Astronomy & Astrophysics

SparSpec: a new method for fitting multiple sinusoids with irregularly sampled data

S. Bourguignon, H. Carfantan, and T. Böhm

Laboratoire d'Astrophysique de Toulouse-Tarbes, UMR 5572, Université Paul Sabatier Toulouse III and CNRS,
14 avenue Édouard Belin, 31400 Toulouse, France
e-mail: [bourgui;Herve.Carfantan;boehm]@ast.obs-mip.fr

ABSTRACT

Context. The location of pure frequencies in the spectrum of an irregularly sampled time series is an important topic in astrophysical data analysis. Especially in the domain of asteroseismology, a highly precise and *unambiguous* study of frequencies in photometric light or radial velocity curves is required.

Aims. Due to sampling irregularities and large observational gaps, the classic methods for frequency estimation (prewhitening techniques, CLEAN, CLEANEST, etc.) sometimes suffer false detections. We propose a new framework for this problem that allows a more precise and unambiguous frequency location.

Methods. Multisine fitting is addressed as the sparse representation of the data in an overcomplete dictionary of frequencies, hence the name *SparSpec* for the method. We model the data as the sum of an arbitrarily large number of pure frequencies, discretised on a fixed grid. Among all the many representations fitting the data, we seek the one with the fewest non-zero amplitudes. This solution can be computed by minimising a convex criterion with no local minima. A computationally efficient and convergent optimisation strategy is derived and a user-friendly software implementing *SparSpec* is provided online at http://www.ast.obs-mip.fr/Softwares.

Results. The method is first illustrated on a simple test example where *SparSpec* correctly locates the frequencies while classic methods fail. Then, simulations on more realistic artificial time series reveal the interest of this new methodology in terms of robustness toward sampling aliases. An application to the radial velocity curve of the pre-main sequence Herbig Ae star HD 104237 is finally presented, where the method is able to determine oscillation frequencies even in the presence of strong low-frequency perturbations such as orbital movements. While *SparSpec* mainly confirms previously published studies for the four more important frequencies, it suggests some ambiguity about the position of a fifth frequency. Additional simulations show that the *SparSpec* results are more plausible.

Key words. methods: data analysis – stars: oscillations – stars: individual: HD 104237

1. Introduction

Detection and estimation of pure frequencies from irregularly sampled time series is a frequent problem in astrophysical data analysis. Especially in the domain of asteroseismology, all science is based on obtaining the largest number of stellar oscillation frequencies by observing different variables such as, e.g., photometric light or radial velocity curves. Knowing the largest number of frequencies with very high accuracy and low false alarm probability provides strong constraints on the subsequent stellar interior modelling. Stellar oscillation codes predict physical oscillation modes and associated frequencies, the goal obviously being to reproduce the observed frequencies by adjusting the stellar structure model and additional input parameters.

From an observational point of view, the observed light or radial velocity curves usually suffer from incomplete time coverage. In particular, periodic gaps caused by the earth's rotation and revolution, may generate high secondary lobes in the spectral window and consequently perturb the interpretation of the Fourier spectrum. Moreover, long-term observations are usually irregularly time-spaced, so that many classic signal processing tools are inapplicable (Kay 1988).

Specific methods of frequency estimation have been developed in the astrophysical context. On the one hand, phase-binning techniques (Lafler & Kinman 1965; Stellingwerf 1978;

Dworetsky 1983) aim at specifically searching *one* short-time periodicity from small data sets with a long time coverage, but cannot be used as general spectral analysis methods. On the other hand, Fourier-based methods are more adapted to the analysis of sinusoidal oscillations. The Lomb-Scargle periodogram (Lomb 1976; Scargle 1982) and the date-compensated discrete Fourier transform (DCDFT, Ferraz-Mello 1981) take the sampling irregularities into account, but are only statistically valid for a single sinusoid (Foster 1996). For multisine estimation, iterative procedures are generally preferred where, at each iteration, a pure frequency is estimated after a *prewhitening* step, i.e., from the data where the contributions of previously estimated frequencies have been removed (Barning 1963; Gray & Desikachary 1973). This strategy, however, is subject to false detections caused by sampling artifacts in the Fourier spectrum. Posterior refinements based on the CLEAN method (Högbom 1974) aim at improving the prewhitening efficiency (Roberts et al. 1987; Foster 1995), but can also lead to unsatisfactory frequency location.

The aim of this paper is to present a new method of estimating the spectral components of a signal from irregularly sampled data. Based on the framework of sparse representations, we named this method *SparSpec*. Multiple sinusoids are estimated by reconstructing a high-dimensional vector of spectral amplitudes corresponding to a whole (discrete) frequency grid.

A *sparse* representation of the spectrum is encouraged, that is, a representation that fits the data with the fewest non-zero spectral amplitudes. This solution can be computed by minimising a convex criterion, so that no local minimum can be reached. Thus, the proposed estimator shows less sensitivity to sampling artifacts in terms of frequency location.

We first recall in Sect. 2 the bases of the classic methods and focus on their limitations. Then, Sect. 3 sets the framework of the *SparSpec* method and Sect. 4 describes the practical implementation. Simulations in Sect. 5 reveal both the advantages of this method compared to the prewhitening techniques on artificial data similar to those used by Foster (1995) and its limitations in case of physical indeterminacy. In Sect. 6, the power of this new methodology is illustrated in application to the radial velocity curve of Herbig Ae star HD 104237 limited to the 1999 data set (Böhm et al. 2004). This analysis confirms the detection of four frequencies with similar parameters to the previous results by Böhm et al. (2004), but it also helps to identify a fifth frequency.

2. Limits of the classic approaches

A direct approach to the estimation of multiple oscillations consists in the identification of a multisine model. This section considers the difficulties that may arise when trying to estimate the corresponding parameters.

2.1. The direct multisine estimation

Let $y_{n,n=1...N}$ represent N samples of a physical quantity measured at instants $t_{n,n=1...N}$ and consider the multisine model $\mathcal{M}_M$:

$$y_n = \sum_{m=1}^{M} A_m \cos\left(2\pi\nu_m t_n + \phi_m\right) + \epsilon_n \tag{1}$$
$$\triangleq \mathcal{M}_M(\boldsymbol{\theta}_M, t_n) + \epsilon_n,$$

where parameters $\boldsymbol{\theta}_M = (A_m, \nu_m, \phi_m)_{m=1...M}$ and the number of frequencies M are unknown, and ϵ_n is a perturbation term that stands for model errors and observation noise. For a fixed number M of sinusoids, the determination of the best least-square fitting, i.e., the minimisation with respect to $\boldsymbol{\theta}_M$ of the quadratic criterion,

$$\mathcal{I}_M(\boldsymbol{\theta}_M) = \sum_{n=1}^{N}\left(y_n - \mathcal{M}_M(\boldsymbol{\theta}_M, t_n)\right)^2, \tag{2}$$

corresponds statistically to the maximum likelihood estimation of $\boldsymbol{\theta}_M$ for additive independent and identically distributed (i.i.d.) Gaussian perturbations ϵ_n. In practice, however, this estimation raises a difficult question:

- for a given M, model $\mathcal{M}_M$ is not linear[1] in the frequencies ν_m, and criterion $\mathcal{I}_M$ has many local minima in these parameters (Stoica et al. 1989). Thus, a *global* optimisation strategy is tricky;
- selecting the number M of searched frequencies is a hard task. One may want to find a relatively small number of significant sinusoids, but in such a way that the corresponding model fits the data with sufficient adequation.

