MNRAS **000**, 1–9 (2020) Preprint 23 April 2020 Compiled using MNRAS LATEX style file v3.0

Self-consistent method to extract non-linearities from pulsating stars light curves I. Combination frequencies.

Lares-Martiz ,M.,[1]★ Garrido, R.,[1] Pascual-Granado, J.[1]
[1]*Instituto de Astrofísica de Andalucía (IAA-CSIC). Glorieta de Astronomía s/n, 18008, Granada. Spain*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Combination frequencies are not solutions of the perturbed stellar structure equations. In dense power spectra from a light curve of a given multi-periodic pulsating star, they can compromise the mode identification in an asteroseismic analysis, hence they must be treated as spurious frequencies and conveniently removed. In this paper, a method based on fitting the set of frequencies that best describe a general non-linear model, like the Volterra series, is presented. The method allows to extract these frequencies from the power spectrum, so helping to improve the frequency analysis enabling hidden frequencies to emerge from the initially considered as noise. Moreover, the method yields frequencies with uncertainties several orders of magnitude smaller than the Rayleigh dispersion, usually taken as the present error in a standard frequency analysis. Furthermore, it is compatible with the classical counting cycles method, the so-called O-C method, which is valid only for mono-periodic stars. The method opens the possibility to characterise the non-linear behaviour of a given pulsating star by studying in detail the complex generalised transfer functions.

Key words: asteroseismology – stars: oscillations – stars: variables: Scuti

1 INTRODUCTION

Combination frequencies arise in the power spectra of some pulsating stars (δ Sct stars, γ Dor stars, Cepheids, β Cep stars, SPB stars, RR lyrae stars or white dwarfs) due to a non-sinusoidal shape of their light curves.

Such deviations may have physical explanations related to the non-linear response of the stellar interior to the oscillation wave, as well as for mechanisms that follow non-linear dependencies (e.g. the T^4 dependency in the Stefan-Boltzmann law for the emergent flux). Their consequences are non-linear interactions of the intrinsic pulsation modes of the star, generating cross-terms at summation, difference and harmonics of them.

These were analysed by different authors: Gillet & Fokin (2014) studied the propagation of a shockwave in RR Lyrae stars; or Brickhill (1992) for DA and DB white dwarfs, where he relates the non-sinusoidal light variations with the interchanges between the convective transport of energy and the radiative one that take place in the deeper layer of the hydrogen ionization zone and where a non-linear relation for the temperature gradient is used. Regardless, in the literature these are generally referred as *non-linear mixing processes* (Breger & Montgomery 2014) or gathered in what is called the *non-linear distortion model* (Degroote et al. 2009).

Combination frequencies have been useful for mode identification in the asteroseismic analysis of the ZZ Ceti stars (e.g. Montgomery 2005). However, Balona (2012) demonstrated that the theory of Brickhill, extended analytically in Wu (2001), is not valid for main-sequence pulsating stars, like δ Sct stars or β Cepheids. The reason is that geometrical effects (i.e. variation in radius and surface normal) have to be taken into account, whereas these effects are negligible in white dwarfs, because of their high density.

Therefore, since they can be a source of confusion in mode identification for main-sequence pulsating stars, they can be considered spurious frequencies and must be properly identified and removed from the power spectra.

Usually, these combination frequencies are identified as such, when in a prewhitening procedure the extracted frequency falls inside an ϵ range (usually the Rayleigh frequency resolution $\epsilon = 1/T$ being T the observation interval) around a previously calculated exact combination value (Zwintz et al. 2020; Saio et al. 2018; Murphy et al. 2013; Pápics 2012; Degroote et al. 2009). The inconsistency of this reasoning, is that when the frequency does not match exactly the value of the combination the residuals after the fit will be correlated with the so selected frequency.

The techniques for identification and extraction of combination frequencies explained in Kurtz et al. (2015), Balona et al. (2012) and Handler et al. (2006) are more sophisticated than the aforecited, but no mathematical frame for this procedure is given.

In this paper we present a novel self-consistent method for removing combination frequencies based on an unbiased estimation

★ E-mail: mlm@iaa.es

of the non-linear solution aimed at yielding residuals uncorrelated with the frequencies generating these non-linearities.

Here, we are continuing the work of Garrido & Rodriguez (1996) where, for the first time, the Volterra expansion is proposed as an option to model the non-linearities present in some pulsating stars.

An advantage of applying this method is that when the combination frequencies have been correctly removed from the time series, a new collection of frequencies can emerge from power spectra, which were previously hidden. In addition, it can give frequency errors much more realistic than the coarse estimation given by the Rayleigh frequency resolution.

The paper is organised as follows: in section 2 a theoretical basis is explained. In section 3 a detailed description of the methodology is presented. The results of applying this technique to different light curves are shown in section 4 followed by a discussion of these results in section 5. Finally, section 6 summarises the conclusions drawn from this study and future work.

2 THEORETICAL BASIS: GENERAL RESPONSE OF A NON-LINEAR SYSTEM

Non-linear systems can be studied and represented by the Volterra series. In the particular case where the input to a non-linear system is a single real-valued sine wave at ω_0 with amplitude A_0, the output can be expressed (see p.29 in Priestley 1988) by:

$$Y(t) = A_0 \cdot \Gamma_1(\omega_0) \cdot e^{i\cdot\omega_0\cdot t+\phi_0} + A_0^2 \cdot \Gamma_2(\omega_0,\omega_0) \cdot e^{2\cdot i\cdot\omega_0\cdot t+2\cdot\phi_0} + A_0^3 \cdot \Gamma_3(\omega_0,\omega_0,\omega_0) \cdot e^{3\cdot i\cdot\omega_0\cdot t+3\cdot\phi_0} + \ldots \quad (1)$$

Where the Γ_j functions are the generalised transfer complex functions. Sub-indexes denote the non-linear order of interactions, i.e. Γ_1 represents the system response for each independent frequency, Γ_2 represents the system response for first order interactions, and so on.

The generalised transfer functions Γ_j can be written as Volterra expansion. These series, although they retain some of the system memory, are non-orthogonal.

To disentangle the correlation between the basis functionals of the Volterra Series, Wiener orthogonal expansion (Wiener 1958) can be used. The ulterior intention is to obtain a physical model that explains non-linearities.

Current efforts are being made towards the aforementioned objective and it will be the aim of future work, but in this paper orthogonality of expression (1) is assumed, enabling to study certain properties that can be observed in a first order approximation.

Although Y(t) is not linear between the input/output spectra, it is linear between Y(t) and the Γ_j (Scargle 2020), so a standard Least-Square procedure will quantify, not only the parameters of the input components, but also the contribution from the generalised transfer functions.

