence de l'activité stellaire ne pose pas ici de problèmes puisqu'elle peut être modélisée comme un processus gaussien au sein du modèle BGE.

Des études plus poussées sont actuellement en cours pour pouvoir montrer son efficacité sur des données mesurées de vitesses radiales, en estimant les paramètres de l'activité, dont des résultats préliminaires sont donnés dans [6].

Références

[1] M. Perryman. *The exoplanet handbook*. Cambridge university press, 2011.

[2] M. Zechmeister and M. Kürster. The generalised lomb-scargle periodogram-a new formalism for the floating-mean and keplerian periodograms. *A&A*, 496(2) :577–584, 2009.

[3] E.B. Ford. Improving the efficiency of Markov Chain Monte Carlo for analyzing the orbits of extrasolar planets. *The Astrophysical Journal*, 642(1) :505, 2006.

[4] B.E. Nelson, E.B. Ford, J. Buchner, et al. Quantifying the evidence for a planet in radial velocity data. *ArXiv eprints*, 2018.

[5] R.D. Haywood, A. Collier Cameron, D. Queloz, et al. Planets and stellar activity : hide and seek in the corot-7 system. *MNRAS*, 443(3) :2517–2531, 2014.

[6] M. Boudineau. *Vers la résolution optimale de problèmes inverses non linéaires parcimonieux grâce à l'exploitation de variables binaires sur dictionnaires continus : applications en astrophysique*. Thèse de doctorat, Université de Toulouse, Toulouse, février 2019.

[7] M. Boudineau, H. Carfantan, S. Bourguignon, and M. Bazot. Sampling schemes and parameter estimation for nonlinear Bernoulli-Gaussian sparse models. In *IEEE SSP Workshop*, 2016.

[8] S. Bourguignon and H. Carfantan. Spectral analysis of irregularly sampled data using a Bernoulli-Gaussian model with free frequencies. In *Proceedings ICASSP*, 2006.

[9] D. A. van Dyk and T. Park. Partially collapsed gibbs samplers : Theory and methods. *Journal of the American Statistical Association*, 103 :790–796, 2008.

[10] D. Ge, J. Idier, and E. Le Carpentier. Enhanced sampling schemes for MCMC based blind Bernoulli-Gaussian deconvolution. *Signal Process.*, 91(4) :759–772, avril 2011.