---

[1] Model (1) is also not linear in the phases ϕ_m. However, it can be written $y_n = \sum_{m=1}^{M} a_m \cos(2\pi\nu_m t_n) + b_m \sin(2\pi\nu_m t_n) + \epsilon_n$ with $a_m = A_m \cos(2\pi\phi_m)$, $b_m = -A_m \sin(2\pi\phi_m)$. Thus, the only important non-linearities to deal with concern the frequency parameters.

Classic *sequential* methods iteratively remove sinusoidal components from the data, which are identified as the maxima of the Fourier spectrum of the residuals after a prewhitening step (Barning 1963; Gray & Desikachary 1973). However, due to sampling artifacts, these maxima may not correspond to true frequencies – see the example introduced in Sect. 2.2. This is the main drawback of the basic prewhitening method as it may lead to fatal error propagation, so the solution may not correspond to the minimum of criterion (2). Two kinds of refinements aim at improving its efficiency:

- the CLEAN procedure as described by Roberts et al. (1987) introduces a *clean gain*, i.e., only removes a fraction g of the sinusoidal component ($0 < g < 1$) at each iteration. This ad-hoc solution tries to prevent the propagation of errors through the algorithm, but it does not provide the identification of model (1) directly, as a single spectral line will be estimated by several components with close frequencies. A posterior reconvolution step is usually performed to make the result more readable, but it reduces the frequency resolution of the estimation and does not directly give values for parameters $\boldsymbol{\theta}_M$;
- the CLEANEST method adds a step of local optimisation of criterion (2) with respect to parameters $\boldsymbol{\theta}_M$ to each iteration. A second refinement of this strategy holds in the use of the date-compensated Fourier transform by Ferraz-Mello (1981), which is more efficient than the basic Fourier spectrum for frequency estimation – see Foster (1995) for a formal description. While satisfactory results are usually achieved in practice, the best fit (i.e., the lowest least-square value of $\mathcal{I}_M$) cannot be guaranteed.

Note that those strategies also suffer from the difficult estimation of model order M, which is related to the number of iterations before algorithm stops: this requires a decision from the user that may be hard to handle. Many empirical or statistical tests have been proposed to associate some *confidence level* to the amplitude of each extracted component (e.g., Breger et al. 1993; Kuschnig et al. 1997; Reegen 2004), but this is beyond the scope of this paper.

2.2. A test example where classic methods fail

The artificial data in Fig. 1 are two noise-free sinusoids irregularly sampled during four days with daily 8-hour gaps. Frequencies are set such that $f_2 - f_1 = 2$ c/d (cycles per day), so that the sidelobes due to the spectral window for each frequency superimpose at $f_w = (f_1 + f_2)/2$, leading to a wrong global maximum of the Fourier spectrum. Because of inaccurate initialisation, the classic prewhitening method, as well as CLEAN and CLEANEST, would lead to erroneous results. Note that here a more sophisticated methodology would be able to identify the two lines, as the *global* CLEANEST described by Foster (1995). Nevertheless, this is an ad-hoc solution that cannot be used automatically for fitting multiple sinusoids.

This simple example shows the sensitivity of prewhitening procedures when they meet sampling artifacts. We will see in Sect. 3 that the proposed formulation is not subject to these limitations.

3. An alternative strategy

The limits of classic identification methods are caused essentially by the multimodality of criterion $\mathcal{I}_M$, which is the
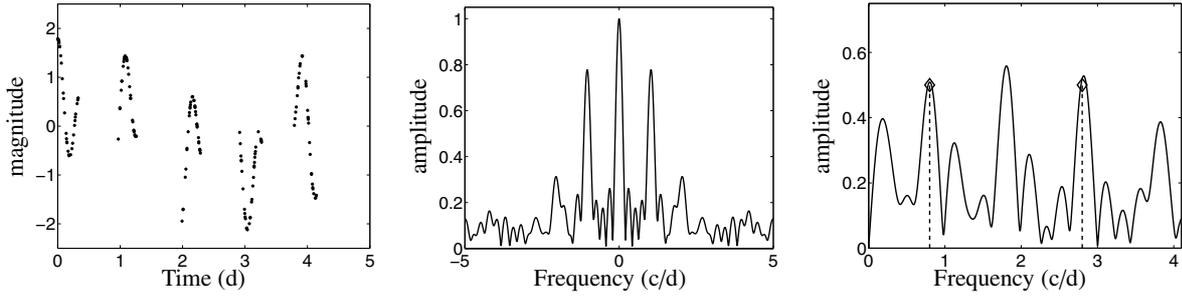
Fig. 1. A pathological test example, showing the irregularly sampled artificial data (*left*), the spectral window (*centre*) and the Fourier spectrum (*right*). Diamonds indicate the two true spectral lines.

consequence of the non-linearity of model (1) in the unknown frequencies ν_m. In this approach, we propose an alternative model that is linear in the frequency parameters.

3.1. A linear model

Data are modelled as the noisy superposition of an *arbitrarily large* number K of sine waves with discretised frequencies $f_k = \frac{k}{K} f_{\max}, k=0\ldots K$ ranging from 0 to a maximum $f_{\max}$. This can be written:

$$y_n = \sum_{k=-K}^{K} x_k \exp\left(j2\pi \frac{k}{K} f_{\max} t_n\right) + \epsilon_n, \quad n = 1\ldots N \qquad (3)$$

where x_k are unknown complex spectral amplitudes and ϵ_n is the usual perturbation term. For real-valued data y_n, one has of course $x_{-k} = x_k^*$, where notation $*$ refers to the conjugate complex number. Then x_k, as well as x_{-k}, represents amplitude and phase information of a potential sinusoid at frequency $f_k = \frac{k}{K} f_{\max}$. Collecting all quantities in matrix-vector forms, model (3) is written as

$$y = Wx + \epsilon \qquad (4)$$

where $y = [y_1 \ldots y_N]^T$, $x = [x_{-K} \ldots x_K]^T$, $\epsilon = [\epsilon_1 \ldots \epsilon_N]^T$, and matrix $W = \{\exp(j2\pi f_k t_n)\}_{k=-K\ldots K, n=1\ldots N}$.

Of course, if we want to estimate the parameters of a multisine model – i.e., a relatively small number of oscillation frequencies – then only a few components in x, say $\{x_{k, k\in\mathcal{K}}\}$, should be non-zero: the representation of the spectrum by vector x is said to be *sparse*. Sinusoidal oscillations are then identified at frequencies $\{f_{k, k\in\mathcal{K}}\}$, with corresponding amplitudes $\{2|x_k|, k\in\mathcal{K}\}$ and phases $\{\arg x_k, k\in\mathcal{K}\}$. This approach allows one to tackle the multisine estimation problem in a different way than the classic use of model (1). Model (3) is linear and considers jointly a high number K of potential frequencies, where the only parameters to estimate are amplitudes. On the contrary, in model (1) the frequency parameters have to be estimated, which makes the problem trickier.