For example, since the Γ_j functions are complex functions, Eq.1 can be rearranged as:

$$Y(t) = \tilde{A_1} \cdot e^{i\,\omega_0\cdot t+\tilde{\phi_1}} + \tilde{A_2} \cdot e^{2\cdot i\cdot\omega_0\cdot t+\tilde{\phi_2}} + \tilde{A_3} \cdot e^{3\cdot i\cdot\omega_0\cdot t+\tilde{\phi_3}} + \ldots \quad (2)$$

Where,

$$\begin{aligned} \tilde{A_1} &= A_0 \cdot |\Gamma_1(\omega_0)|, \\ \tilde{A_2} &= A_0^2 \cdot |\Gamma_2(\omega_0,\omega_0)|, \\ \tilde{A_3} &= A_0^3 \cdot |\Gamma_3(\omega_0,\omega_0,\omega_0)| \end{aligned} \quad (3)$$

And

$$\begin{aligned} \tilde{\phi_1} &= \phi_0 + arg\{\Gamma_1(\omega_0)\}, \\ \tilde{\phi_2} &= 2\cdot\phi_0 + arg\{\Gamma_2(\omega_0,\omega_0)\}, \\ \tilde{\phi_3} &= 3\cdot\phi_0 + arg\{\Gamma_3(\omega_0,\omega_0,\omega_0)\} \end{aligned} \quad (4)$$

Examining the case when the input is composed of two real-valued sine waves at frequencies ω_0 and ω_1, the output of a non-linear system modeled by a Volterra series is:

$$\begin{aligned} Y(t) = A_0 \cdot \Gamma_1(\omega_0) \cdot e^{i\cdot\omega_0\cdot t+\phi_0} + A_1 \cdot \Gamma_1(\omega_1) \cdot e^{i\cdot\omega_1\cdot t+\phi_1} \\ +A_0^2 \cdot \Gamma_2(\omega_0,\omega_0) \cdot e^{2\cdot i\cdot\omega_0\cdot t+2\cdot\phi_0} + A_1^2 \cdot \Gamma_2(\omega_1,\omega_1) \cdot e^{2\cdot i\cdot\omega_1\cdot t+2\cdot\phi_1} \\ +A_0 \cdot A_1 \cdot \Gamma_2(\omega_0,\pm\omega_1) \cdot e^{i\cdot(\omega_0\pm\omega_1)\cdot t+(\phi_0\pm\phi_1)} \\ +A_1 \cdot A_0 \cdot \Gamma_2(\omega_1,\pm\omega_0) \cdot e^{i\cdot(\omega_1\pm\omega_0)\cdot t+(\phi_1\pm\phi_0)} + \ldots \end{aligned} \quad (5)$$

Eq.5 can be rearranged as:

$$\begin{aligned} Y(t) = \tilde{A_1} \cdot e^{i\cdot\omega_0\cdot t+\tilde{\phi_1}} + \tilde{A_2} \cdot e^{i\cdot\omega_1\cdot t+\tilde{\phi_2}} + \\ \tilde{A_3} \cdot e^{2\cdot i\cdot\omega_0\cdot t+\tilde{\phi_3}} + \tilde{A_4} \cdot e^{2\cdot i\cdot\omega_1\cdot t+\tilde{\phi_4}} + \\ \tilde{A_5} \cdot e^{i\cdot(\omega_0\pm\omega_1)\cdot t\pm\tilde{\phi_5}} + \ldots \end{aligned} \quad (6)$$

where,

$$\begin{aligned} \tilde{A_1} &= A_0|\Gamma_1(\omega_0)|, & \tilde{A_4} &= A_1^2|\Gamma_2(\omega_1,\omega_1)|, \\ \tilde{A_2} &= A_1|\Gamma_1(\omega_1|, & \tilde{A_5} &= A_0A_1|\Gamma_2(\omega_0,\omega_1)|, \\ \tilde{A_3} &= A_0^2|\Gamma_2(\omega_0,\omega_0)| & or\ \tilde{A_5} &= A_1A_0|\Gamma_2(\omega_1,\omega_0)|\ ^{*1} \end{aligned} \quad (7)$$

And

$$\begin{aligned} \tilde{\phi_1} &= \phi_0 + arg\{\Gamma_1(\omega_0)\}, & \tilde{\phi_4} &= 2\cdot\phi_1 + arg\{\Gamma_2(\omega_1,\omega_1)\}, \\ \tilde{\phi_2} &= \phi_1 + arg\{\Gamma_1(\omega_1)\}, & \tilde{\phi_5} &= \phi_0 \pm \phi_1 + arg\{\Gamma_2(\omega_0,\pm\omega_1)\}, \\ \tilde{\phi_3} &= 2\cdot\phi_0 + arg\{\Gamma_2(\omega_0,\omega_0)\}, & or\ \tilde{\phi_5} &= \phi_1 \pm \phi_0 + arg\{\Gamma_2(\omega_1,\pm\omega_0)\}^{*1} \end{aligned} \quad (8)$$

^{*1}: since no condition of symmetry is yet imposed.

These amplitude and phase relations are considerably different to the ones suggested by the so called *simple model* in Breger & Montgomery (2014). Analysis of amplitude and phase relations affected by the Γ_j functions (Eqs. 3,7 and 4,8 for the mono-periodic and double mode cases respectively) are going to be studied in the second part of this publication.

Notice that the generalised transfer functions will not influence the frequencies, topic of this first part of the overall study. What is relevant of this formulation for the present paper is that this output could describe the rapid increases and slow decreases of the luminosity observed in some light curves.

A Fourier Transform of Eq.6 would result in a power spectrum with peaks at the independent frequencies ω_0 and ω_1 (called parent frequencies in this paper), as well as on the frequencies of the non-linear terms, which are the combination frequencies (called children in this paper). This is a well-known phenomenon in system and signal analysis called *intermodulation distortion* (in the case of

one parent frequency is called *frequency multiplication* or *harmonic distortion*).

Modelling non-linearities by treating the variable star as a non-linear but stationary system is the mathematical foundation for the method here presented, assuming a known input equal to a basic harmonic signal.

3 METHODOLOGY: THE 'BEST' PARENTS

Under the hypothesis that Eqs.2 and 6 are approximation functions modelling the non-linearities present in the light curves of a mono-periodic and double-mode variable stars respectively, a least-square fit of them should yield residuals uncorrelated with the combination frequencies.

The variance of such residuals (in comparison with the variance of the original light curve) quantifies to which extend the parents and children fitted explain the signal as non-linearities: the lower the variance value the better the fit.

We define V as the continuous function of the variance related to different fittings of parent frequencies (and their statistically significant combinations) to a given light curve. Therefore, V is only a function of the parent frequencies. The aim is to find a minimum in V.

There are several procedures for finding the minimum in a non-linear least-square fitting (when the fitted functions are non-linear in the parameters) (Bevington & Robinson 2003), but in the method here presented we follow an empirical approach which consists in exploring the n-dimensional independent frequencies surface.

In the case of a single parent frequency (e.g. a Cepheid or a high amplitude mono-periodic variable), we select the highest peak in the power spectrum as a first approximation to the parent frequency with a coarse precision. Then, V(ω) will be sampled with a step progressively smaller until the minimum is reached.