3.2. Sparse representations

The frequency resolution in model (3) is intrinsically limited by the discretisation step of the frequency grid $f_{\max}/K$. Thus, the size of unknown vector x, that is, $2K + 1$, must be very large to yield a resolution comparable to that of classic prewhitening methods, even larger than the amount of data (typically, a vector with several thousand unknown parameters is used). In this case, an infinite number of vectors x perfectly fit the data with

model (3), i.e., solve equation $y = Wx$. Among all possible solutions, our aim is to obtain the *sparsest* one, i.e., the solution of (3) with the fewest non-zero complex amplitudes. In the *noise-free* case, this problem is written as

(P0) : find $\arg \min \|x\|_0$ under the constraint $y = Wx$,

where the ℓ^0-norm $\|x\|_0$ is the number of non-zero components in x. Fuchs (2004) has shown under some conditions that solving (P0) is equivalent[2] to solving (P1):

(P1) : find $\arg \min \|x\|_1$ under the constraint $y = Wx$,

substituting $\|x\|_0$ by the ℓ^1-norm: $\|x\|_1 = \sum_k |x_k|$. For *noisy data* – our case – this problem is converted into the minimisation of a composite criterion (Donoho et al. 2006), which defines the spectral estimator as:

$$\widehat{x} = \arg \min \; J(x), \quad J(x) = \frac{1}{2}\|y - Wx\|^2 + \lambda \sum_{k=-K}^{K} |x_k|, \qquad (5)$$

where the value of $\lambda > 0$ balances between data fidelity – with a low value of the quadratic misfit measure – and prior sparsity – with a low ℓ^1-norm. Definition (5) is the core of the *SparSpec* estimation.

Definition (5) of the spectral estimator can be described within the Tikhonov regularisation framework (Tikhonov & Arsenin 1977; Demoment 1989), where the second term in criterion J is called the *penalisation* function. Similar formulations of the spectral analysis problem have been proposed during the past ten years, using other penalisation functions (Sacchi et al. 1998; Ciuciu et al. 2001). In our case, the choice of the ℓ^1-norm was motivated by the strong theoretical background concerning sparse representations, as well as for technical reasons. First, it leads to a convex criterion, so its global minimum can be reached by descent techniques. Second, a computationally efficient algorithm can be obtained (see Sect. 4.2), whereas algorithms proposed by the aforementioned authors are only efficient for regularly sampled data.

3.3. Comparison with the classic approach

Compared with model (1), the association of the *linear* model (3) with sparse representations leads to several practical consequences.

---

[2] Note that the sparsity properties of this minimiser have been theoretically studied for *real-valued* parameters (Fuchs 2004; Donoho et al. 2006). In practice, sparse estimators are also obtained in the considered complex case.
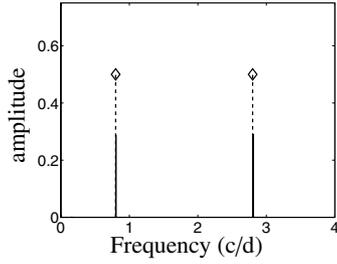
Fig. 2. *SparSpec* estimator $|\widehat{x}_\lambda|$ obtained from data in Fig. 1 for $\lambda = 30$ (full line). The two true lines (diamonds) are correctly estimated in frequency.

i) As the complex modulus function is a convex function, so is the criterion to minimise. Thus, there are no local minima and local optimisation algorithms are guaranteed to converge toward the lowest value of the criterion. This is not the case for model (1) where a least-square fitting procedure can be trapped in local modes. Figure 2 illustrates this point in application to the example introduced in Fig. 1, where prewhitening methods failed. For an adequate tuning of parameter λ – as will be discussed in Sect. 4.1 – the *SparSpec* estimation enables the correct identification of the two spectral lines and does not provide false detections. Note that amplitudes are systematically underestimated by this method; this point will be studied in Sect. 5.

ii) Since J is not strictly convex, the unicity of minimiser $\widehat{x}$ cannot be theoretically guaranteed. It can be shown, however, that J has a unique minimiser with less than $N/2$ non-zero components, where N is the number of data (Bourguignon 2006). In other words, if the minimiser $\widehat{x}$ of criterion (5) has less than $N/2$ non-zero components – if it is sparse enough – then it is the only sparse solution.

iii) Here order selection – i.e., estimation of the number of sinusoids extracted from the data – is a direct byproduct of the estimation result since all components of vector x are estimated *jointly*. With this approach, the number of frequencies corresponds to the amount of sparseness in x, which is controlled by the value of λ.

Of course, the price to pay for these properties is consequently high, as it results in a considerable increase in the number of parameters to estimate. Thus, the estimation procedure may require a much higher computational time than, e.g., sequential prewhitening methods do. However, the efficient optimisation procedure proposed in Sect. 4.2 allows one to deal with high-dimensional unknown vectors x – that is, high spectral resolution – at a reasonable computational cost.

4. Practical implementation

Definition (5) of the *SparSpec* estimator requires two essential points to be examined with precision. First, a practical rule is needed for tuning the λ parameter. Then, an efficient optimisation strategy must be defined to compute the minimiser $\widehat{x}$ at low computational cost.

4.1. Tuning the λ hyperparameter

We first give a statistical interpretation of λ using the Bayesian framework. Then we show that, in practice, optimal λ should belong to some interval $[\lambda_{\min}, \lambda_{\max}]$ depending on the data. Based

on this, we propose some heuristic rules to tune the λ parameter satisfactorily.

Bayesian interpretation

The proposed estimation procedure can be described through the Bayesian statistical framework (Idier 2001). We suppose here that additive perturbations ϵ_n are i.i.d. centred Gaussian random variables with variance σ_ϵ^2. Minimiser $\widehat{x}$ of the penalised criterion writes equivalently as the maximiser of:

$$\exp\left(-\frac{1}{2\sigma_\epsilon^2}\|y - Wx\|^2\right) \exp\left(-\frac{\lambda}{\sigma_\epsilon^2}\|x\|_1\right) \qquad (6)$$

where first term corresponds, up to a normalisation coefficient, to the likelihood $\mathcal{L}(y; x)$ of model (3).

The Bayesian statistics stipulates that the posterior probability of unknown vector x given the observations y is proportional to the product of the likelihood by the prior probability distribution on x: $p(x|y) \propto \mathcal{L}(y; x)\, p(x)$. Within this framework, estimator $\widehat{x}$ corresponds to the maximiser of the posterior probability distribution $p(x|y)$, where the second term in expression (6) can be identified with a prior distribution $p(x)$ on the unknown vector that writes (up to a normalisation coefficient):

$$p(x) \propto \exp\left(-\frac{\lambda}{\sigma_\epsilon^2}\|x\|_1\right) = \prod_{k=-K}^{K} \exp\left(-\frac{\lambda|x_k|}{\sigma_\epsilon^2}\right).$$

That is, $\widehat{x}$ can be defined as the Bayesian maximum a posteriori estimator assuming that the modulus of the spectral amplitudes are i.i.d. according to a centred *Laplace* prior distribution with variance $\mathbb{E}\left[|x_k|^2\right] = 2\sigma_\epsilon^4/\lambda^2$. Then, the *signal power*, which is the power due to the spectral content of x in every sample y_n, reduces to $(2K + 1)\mathbb{E}\left[|x_k|^2\right] = 2(2K + 1)\sigma_\epsilon^4/\lambda^2$, and the signal-to-noise ratio is finally written as

$$SNR = \frac{signal\ power}{noise\ power} = 2(2K + 1)\frac{\sigma_\epsilon^2}{\lambda^2}. \qquad (7)$$

Consequently, the prior knowledge of the signal-to-noise ratio and the noise power gives a statistically meaningful value for parameter λ. As this information may not be available, we propose heuristic methods for efficiently tuning the λ parameter in the following.