In this way, the minimum V value will give us a much more precise frequency. Often, the frequency error will be smaller than the nominal 1/T Rayleigh dispersion. As we will see in Section 4.1, the frequency value will be compatible with that obtained by the O-C method (Sterken 2005), which is basically how the times of arrival of the pulses are determined in the Pulsar Timing analysis technique, known for yielding extraordinary precision when calculating the period of a pulsar (Lorimer & Kramer 2004).

Now that we know the 'best' parent frequency, in the case of a single parent mode, we calculate the set of combinations (in this case, harmonics) not exceeding the Nyquist frequency, to test whether they are statistically significant by giving any statistical test (e.g. Student's t or Snedecor's-F). It is important to highlight that the applicability of this method is not just for evenly-spaced data. When dealing with unequally-spaced data, the set of combinations to calculate and fit can be accordingly chosen up to any convenient frequency.

In the case of two parent frequencies, e.g. double-mode Cepheids, High Amplitude δ Sct stars (HADS) or RRLyrae, or multiperiodic stars such as Low Amplitude δ Sct stars (LADS), γ Dor, etc.) the procedure is the same for finding the 'best' parents but for calculating all the combination values we use the form:

$$\omega_k = \left|\pm n \cdot \omega_i \pm m \cdot \omega_j\right| \tag{9}$$

Where m and n are integer numbers under the condition that ω_k don't go beyond the Nyquist frequency. The absolute sum of n and m defines the combination order. We choose to limit equation (9) to a two-termed expression to avoid increasing the probabilities for false identifications.

A fit to the light curve of all the calculated combinations and their parents is computed. In this paper, a Student's t criterion of significance with a level of confidence of 99.9% is used.

Finally, a Fast Fourier Transform (FFT) of the residual light curve will result in the power spectrum free of combination frequencies which is the aim of this work.

In summary, the method consists essentially in two steps: finding the parent frequencies that best describe the signal's non-linearities (we call this first step the 'best' parent method, BPM for now on), and then fitting them together with the set of children frequencies originated by them.

4 RESULTS

Although this non-linearity analysis applies to every type of pulsator with combination frequencies, we focused in δ Sct stars due to their dense and unexplained power spectra (Poretti et al. 2009; Mantegazza et al. 2012; García Hernández et al. 2013).

To test its performance, three δ Sct stars, of different pulsational content, were chosen: First, a mono-periodic δ Sct variable, in order to verify the process of finding the 'best' parent. Then, a double-mode pulsator, and a multi-periodic δ Sct stars, where more complex non-linearities can be present and where their extraction could be critical to the frequency analysis.

Details of the physical parameters and relevant information describing the light curves are listed in Table 8, located in the Appendix 7.2.

It is known that gaps in time series are a source of uncertainties and error in harmonic analysis, see for instance Pascual-Granado et al. (2018). Therefore, gaps in the light curves used in this study where filled with MIARMA algorithm (Pascual-Granado et al. 2015), which remove the effects of the observational window while preserving the frequency content of the star. When the gap is sufficiently small (e.g. in the KIC 5950759 Kepler light curve only a few points are missing), then a linear interpolation was enough.

4.1 Mono-periodic Stars: The δ Sct case

The method applied to a light curve of a mono-periodic pulsating star results in a local minimum of the V function, which determines the fundamental period. This is shown for the mono-periodic δ Sct star TIC 9632550 (see Fig. 1, upper panel), observed by The Transiting Exoplanet Survey Satellite (TESS, Ricker et al. 2014).

The iterative process of searching for the 'best' parent explained in Section 3, refines the frequency until the variance value (V) does not change. In this particular case, it happened in the 14^{th} iteration (see Table 1). However, the list is truncated at the 5^{th} iteration due to the numerical uncertainty (see Appendix 7.1 for a full explanation).

Our result for the fundamental frequency is compatible with the one obtained by the O-C procedure (see Figure 2):

$$\omega_0 = (5.05496 \pm 0.00002)\ \text{c/d}.$$

The FFT of the residual light curve after fitting the fundamental frequency and its harmonics is almost cleaned from all signal contribution. This is graphically represented in the lower panel of Figure 1 and quantitatively expressed by the high percentage of the original power explained by these non-linear effects and their parents (expressed as %CF in Table 2).

Table 1. The 'best' parent search tree for the mono-periodic δ Sct star TIC 9632550. First column quantifies the number of statistically significant frequencies, or children, detected with the parent frequency specified in the third column, in cycles per days units (zeros omitted for the sake of clarity). Second column is the variance after fitting the parent and combination frequencies (in this case, only harmonics of the highest one).

N of fitted frequencies	V value	Frequency[c/d]
1	3155.844405793707201	5.0
5	948.937738175725485	5.05
14	33.629889361388315	5.055
14	33.629889361388315	5.055
14	**32.871889584474623**	**5.05496**
14	32.862102488776905	5.054964
14	32.862102488776905	5.054964
14	32.862101660828912	5.05496404
14	32.862101656257245	5.054964037
14	32.862101656225789	5.0549640372
14	32.862101656224368	5.05496403722
14	32.862101656224311	5.054964037229
14	32.862101656223409	5.0549640372274
14	32.862101656222627	5.05496403722726
14	32.862101656222627	5.05496403722726

Table 2. Results of the combination frequencies extraction process for the mono-periodic δ Sct star TIC 9632550. First column show the 'best' parent from the search tree in cycles per day. Second column specifies the number of statistically significant frequencies, or children, extracted. The %CF (third column) quantifies the percentage of initial power due to the combination frequencies and their parents.

TIC 9632550			
Tag	'best' parent [c/d]	Combinations extracted	%CF
f0	5.05496	13 Harmonics	98.98

There are three peaks remaining in the residual power spectrum: the first one correspond to the first frequency bin and it is the residual of a second order polynomial fitting performed in order to remove any trend in the light curve; the second peak, appearing next to the fundamental frequency, and the third one, which is next to the first harmonic, are possibly explained by an amplitude modulation of the fundamental frequency. An alternative explanation of these two peaks appearing at about the same frequencies as the fitted ones, might be the fractal property studied by De Franciscis et al. (2018), which is impossible to reproduce by using a Fourier representation. In any case, the logarithmic scale shows 5 orders of magnitude difference between the original and the residual power, which is in very good agreement with our expectation of uncorrelated residuals.

4.2 Double mode stars: The HADS case

We will now extend the procedure described in the previous section for the 'best' parent search to double mode stars.

There is a slight tendency for HADS stars to have higher number of non-linearities (Balona 2016). This is why we chose to apply the method to KIC 5950759, a HADS star observed by the Kepler satellite (Gilliland et al. 2010).

The original power spectrum (blue in Figure 3) shows a very regular structure where the first two highest peaks follow the fundamental period and first overtone ratio expected to occur for δ Sct stars (see Stellingwerf 1979).