Acceptable bounds on λ

It can be shown (Bourguignon 2005) that a minimum point of criterion (5) is characterised by the following condition:

$$\widehat{x}\ \text{minimises}\ J \ \Leftrightarrow \ \begin{cases} (i)\ \forall k\, /\, |\widehat{x_k}| = 0: \ |r_k| \le \lambda, \\ (ii)\ \forall k\, /\, |\widehat{x_k}| \ne 0: \ r_k + \lambda\dfrac{x_k}{|x_k|} = 0, \end{cases} \qquad (8)$$

with $r = W^\dagger(W\widehat{x} - y)$, where notation $W^\dagger$ designs the conjugate transpose of matrix W. This is a generalisation of the condition that holds for a differentiable criterion, which stipulates that the gradient has to be zero at the minimum. Property (8) allows reasonable bounds to be derived for λ:

– it is straightforward to show that $\widehat{x}$ is identically zero if $\lambda > \lambda_{\max} = \max\left|W^\dagger y\right|$;

– $\boldsymbol{r}$ is merely the discretised Fourier spectrum of the residuals $\mathsf{W}\widehat{\boldsymbol{x}} - \boldsymbol{y}$, and property (8) imposes $|r_k| \leq \lambda$: the λ parameter can be viewed as the maximum peak amplitude allowed in these residuals. Thus, it is reasonable to set a minimum value for λ to some small fraction of the initial Fourier spectrum, say 1%: $\lambda_{\min} = \max \left| \mathsf{W}^\dagger \boldsymbol{y} \right| / 100$. Note that estimator $\widehat{\boldsymbol{x}}$ naturally suffers some bias in the amplitude estimation (see Sect. 5), so $\boldsymbol{r}$ is not exactly the residual spectrum after removing the extracted frequencies, in the sense of the residual after prewhitening steps for sequential methods.

Practical use

In practice, efficient tuning of λ can be obtained using the following principle: first, compute the minimiser of criterion (5) for several values of $\lambda \in [\lambda_{\min}, \lambda_{\max}]$, say $\widehat{\boldsymbol{x}}_\lambda$, and then decide which is the best. The heuristic *L*-curve technique uses a plot of the value of the penalisation function at the solution $R(\widehat{\boldsymbol{x}}_\lambda)$ versus the goodness-of-fit measure $Q(\widehat{\boldsymbol{x}}_\lambda) = \frac{1}{2} \left\| \boldsymbol{y} - \mathsf{W}\widehat{\boldsymbol{x}}_\lambda \right\|^2$. The resulting curve was shown to have an *L* shape in the case of quadratic penalisation (Hansen 1992); selecting a solution corresponding to the corner of the *L* then gives a satisfactory compromise between both Q and R terms. In order to plot energetic quantities on both axes, we prefer to use the curve $\left(Q(\widehat{\boldsymbol{x}}_\lambda), R(\widehat{\boldsymbol{x}}_\lambda)^2 \right)$ here. Although the use of this method for non-quadratic ℓ^1 penalisation is not based on any theoretical property, satisfactory results were obtained in practice.

Another possibility is simply to select among all $\widehat{\boldsymbol{x}}_\lambda$ those that work the best visually. Start with $\lambda = \lambda_{\max}$ and then progressively decrease λ until *under-regularisation* is detected, characterised by the presence of spurious peaks along the full reconstruction grid. Examples of a correct selection of λ using these methods will be given in Sect. 5.

4.2. Optimisation algorithm

As function $J(\boldsymbol{x})$ is not differentiable for every vector containing a zero value (the complex modulus function is not differentiable at zero), classic gradient-based descent techniques cannot be applied. In the case of real-valued variables, the minimisation of criterion J can be written as a constrained quadratic optimisation problem (Fuchs 2004), which can be solved using standard algorithms from many programming libraries. In our case, variables x_k are complex-valued and this is not true any more.

To compute the *SparSpec* solution, we minimise criterion (5) by an iterative coordinate descent (ICD) procedure (Bertsekas 1995), which consists in performing successive one-dimensional minimisation steps with respect to each complex-valued parameter x_k. As criterion J is convex, the convergence of such algorithm toward the minimum of the criterion is ensured (Alliney & Ruzinsky 1994). The use of an ICD algorithm to minimise the ℓ^1-penalised criterion leads to a computationally attractive implementation. Note that a similar strategy was proposed by Alliney & Ruzinsky (1994) for real-valued parameters, but our implementation outperforms their procedure, thanks to the following property: every scalar minimisation has an explicit solution that can be computed almost instantaneously. Indeed, let $\boldsymbol{w}_k$ design the column with index k in matrix W and $\boldsymbol{e}_k = \boldsymbol{y} - \sum_{\ell \neq k} \boldsymbol{w}_\ell x_\ell$ and consider criterion J only as a function of

parameter x_k. It can be shown (Bourguignon 2005) that

$$x_k^{\min} = \arg\min_{x_k} J(\boldsymbol{x}) \;\Leftrightarrow\; \begin{cases} \text{if } |\boldsymbol{w}_k^\dagger \boldsymbol{e}_k| \leq \lambda \;:\; x_k^{\min} = 0, \\ \text{if } |\boldsymbol{w}_k^\dagger \boldsymbol{e}_k| > \lambda \;:\; \arg x_k^{\min} = \arg \boldsymbol{w}_k^\dagger \boldsymbol{e}_k \\ \quad \text{and } |x_k^{\min}| = \frac{1}{N} \left(|\boldsymbol{w}_k^\dagger \boldsymbol{e}_k| - \lambda \right). \end{cases}$$
$$(9)$$

In practice, iterations like (9) are repeated until condition (8) is verified up to a numerical threshold. One implementation of *SparSpec* is available online at `http://www.ast.obs-mip.fr/Softwares`.

5. Simulation results

The efficiency of *SparSpec* is evaluated using two artificial data sets similar to those proposed by Foster (1995), which consist in three sinusoids with periods $\nu_1^{-1} = 370$, $\nu_2^{-1} = 230$, and $\nu_3^{-1} = 100$ days and amplitudes 3, 2.828, and 3, respectively. A constant value of 10 was also added. An initial data set was generated with 200 points sampled every 10 days. Data set (A) shows gaps of 100 days every 365 days and gaps of 10 days every 30 days. Data set (B) shows gaps of 200 days every 365 days and gaps of 10 days every 30 days. Data sets (A) and (B) correspond in Foster (1995) to data sets 2 and 3, respectively. To make the problem a bit trickier, a fourth sinusoid was added with period $\nu_1'^{-1} = 122.5$ days and amplitude 3, such that $\nu_1' - \nu_1 = 2\nu_{\text{year}}$ with $\nu_{\text{year}}^{-1} = 365$ days, so that the sidelobes caused by the annual gaps superimpose at $\nu_{\text{alias}} = (\nu_1 + \nu_1')/2$. White Gaussian noise with standard deviation $\sigma_\epsilon = 0.3$ was also added. The top and bottom of Fig. 3 show the corresponding time series, the spectral windows, and the Fourier spectra for data sets (A) and (B), respectively. Because of high sidelobes in the spectral windows, both Fourier spectra show many aliased peaks.