The 'best' parents (see Table 3) are compatible with the ones

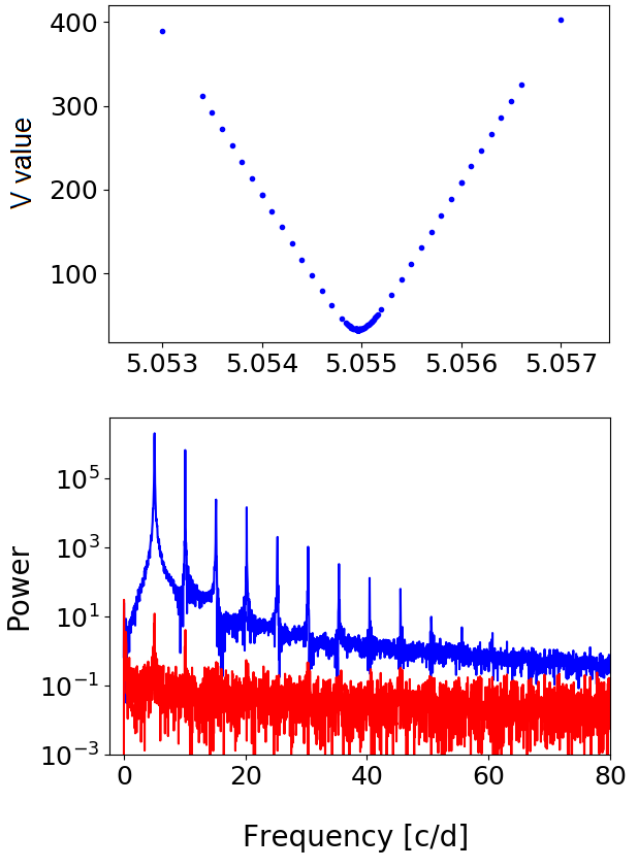

Figure 1. Upper panel: Fundamental period found as a local minimum at 5.054963644 c/d. Lower panel: in blue, the FFT of the original light curve. In Red, the FFT of the residuals after fitting the fundamental frequency and 13 statistically significant harmonics.

Table 3. Results of the combination frequencies extraction process for the double-mode HADS star KIC 5950759. First column show the 'best' parents from the search tree in cycles per day. Second column specifies the number of statistically significant frequencies, or children, extracted. The %CF (third column) quantifies the percentage of initial power due to the combination frequencies and their parents.

KIC 5950759			
Tag	'best' parents [c/d]	Combinations extracted	%CF
f0 f1	14.22136 18.33722	177 in total: 17 harmonics 92 sums and 68 differences	97.48

presented by Yang et al. (2018):

$$\omega_0 = (14.221367 \pm 0.000015)\ c/d$$
$$\omega_1 = (18.337228 \pm 0.000023)\ c/d$$

These frequency errors were calculated according to a heuristically derived formula for the upper limit of the frequency uncertainty in Kallinger et al. (2008). The precision reached with the 'best' parents search is also empirically justified in Appendix 7.1.

Due to the high power of the initial components, the fitting was completed in three steps, continuing the extraction in the residual light curve. After the fitting, almost every contribution from

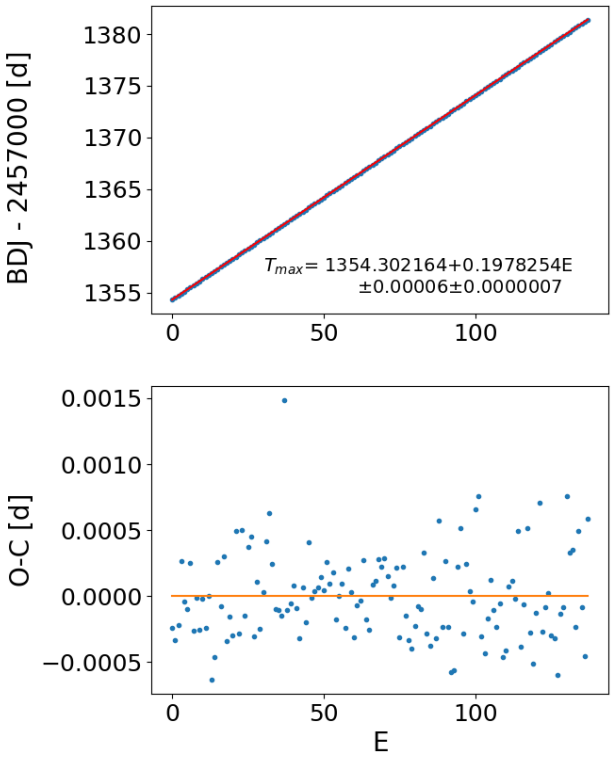

Figure 2. Upper panel: in blue, the times of the light maximum, corresponding to the maximum value of a parabola fitted to each cycle. Red: Regression line $T_{max} = T_0 + PE$, where P is the trial period, T_0 is the zero epoch, and E an integer number of cycles elapsed since the zero epoch. Fundamental frequency ($\omega_0 = 1/P$) calculated by the O-C method: $\omega_0 = (5.05496 \pm 0.00002)\ c/d$. Lower panel: residuals of the regression.

combination frequencies where effectively extracted and, as a consequence, a new frequency structure (that was previously hidden) emerges in its power spectrum (see middle panel of Fig. 3).

The frequency $\omega_m \approx 0.32$ c/d , first claimed by Yang et al. (2018) to be modulating the entire spectrum, is now significant according to the Reegen (2007) criteria of signal/noise > 12.57 in the power domain.

The new frequency structure seen for KIC 5950759 deserves further studies regarding its origin. Yang et al. (2018) explored several physical explanations, proposing as the most likely to identify ω_m as the rotation of the star, meaning that the dashed lines around ω_0 in the middle panel of Figure 3 corresponds to the modulation of the main pulsation modes with rotation frequency. But the high power for the frequency ω_0+ω_m suggests that it could still be an independent mode, resonantly coupled with the combination coming from the modulation effect. This possibility is going to be analysed in upcoming studies.

In any case, the results shown in Figure 3 provide a new level of relevance to this method: with a correct extraction of the combination frequencies, like the one proposed in this work, it is possible to unveil frequency structures that previously did not exceed the detection threshold. In this particular case, we were able to detect the new structure using short cadence (SC) data, while in Yang et al. (2018) long cadence (LC) data were used and a super-Nyquist and alias analysis were necessary in order to identify it.