Results obtained for data set (A) using basic prewhitening[3] and CLEANEST[4] strategies are plotted in Fig. 4. Both methods fail because of the aliased frequency at ν_{alias}. The basic prewhitening retrieves the four true frequencies but also three wrong ones with a higher amplitude than ν_1, whereas CLEANEST is not able to retrieve ν_3. Figure 5 plots the *SparSpec* estimation with $P = 4000$ frequencies ranging from 0 to $f_{\max} = 0.02$ c/d for $\lambda = 10$, i.e., $\lambda/\lambda_{\max} \simeq 0.05$ (in this example $\lambda_{\max} = 206$ and $\lambda_{\min} = 2.06$). The only non-zero components are located at zero and around each of the four true frequencies, and a non-zero value is also obtained at ν_{alias}, but with very low amplitude. Note that the raw estimator suffers a loss in the amplitude estimation, which is inherent to the regularisation procedure: the minimiser of criterion (5) no longer minimises the least-square term for data fidelity. However, once the frequencies are correctly determined, posterior amplitude re-estimation is straightforward by least-square fitting the amplitudes of a multisine model such as (1) with *estimated* frequencies. As the true frequencies do not belong exactly to the reconstruction grid, a frequency may be estimated by two adjacent non-zero values in $\widehat{\boldsymbol{x}}$: if so, the estimated frequency can be obtained by averaging these two adjacent components, weighted by their corresponding amplitude. The final result is plotted in Fig. 5 bottom, which shows an accurate estimation for both frequency and amplitude parameters.

---

[3] The basic sequential prewhitening procedure was implemented as proposed by Roberts et al. (1987), with no CLEAN gain.

[4] The CLEANEST procedure was performed with the *Period04* package: `http://www.univie.ac.at/tops/Period04/`
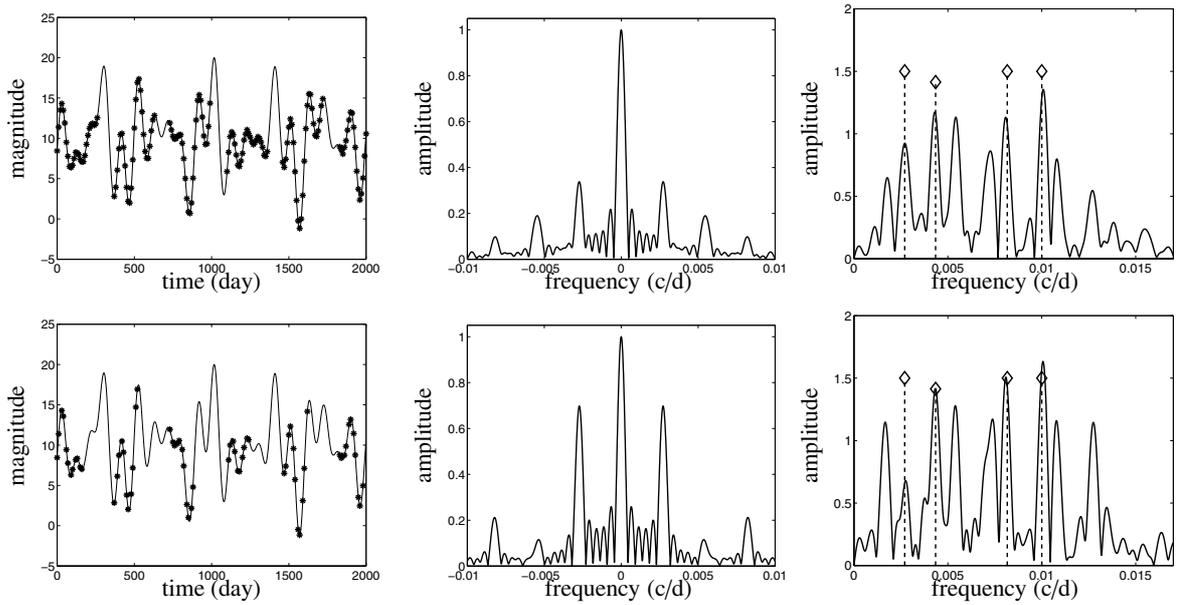
Fig. 3. Artificial data sets (A) (*top*) and (B) (*bottom*). *Left*: available data (circles) and full signal (solid line). *Centre*: spectral window. *Right*: Fourier spectrum (solid line) and true spectral lines (diamonds).
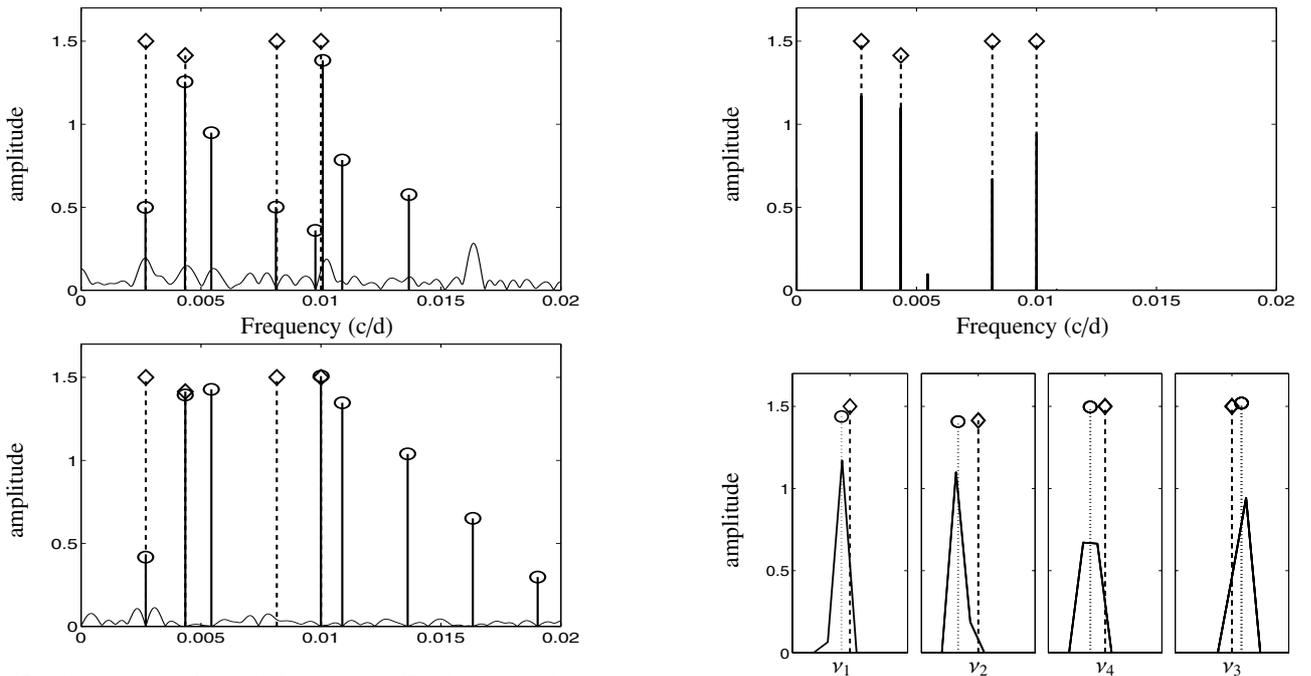

Fig. 4. Prewhitening results with data set (A). *Top*: basic prewhitening. *Bottom*: CLEANEST prewhitening. Circles correspond to the extracted components, diamonds to the true components, and the solid line to the residual spectrum.

Fig. 5. *SparSpec* results for data set (A). *Top*: estimator $|\widehat{x}_\lambda|$ for $\lambda/\lambda_{\max} = 0.05$ (solid line) and true lines (diamonds). *Bottom*: zoom around each true frequency. Width of every window is 4.10^{-5} c/d. The circles correspond to the posterior least-square estimation of amplitudes once frequencies have been determined.

Efficient tuning of the λ parameter is illustrated in Fig. 6. For $\lambda/\lambda_{\max} \in [0.04, 0.45]$, that is, in the corner of the *L*-curve plotted in Fig. 6 top, $\widehat{x}_\lambda$ locates the same non-zero components. Figure 6 bottom shows estimators $\widehat{x}_\lambda$ for λ values out of this range, which clearly characterise over-regularisation (left) and under-regularisation (right).

More ambiguous results are obtained on the more critical data set (B), as shown in Fig. 7: the frequency at $\nu_1 = 1/370$ c/d is not retrieved, but is replaced by two wrong components at higher frequencies. Figure 8 shows, however, that the solution given by *SparSpec* models the available data quite well, while the reconstruction obtained in the gapped periods is inaccurate. The lack of success here is caused by some physical indeterminacy. Although the sparsest solution to criterion (5) is unique, in this case other multisine functions exist that are modelling the
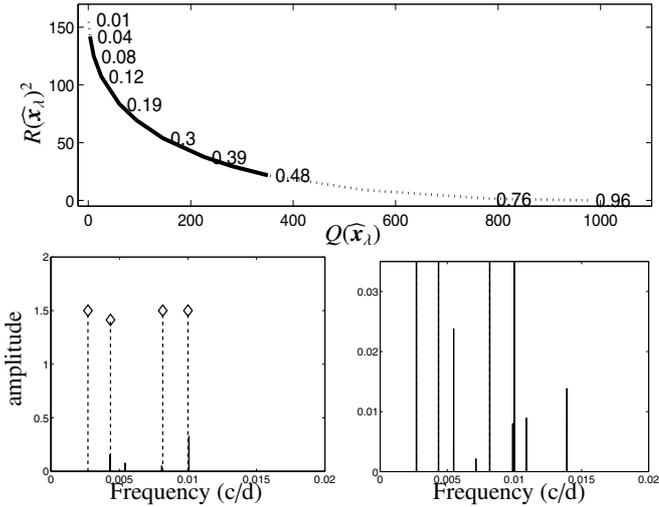
Fig. 6. Sensitivity of *SparSpec* to parameter λ. *Top*: *L*-curve $(Q(\widehat{x}_\lambda), R(\widehat{x}_\lambda)^2)$ parameterised by λ/λ_{max}. The solid line indicates acceptable solutions. *Bottom left*: over-regularisation ($\lambda/\lambda_{max} = 0.76$) and *bottom right*: under-regularisation ($\lambda/\lambda_{max} = 0.01$), with a zoom on the small amplitudes. $|\widehat{x}_\lambda|$ is plotted in solid line and the diamonds represent the true spectral lines.
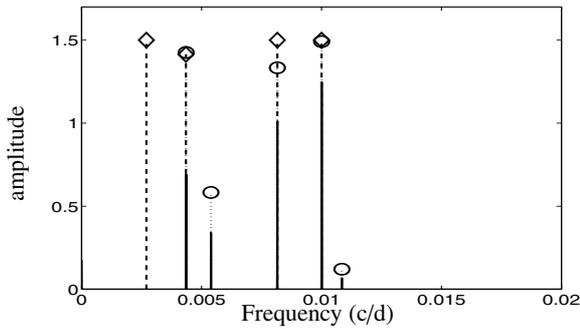
Fig. 7. *SparSpec* results for data set (B). Estimator $|\widehat{x}_\lambda|$ for $\lambda/\lambda_{max} = 0.05$ (solid line), amplitude posterior reestimation (circles), and true lines (diamonds).
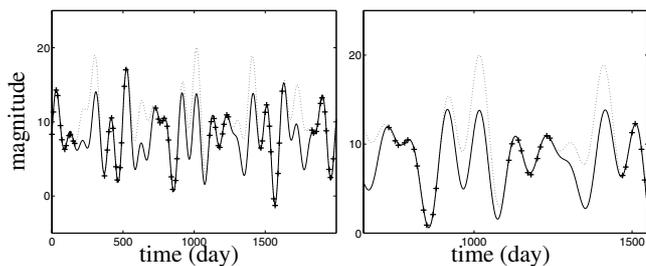
Fig. 8. Signal reconstruction from *SparSpec* estimation (solid line) for data set (B). *Left*: total time coverage. *Right*: zoom on a portion. The plus represent the available data and the dotted line corresponds to the true fully-sampled signal.

data (including of course the true one), and data do not contain enough information to distinguish those solutions.

6. Application to HD 104237 data

Figure 9 shows the radial velocity curve of the Herbig Ae star HD 104237 from Böhm et al. (2004). These data correspond to five observing nights of high resolution spectroscopy at SAAO (South Africa) during April 1999. The individual radial

velocities were determined by adjusting a Gaussian to the resolution-dominated spectral profiles and taking its centroid, more precisely by adjusting it to the least-square deconvolved profiles (LSD, Donati et al. 1997).

Variations are dominated by the low-frequency orbital movement caused by the system multiplicity and, as the corresponding Fourier spectrum shows, direct application of a prewhitening procedure to these data would lead to unsatisfactory results. Since HD 104237 was observed during a second observation campaign in April and May 2000, Böhm et al. (2004) were able to determine the orbital parameters with high accuracy by a binary approximation merging both data sets. Then, after removing the main part of the orbital component (see Fig. 9 right), application of a prewhitening method was possible, allowing these authors to detect five oscillations between 28.5 c/d and 35.6 c/d.

As time sampling is irregular, there is no more Nyquist limit as in the regular sampling case, and aliasing is pushed at much higher frequencies (Eyer & Bartholdi 1999). This offers the possibility of searching for oscillations in a much wider frequency range. Note that aliasing is caused by the periodicity of the spectral window $W(f) = \sum_n e^{j2\pi f t_n}/N$, for which no general properties can be established in the irregular sampling case, except hermitian symmetry. Thus, a plot of the spectral window for $f \in [-f_{max}, f_{max}]$ is necessary to ensure that the bandwidth $[0, f_{max}]$ is alias-free. The template should be free of any periodicity or pseudo-periodicity. For this sampling scheme, Fig. 10 (top) shows that the first pseudo-periodicity occurs at 240 c/d. A first try with f_{max} at 100 c/d was performed, where no spectral component was found above 40 c/d. In the following results, f_{max} was set to 40 c/d to reduce the computational cost. The grid size here is $K = 5000$ so that frequency precision is $f_{max}/K = 0.008$ c/d. A zoom on the central template of the spectral window is plotted in Fig. 10 (bottom), showing the high sidelobes caused by the daily gaps.

SparSpec was applied on both raw and corrected data for HD 104237 (Figs. 9a) and c), respectively, and very similar results were obtained for both data sets (except for low frequencies, of course). This shows the robustness of *SparSpec* toward low-frequency perturbations, which is a consequence of using model (3) with fixed frequencies over the whole grid.

Figure 11 shows estimator $|\widehat{x}_\lambda|$ obtained by *SparSpec* from the corrected data for $\lambda/\lambda_{max} = 0.08$. The lines detected by Böhm et al. (2004) using a CLEANEST algorithm after orbital correction are also plotted in this figure[5]. Note that a basic prewhitening strategy led to results that are similar to CLEANEST. Numerical values are reported in Table 1, where frequencies are arranged in their order of appearance in the sequential prewhitening procedure. Both estimations coincide very well for the four main frequencies. The fifth frequency detected by *SparSpec* at $\nu_5 = 34.6$ c/d differs quite significantly from the one detected at $\nu_5 = 33.862$ c/d by Böhm et al. (2004).