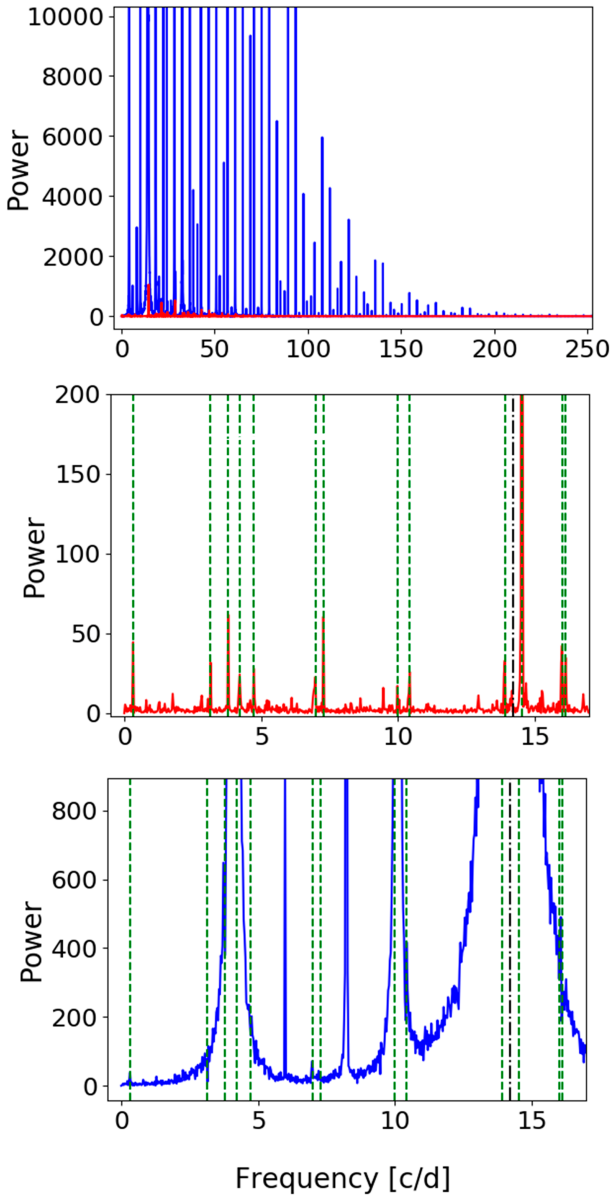

Figure 3. Blue: FFT of the original light curve. Red: FFT of the residuals after fitting the 'best' parents and the combinations generated by them. Green dashed: new frequencies detected in the residuals of the fitting. Black dash-dotted: 'best' parent frequencies. Notice in middle panel the significant peak corresponding to $\omega_m \approx 0.32$ c/d

4.3 Multi-periodic stars: The LADS case.

For more than two parents, exploring the frequency space of the V function recursively to find its minimum value can be computationally expensive, but the implications of applying the BPM can be crucial for an asteroseismic analysis, as we showed in the previous section.

We applied the method to the light curve of HD 174966 (see Fig. 4), a LADS (or simply δ Sct star) observed by the CoRoT satellite (Auvergne et al. 2009). This star was studied in García Hernández et al. (2013), who found that the 5th highest peak in the amplitude spectrum was very near to the estimated fundamental radial mode (17.3 ± 2.5 c/d).

Table 4. BPM for every possible couple of the first five high power peaks in the HD 174966 power spectrum

Couple Tag	Couple Frequencies [c/d]
(f0, f1)	(17.62288, 23.19479)
(f0, f2)	(17.62298, 21.42079)
(f0, f3)	(17.62280, 26.95853)
(f0, f4)	(17.62291, 27.71456)
(f1, f2)	(23.19477, 21.42097)
(f1, f3)	(23.19479, 26.95851)
(f1, f4)	(23.19477, 27.71503)
(f2, f3)	(21.42078, 26.95853)
(f2, f4)	(21.42078, 27.71463)
(f3, f4)	(26.95851 , 27.71487)

Table 5. Results of the combination frequencies extraction process for the multi-mode δ Sct star HD 174966. First column show the 'best' parents from the search tree in cycles per day. Second column specifies the number of statistically significant frequencies, or children, extracted. The %CF (third column) quantifies the percentage of initial power due to the combination frequencies and their parents.

HD 174966			
Tag	'best' parents [c/d]	Combinations extracted	%CF
f0	17.6230	118 in total:	93.01
f1	23.1948	1 harmonic	
f2	21.4210	11 sums	
f3	26.9585	106 differences	
f4	27.7150		

Therefore, we choose as independent frequencies the first five peaks of highest power but, instead of searching the combinations of the full set of five independent frequencies, we decided to divide the search in couples in order to reduce the computational cost of the BPM solution.

In Table 4, the results of the BPM for every possible couple with the precision adopted in this work (see 7.1) show small differences. In this case, we select as the 'best' parents for the 5 highest amplitude peaks the values with the best precision in which they are compatible in each search.

Particularly in this star, 118 combinations were statistically significant (see Table 5). Notice that the number of differences is higher than the number of sum combinations. This may raise concerns about the reality of these identifications in the lower frequency range.

The majority of these differences correspond to high order combinations (see Table 11), but harmonics ($n \cdot \omega_i$ or $m \cdot \omega_j$) with such high n and m, are not statistically significant (only 2f3 is detected). Alternatively, these significant differences could simply be false identifications due to the fact that at higher n and m more combinations are tested, increasing the probability of a match, as well as the possibility of choosing as parent frequencies a combination frequency.

In this work, we do not exclude any match since we are interested in finding the set of combination frequencies that could explain the most of the signal as non-linearities. Besides, Kurtz et al. (2015) states that the amplitude of a child frequency could be higher than their parents amplitudes, which could explain the missing high order harmonics issue.

As previously mentioned, further discriminating criteria (apart from their frequency value) are required for an unambiguous identification, in this way avoiding false identifications due to high values of n and m.

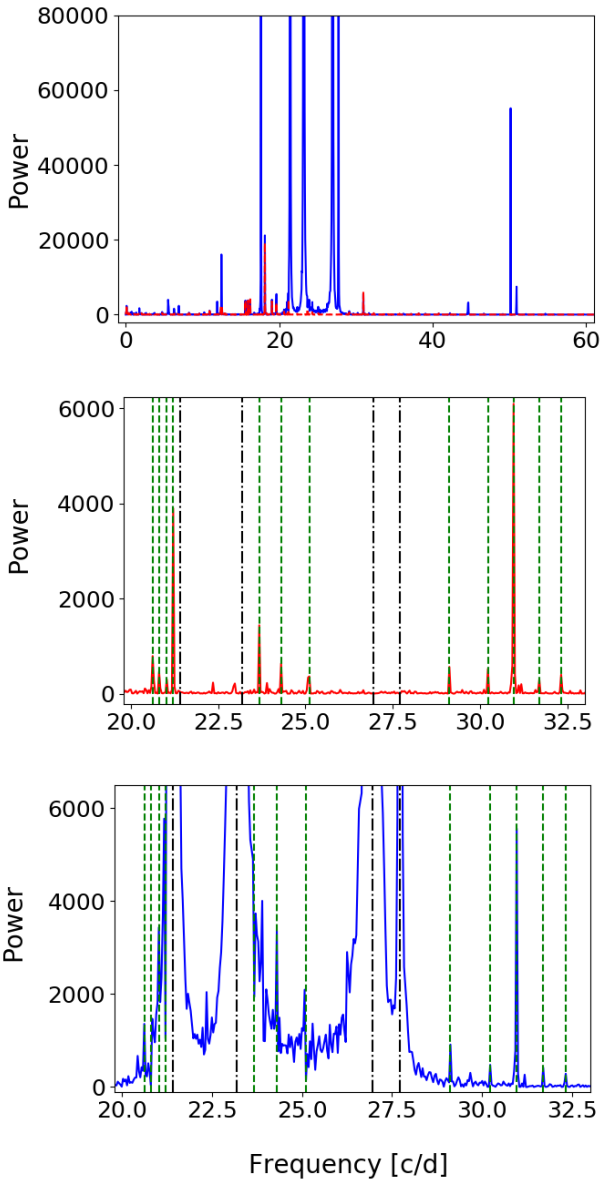

Figure 4. Blue: FFT of the original light curve. Red: FFT of the residuals after fitting the 'best' parents and the series of combinations originated by them. Green dashed: new frequencies detected after the fit. Black dash-dotted: 'best' parent frequencies.

Nevertheless, this example shows that in no case the method is introducing new frequencies and that even when is not clear that the arbitrarily chosen as parent frequencies are actual oscillation modes of the star (which is often the case when dealing with multi-periodic star), the set of significant combinations resulting from the algorithm can still be useful to test if extracting them has simplified the power spectrum in agreement with a solution from a linear stellar oscillation model (see Fig. 4, where some of the green dashed lines are equally spaced, possibly identifiable with non-radial frequency structures or rotational splittings).

5 DISCUSSION : FREQUENCIES INSIDE THE RAYLEIGH INTERVAL AND BIASED LEAST-SQUARE SOLUTIONS.

In this paper, we focus in discriminating combination frequencies from oscillation modes of the star only by the frequency relationship between parents and children (see Eq.9). However, pulsation modes could still have the same frequency value as a combination by chance, being this issue the main limitation of this method. Other considerations regarding the phases and amplitudes of the non-linearities, described by the generalised transfer functions Γ_j, will be examined in upcoming studies, looking for an unambiguous identification of a non-linearity.

Nevertheless, this study revealed the necessity of revising some common misconceptions in the field in order to obtain well-defined results.

First of all, we consider spurious peaks as significant and non-significant maxima in the frequency spectrum that do not correspond to any oscillation mode of the pulsating star(Suárez et al. 2020). In this sense, peaks associated to non-linear effects should not be considered spurious since they are originated by oscillations themselves. However, no closed-form formula like, for instance, Eqs. 1 and 5, have been obtained so far to characterise these non-linear effects. Therefore, until such expression is developed, we can consider non-linear components appearing in the frequency spectrum as spurious peaks.

In practice, the Rayleigh frequency resolution is often used as the error for the identification of combination frequencies. Frequencies extracted from a prewhitening cascade are found to be a combination when they fall inside the Rayleigh frequency range. However, after some numerical exercise performed for this particular case, we have found that fitting different sets of combination frequencies inside the Rayleigh interval, showed significantly different residuals. Therefore, Rayleigh frequency dispersion cannot be considered as a real frequency error.

The Rayleigh frequency resolution becomes a good uncertainty estimator when dealing with frequencies closer than $1/T$ to each other, but it could be exceeded when there is only one frequency component inside the Rayleigh interval. Resolution is not the same as precision.

On the other hand, extracting combination frequencies in a least-square sense as a first step before undertaking the frequency analysis can expose pulsation modes or frequency spacing patterns in which we are interested on, otherwise hidden. We have verified this in the results of the double-mode HADS, where the modulating frequency ω_m could be detected.

Finally, it is important to note that minimising the residuals does not guarantee that a real solution has been found. A least-square fit, exploring all free parameters (frequencies, amplitudes and phases) without preserving that the combinations are integer times the parent frequencies, can result in smaller residuals, but these cannot be explained with a closed-form formula.

6 CONCLUSIONS AND FUTURE WORK

Fundamental frequencies of mono-periodic stars analysed by this method are compatible with those calculated with the commonly used O-C method. The extension to stars pulsating in two or more parent frequencies is envisaged in this manuscript.

Although we support a conservative approach regarding the numerical precision (of the order of $\approx 10^{-5} d^{-1}$, see Sec.7.1), the method yields more realistic frequency uncertainties than the usual Rayleigh dispersion, a rough estimator of the frequency error.

Moreover, fitting combination frequencies with this method could help to identify pulsation modes by possibly revealing radial and non-radial frequency patterns or rotational splittings in the periodogram of the residuals. An example of this was shown in the results for the SC observations of the HADS star KIC 5950759, where the modulating frequency ω_m found by Yang et al. (2018) became detectable.

In the light of these conclusions, some common misconceptions when analysing non-linearities should be reconsidered: the Rayleigh dispersion is not necessarily the inaccuracy associated to the determination of combination frequencies ω_k (9).

This method partially overcomes the necessity to correct non-linearities in order to achieve an efficient mode identification with a linear pulsation model. Future research could explore the link between the theoretical foundations of this paper and the new MESA functionality regarding non-linear radial stellar pulsations (Paxton et al. 2019).

In essence, we can now identify the set of combination frequencies that best describe the signal, but the method do not supply a physical mechanism to explain their visibilities. There is still a possibility for the combination frequency to not correspond to an interaction between modes due to their intrinsic non-linear behaviour. It could be a new unstable mode (very near the exact combination frequency value) that undergoes amplitude enhancement due to the resonance mode coupling mechanism (Dziembowski & Krolikowska 1985; Dziembowski et al. 1988). This is an important aspect to be determined since it could clarify why stars with similar parameters exhibit very different power spectra. (Balona & Dziembowski 2011). Consequently, extra criteria for an unambiguous identification constitute the aim of the second part of this paper.

This method is shown to have a very high potential in exploring combination frequencies as the output of a non-linear system. The underlying physical meaning of the generalised transfer functions Γ_j defined in Eqs. (1) and (5) is the next step of this research. Future studies will consider orthogonal expansion in terms of a Wiener series opening a new insight on the relationships between amplitudes and phases as first envisioned by Garrido & Rodriguez (1996). The final objective is to properly characterise the non-linear response of a pulsating star.

ACKNOWLEDGEMENTS

MLM, RGH and JPG acknowledge financial support from the State Agency for Research of the Spanish MCIU through the "Center of Excellence Severo Ochoa" award to the Instituto de Astrofísica de Andalucía(SEV-2017-0709). Authors acknowledge funding support from Spanish public funds for research under project ESP2017-87676-C5-5-R.