It is quite difficult to distinguish both fits, and one can even think of some physical indeterminacy. To this end, complementary simulations were performed. *SparSpec* and the sequential prewhitening method were run on two simulated data sets using the HD 104237 data time sampling and both sets of spectral components reported in Table 1. Figure 12 shows the corresponding four estimates: we can see that in both cases *SparSpec* correctly locates the five components, whereas the prewhitening method fails to retrieve the line at $\nu_5 = 34.6$ c/d for the simulated data corresponding to the *SparSpec* fit. This shows that a

---

[5] We only keep in this comparison the five frequencies to which a high confidence level was associated by Böhm et al. (2004).
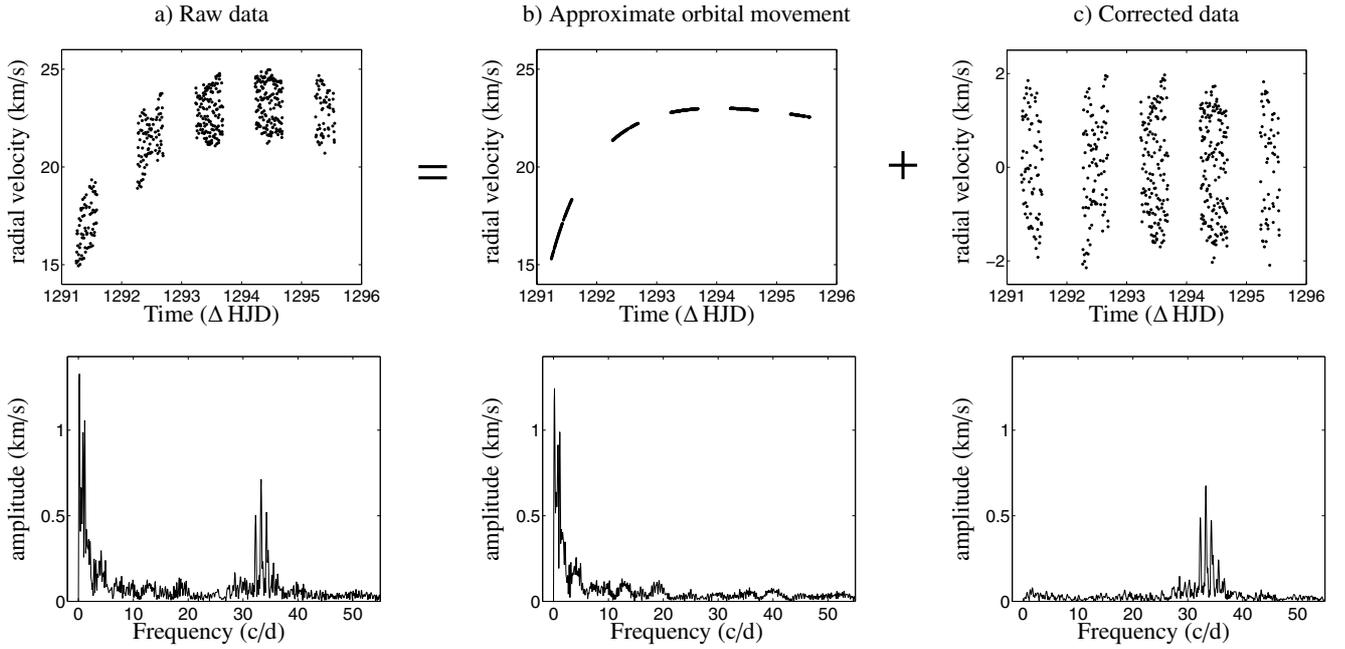
Fig. 9. HD 104237 data from Böhm et al. (2004): raw data, extracted orbital movement and corrected data. *Top*: radial velocity curves (km s^{-1}). Time is expressed in HJD, heliocentric Julian date, with HJD = 2 450 000 + Δ HJD. *Bottom*: moduli of their Fourier spectra.
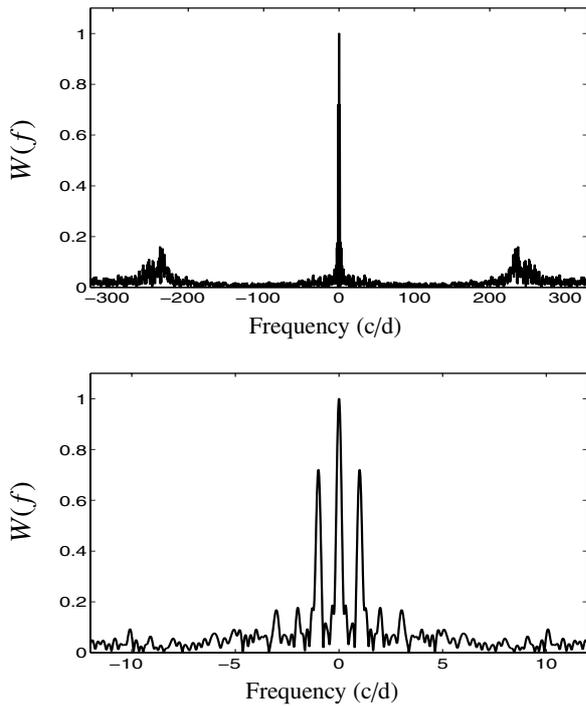
Fig. 10. Spectral window associated to the HD 104237 data in Fig. 9. *Top*: for $f \in [-300, 300]$ c/d. *Bottom*: for $f \in [-10, 10]$ c/d.
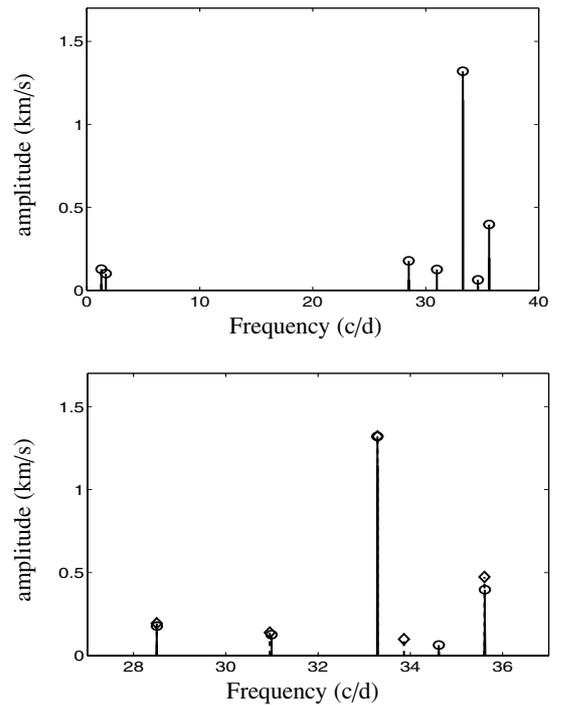
Fig. 11. *Top*: *SparSpec* estimator $|\widehat{x}_\lambda|$ obtained from HD 104237 corrected data for $\lambda/\lambda_{\max} = 0.08$ (circles). *Bottom*: zoom on the frequency range [27 c/d, 37 c/d]. The diamonds indicate former estimation results by Böhm et al. (2004).

prewhitening strategy is unable to detect the fifth frequency at 34.6 c/d if it was present in the data. In contrast, if a frequency at 33.862 c/d really existed in the data, then *SparSpec* should be able to retrieve it. As a conclusion, the frequencies identified from the *SparSpec* estimation – in particular, the detection of an oscillation frequency at $\nu_5 = 34.6$ c/d – is a more reliable result.
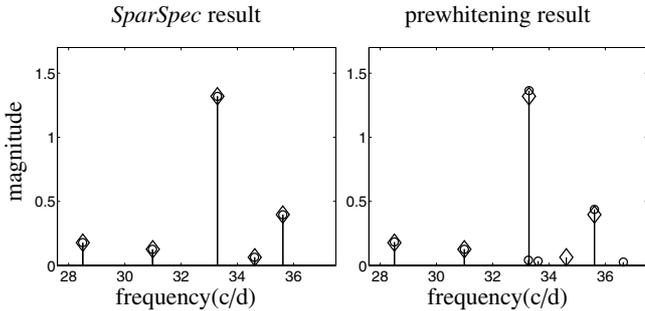
7. Conclusion

In this paper we presented an alternative approach to identifying multiple sinusoids. A linear model was proposed, where the spectrum is discretised on a fixed frequency grid with arbitrarily high frequency precision. Using the sparse representations framework, the *SparSpec* spectral estimator was defined as the
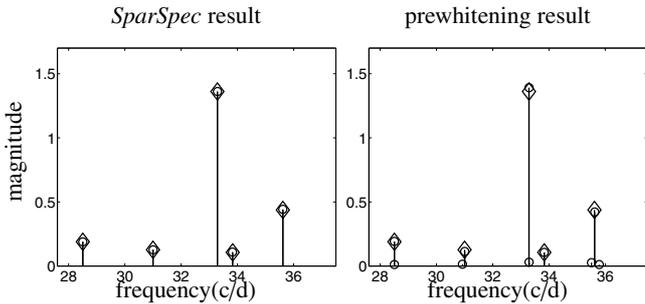
Table 1. Numerical comparison between the *SparSpec* results obtained from HD 104237 corrected data and results previously obtained by Böhm et al. (2004).

	SparSpec precision $\frac{f_{max}}{K} = 0.008$ c/d		CLEANEST (Böhm et al. 2004) precision 0.04 c/d	
	Frequency (c/d)	Amplitude (km s^{-1})	Frequency (c/d)	Amplitude (km s^{-1})
ν_1	33.288	1.320	33.289	1.320
ν_2	35.613	0.397	35.606	0.474
ν_3	28.504	0.178	28.503	0.195
ν_4	30.992	0.127	30.954	0.139
ν_5	34.616	0.064	33.862	0.099

Fig. 12. Comparison of *SparSpec* and prewhitening results using artificial data simulated from the *SparSpec* fit (*top*) and the previous CLEANEST fit (*bottom*) of HD 104237 data. Circles correspond to the estimated components and diamonds to the true values.

global minimiser of a convex criterion, for which an efficient algorithm was described to perform the estimation at a reasonable computational cost. Compared with classic sequential methods, the determination of the frequency locations with *SparSpec* is less ambiguous. Indeed, the criterion minimised by classic methods has numerous local minima in which optimisation algorithms can be trapped, leading to false frequency detections.

The proposed procedure was validated on simulations where spectral lines were correctly estimated, whereas results of prewhitening methods are less satisfactory because of strong sampling aliases. Finally, an application of *SparSpec* to the HD 104237 radial velocity data was presented and compared to former results by Böhm et al. (2004), obtained with a CLEANEST

procedure after orbital correction. While similar results were achieved for the four main frequencies, both methods locate a fifth frequency differently. Additional simulations showed that the *SparSpec* result was more plausible. To summarise, the main advantages of *SparSpec* consist in:

– more robustness than usual sequential prewhitening methods toward sampling artifacts;
– accurate estimation of both frequencies and amplitudes;
– insensitivity to low-frequency perturbations, e.g. to those caused by orbital movements.

As a conclusion, *SparSpec* may become an efficient alternative to prewhitening methods, especially in the case of ambiguous frequency location. Researchers are invited to test it intensively and send their comments to the authors[6] for further improvements.

References

Alliney, S., & Ruzinsky, S. A. 1994, IEEE Trans. Signal Processing, 42, 618
Barning, F. J. M. 1963, Bull. Astron. Inst. Netherlands, 17, 22
Bertsekas, D. 1995, Nonlinear Programming (Belmont, MA: Athena Scientific)
Böhm, T., Catala, C., Balona, L., & Carter, B. 2004, A&A, 427, 907
Bourguignon, S. 2005, PhD thesis, Université Paul Sabatier – Toulouse III, France
Bourguignon, S. 2006, Sur l'unicité du minimiseur d'un critère quadratique pénalisé par la norme ℓ^1, Tech. rep., Laboratoire d'Astrophysique de Toulouse-Tarbes
Breger, M., Stich, J., Garrido, R., et al. 1993, A&A, 271, 482
Ciuciu, P., Idier, J., & Giovannelli, J.-F. 2001, IEEE Trans. Signal Processing, 49, 2201
Demoment, G. 1989, IEEE Trans. Acoustics, Speech, and Signal Processing, 37, 2024
Donati, J.-F., Semel, M., Carter, B. D., et al. 1997, MNRAS, 291, 658
Donoho, D., Elad, M., & Temlyakov, V. 2006, IEEE Trans. Information Theory, 52, 6
Dworetsky, M. M. 1983, MNRAS, 203, 917
Eyer, L., & Bartholdi, P. 1999, A&AS, 135, 1
Ferraz-Mello, S. 1981, AJ, 86, 619
Foster, G. 1995, AJ, 109, 1889
Foster, G. 1996, AJ, 111, 541
Fuchs, J.-J. 2004, IEEE Trans. Inf. Theory, 50, 1341
Gray, D. F., & Desikachary, K. 1973, ApJ, 181, 523
Högbom, J. A. 1974, A&AS, 15, 417
Hansen, P. 1992, SIAM Review, 34, 561
Idier, J. (ed.) 2001, Approche bayésienne pour les problèmes inverses (Paris: Traité IC2, Série traitement du signal et de l'image, Hermès)
Kay, S. 1988, Modern Spectral Estimation (Englewood Cliffs: Prentice-Hall)
Kuschnig, R., Weiss, W. W., Gruber, R., Bely, P. Y., & Jenkner, H. 1997, A&A, 328, 544
Lafler, J., & Kinman, T. D. 1965, ApJS, 11, 216
Lomb, N. R. 1976, Ap&SS, 39, 447
Reegen, P. 2004, in IAU Symp., ed. J. Zverko, J. Ziznovsky, S. J. Adelman, & W. W. Weiss, 791–798
Roberts, D. H., Lehar, J., & Dreher, J. W. 1987, AJ, 93, 968
Sacchi, M., Ulrych, T., & Walker, C. 1998, IEEE Trans. Signal Processing, 46, 31
Scargle, J. D. 1982, ApJ, 263, 835
Stellingwerf, R. F. 1978, ApJ, 224, 953
Stoica, P., Moses, R. L., Friedlander, B., & Söderström, T. 1989, IEEE Trans. Acoustics, Speech and Signal Processing, 37, 378
Tikhonov, A., & Arsenin, V. 1977, Solutions of ill-posed problems (Washington, DC: V.H. Winston & Sons)

---

[6] Send e-mail to SparSpec@ast.obs-mip.fr