REFERENCES

Auvergne M., et al., 2009, Astronomy & Astrophysics, 506, 411
Balona L. A., 2012, Monthly Notices of the Royal Astronomical Society, 422, 1092
Balona L. A., 2016, Monthly Notices of the Royal Astronomical Society, 459, 1097
Balona L. A., Dziembowski W. A., 2011, Monthly Notices of the Royal Astronomical Society, 417, 591

Balona L. A., et al., 2012, Monthly Notices of the Royal Astronomical Society, 419, 3028
Bevington P. R., Robinson D. K., 2003, Data reduction and error analysis for physical sciences.. McGraw-Hill, New York, NY
Breger M., Montgomery M. H., 2014, The Astrophysical Journal, 783, 89
Brickhill A. J., 1992, Monthly Notices of the Royal Astronomical Society, 259, 519
De Franciscis S., Pascual-Granado J., Suárez J. C., García Hernández A., Garrido R., 2018, Monthly Notices of the Royal Astronomical Society, 481, 4637
Degroote P., et al., 2009, Astronomy & Astrophysics, 506, 471
Dziembowski W. A., Krolikowska M., 1985, Acta Astronómica, 35, 5
Dziembowski W. A., Krolikowska M., Kosovichev A. G., 1988, Acta Astronómica, 38, 61
García Hernández A., et al., 2013, Astronomy & Astrophysics, 559, A63
Garrido R., Rodriguez E., 1996, Monthly Notices of the Royal Astronomical Society, 281, 696
Gillet D., Fokin A. B., 2014, Astronomy & Astrophysics, 565, A73
Gilliland R. L., et al., 2010, Publications of the Astronomical Society of the Pacific, 122, 131
Handler G., et al., 2006, Monthly Notices of the Royal Astronomical Society, 365, 327
Kallinger T., Reegen P., Weiss W. W., 2008, Astronomy & Astrophysics, 481, 571
Kurtz D. W., Shibahashi H., Murphy S. J., Bedding T. R., Bowman D. M., 2015, Monthly Notices of the Royal Astronomical Society, 450, 3015
Lorimer D. R., Kramer M., 2004, Handbook of Pulsar Astronomy. Vol. 4
Mantegazza L., Poretti E., Michel E., others 2012, Astronomy & Astrophysics, 542, A24
Montgomery M. H., 2005, The Astrophysical Journal, 633, 1142
Murphy S. J., et al., 2013, Monthly Notices of the Royal Astronomical Society, 432, 2284
Pápics P. I., 2012, arXiv e-prints, p. arXiv:1210.5834
Pascual-Granado J., Garrido R., Suárez J. C., 2015, Astronomy & Astrophysics, 575, A78
Pascual-Granado J., Suárez J. C., Garrido R., Moya A., García Hernández A., Rodón J. R., Lares-Martiz M., 2018, Astronomy & Astrophysics, 614, A40
Paxton B., et al., 2019, The Astrophysical Journal Supplement Series, 243, 10
Poretti E., et al., 2009, Astronomy & Astrophysics, 506, 85
Priestley M. B., 1988, Non-linear and non-stationary time series analysis. Academic Press
Reegen P., 2007, Astronomy & Astrophysics, 467, 1353
Ricker G. R., et al., 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
Saio H., Bedding T. R., Kurtz D. W., Murphy S. J., Antoci V., Shibahashi H., Li G., Takata M., 2018, Monthly Notices of the Royal Astronomical Society, 477, 2183
Scargle J. D., 2020, Preprint (arXiv-ph.IM/08314v1)
Stellingwerf R. F., 1979, The Astrophysical Journal, 227, 935
Sterken C., 2005, in Sterken C., ed., Astronomical Society of the Pacific Conference Series, Vol. 335, The Light-Time Effect in Astrophysics, Proceedings of ASP Conference Series, Vol. 335, held in Brussels 19-22 July 2004. Edited by C. Sterken. San Francisco: Astronomical Society of the Pacific, 2005, p. 3.. p. 3
Suárez J. C., Garrido R., Pascual-Granado J., García Hernández A., de Franciscis S., Lares-Martiz M., Rodón J. R., 2020, Frontiers in Astronomy and Space Sciences, 7, 12
Wiener N., 1958, Nonlinear Problems in Random Theory., doi:10.1063/1.3060939.
Wu Y., 2001, Monthly Notices of the Royal Astronomical Society, 323, 248
Yang T., et al., 2018, The Astrophysical Journal, 863, 195
Zwintz K., Neiner C., Kochukhov O., Ryabchikova T., Pigulski A., Müllner M., Steindl T., Kuschnig R., 2020, β Cas: the first δ Scuti star with a dynamo magnetic field, http://smei.ucsd.edu/new_smei/index.html

Table 6. The 'best' parent search tree for the synthetic light curve build from TIC 9632550 data. First column quantifies the number of statistically significant frequencies, or children, detected with the parent frequency specified in the third column, in cycles per days units (zeros omitted for the sake of clarity). Second column is the variance after the fit of the parent and combination frequencies (in this case, only harmonics of the highest one).

N of fitted frequencies	V value	Frequency [c/d]
1	3156.591456884018044	5.0
5	948.685924387723073	5.05
14	7.165213986246192	5.055
14	7.165213986246192	5.055
14	0.804372110830800	5.05496
14	0.007417117763935	5.054964
14	0.007417117763935	5.054964
14	0.000552438323944	5.05496404
14	0.000045278362580	5.054964037
14	0.000005430569507	5.0549640372
14	0.000000546520136	5.05496403723
14	0.000000051236668	5.054964037227
14	0.000000008582531	5.0549640372273
14	**6.02951e-10**	**5.05496403722726**

7 APPENDIX

7.1 Uncertainties in frequencies.

The method described in this manuscript for studying combination frequencies mainly relies in how we determine the 'best' parents. Progressively increasing the precision in frequency (when searching for the minimum of the variance of the residuals) involves getting closer to the floating point number precision, which implies that numerical errors are an important source of uncertainty.

Finding when this numerical effects are hampering the 'best' parents computations, will provide us with an estimate of the upper limit in the uncertainty of the frequencies.

We find this limit by building a synthetic light curve in this way:

$$S(t) = \sum_{k=1}^{n} A_k \cos(2\pi k\omega t + \phi_k), \quad k, n \in \mathbb{N} \tag{10}$$

where ω is the parent frequency for a mono-periodic variable up to n harmonics. The input Fourier parameters of the synthetic light curve have been obtained applying the method initially to real data (the 'best' parent frequencies and their combinations). The synthetic light curve will have the same number of data points as the observations and no added noise.

The output of the method applied to the simulated light curve will have to converge to the 'best' parents initial values (i.e. the input parent and combination frequencies of the synthetic light curve). The variance at that point (V value), theoretically expected to be zero, will reveal the error in machine calculations.

Results of this test using the components extracted form TIC 9632550, the mono-periodic δ Sct star observed by TESS, are listed in Table 6. The initial 'best' parent is reached with the V value of order $\approx 6 \cdot 10^{-10}$. Consequently, V values smaller than this number are compromised by the numerical errors intrinsic to the machine calculations.

In addition, we divided the real light curve of TIC 9632550 in four sections and find the 'best' parent in each of this partitions. In spite of the reduced frequency resolution, due to the smaller time interval of the light curve, the parent frequency found for each partition is similar, and also compatible with the O-C method up to

Table 7. Results of the fundamental frequency determination by the 'best' parent search and O-C method for the 4 partitions of the light curve of the mono-periodic δ Sct star, TIC 9632550. Each section is ≈7 days long.

Section	'Best' parent [c/d]	O-C Frequency [c/d]
S1	5.05491643	5.0548 ± 0.0001
S2	5.05490929	5.0550 ± 0.0001
S3	5.05501342	5.0549 ± 0.0002
S4	5.05501167	5.0548 ± 0.0002

the 4^{th} decimal (see Table 7). This test confirm the robustness of the BPM search .

In order to test if the effect of leakage had something to do with the small variation in the 4^{th} decimal (i.e. ≈4e-6 s period variations), we applied the BPM to the mono-periodic light curve with exactly an integer number of cycles and with an integer number of cycles plus half a cycle.

Results of this test highlighted the influence of the number of cycles on the determination of the period. Future work could explore this issue to give a precise lower limit to the frequency uncertainty. In this regard, a conservative approach is adopted in this paper expressing results with the precision that the O-C method achieves.

As a last remark, the duration of the observation is a relevant parameter for estimating the frequency uncertainty when dealing with two close frequencies (closer than $1/T$ to each other), and so, the Rayleigh frequency resolution becomes a good estimator. But, precision can go further the Rayleigh frequency resolution if there is only one frequency component inside the Rayleigh interval. We just proved this for each ≈7 days long sections of the light curve of the mono-periodic δ Sctstar, TIC 9632550. The numerical precision reached is $\pm 1 \cdot 10^{-8}$ and is not "as if we were observing ≈274000 years". Rayleigh resolution remains the same: ≈1/7 days, that is, ≈0.14 c/d.

7.2 Stellar parameters and relevant information of the time series used.

7.3 Combination frequencies detected.

This paper has been typeset from a TEX/LATEX file prepared by the author.

Table 8. Stellar parameters from Gaia DR2 catalogue and time series information from each space satellite. T is the length of the observation in days and δ_t is the cadence or sampling rate in seconds. For the TESS and Kepler light curves, we used the instrumental effects free light curve, resulting from the Pre-Search data Conditioning (PDC) pipeline, accessible in the Mikulski Archive for Space Telescopes (MAST: https://archive.stsci.edu/).

Stellar Parameters				Time series Parameters		
Name	Spectral Type	Magnitude m_v [mag]	Effective Temp. $T_{\rm eff}$ [K]	T [d]	δ_t [s]	Obs. Sequence
TIC 9632550	A8III	9.317 ± 0.009	$7009.25^{+7210.64}_{-6800.90}$	27.41	120.01	Sector 2
KIC 5950759	–	13.8271 ± 0.0141	$7842.5^{+7995.00}_{-7595.00}$	31.04	58.85	Quarter 4
HD 174966	A7III/IV	7.6498 ± 0.0005	$7446.67^{+7583.33}_{-7291.50}$	27.20	31.99	Run SRc01

Table 9. Tags of the statistically significant combination frequencies for the mono-periodic δ Sct star TIC 9632550. The frequency values can be calculated with the given parent frequencies resulting from BPM since the fitted values are the exact combination frequency values

Non-linearities of TIC 9362550		
2f0	7f0	12f0
3f0	8f0	13f0
4f0	9f0	14f0
5f0	10f0	
6f0	11f0	

Table 10. Tags of the statistically significant combination frequencies for the double mode HADS star KIC 5059759. The frequency values can be calculated with the given parent frequencies resulting from BPM since the fitted values are the exact combination frequency values

Non-linearities of KIC 5059759					
2f0	3f0+2f1	7f0+5f1	11f0+7f1	6f0-3f1	4f1-1f0
3f0	3f0+3f1	7f0+6f1	12f0+1f1	7f0-1f1	4f1-2f0
4f0	3f0+4f1	7f0+7f1	12f0+2f1	7f0-2f1	4f1-3f0
5f0	3f0+5f1	8f0+1f1	12f0+3f1	7f0-3f1	4f1-5f0
6f0	3f0+6f1	8f0+2f1	12f0+4f1	7f0-4f1	5f1-1f0
7f0	3f0+7f1	8f0+3f1	12f0+5f1	7f0-5f1	5f1-2f0
8f0	4f0+1f1	8f0+4f1	12f0+6f1	8f0-1f1	5f1-3f0
9f0	4f0+2f1	8f0+5f1	13f0+1f1	8f0-2f1	5f1-4f0
10f0	4f0+3f1	8f0+6f1	13f0+2f1	8f0-4f1	5f1-6f0
11f0	4f0+4f1	8f0+7f1	13f0+3f1	8f0-6f1	6f1-4f0
12f0	4f0+5f1	9f0+1f1	13f0+4f1	9f0-1f1	6f1-5f0
13f0	4f0+6f1	9f0+2f1	13f0+5f1	9f0-2f1	6f1-6f0
14f0	4f0+7f1	9f0+3f1	13f0+6f1	9f0-6f1	6f1-7f0
2f1	5f0+1f1	9f0+4f1	14f0+1f1	10f0-1f1	7f1-6f0
3f1	5f0+2f1	9f0+5f1	14f0+2f1	10f0-2f1	7f1-7f0
4f1	5f0+3f1	9f0+6f1	14f0+3f1	10f0-4f1	7f1-8f0
5f1	5f0+4f1	9f0+7f1	14f0+4f1	10f0-6f1	7f1-9f0
1f0+1f1	5f0+5f1	10f0+1f1	14f0+5f1	10f0-7f1	8f1-7f0
1f0+2f1	5f0+6f1	10f0+2f1	15f0+4f1	11f0-1f1	8f1-6f0
1f0+3f1	6f0+1f1	10f0+3f1	2f0-1f1	11f0-2f1	8f1-8f0
1f0+4f1	6f0+2f1	10f0+4f1	3f0-1f1	11f0-7f1	8f1-9f0
1f0+5f1	6f0+3f1	10f0+5f1	3f0-2f1	11f0-4f1	9f1-7f0
1f0+6f1	6f0+4f1	10f0+6f1	4f0-1f1	12f0-1f1	9f1-8f0
2f0+1f1	6f0+5f1	10f0+7f1	4f0-2f1	20f0-14f1	15f1-19f0
2f0+2f1	6f0+6f1	11f0+1f1	4f0-3f1	1f1-1f0	16f1-19f0
2f0+3f1	6f0+7f1	11f0+2f1	5f0-1f1	2f1-1f0	17f1-19f0
2f0+4f1	7f0+1f1	11f0+3f1	5f0-2f1	2f1-2f0	18f1-20f0
2f0+5f1	7f0+2f1	11f0+4f1	5f0-3f1	3f1-1f0	
2f0+6f1	7f0+3f1	11f0+5f1	6f0-1f1	3f1-2f0	
3f0+1f1	7f0+4f1	11f0+6f1	6f0-2f1	3f1-3f0	

Table 11. Tags of the statistically significant combination frequencies for the multi-periodic δ Sct star HD 174966. The frequency values can be calculated with the given parent frequencies resulting from BPM since the fitted values are the exact combination frequency values

Non-linearities of HD 174966			
2f3	6f0-5f2	9f1-9f2	8f2-7f3
1f0+1f1	7f0-6f2	1f1-1f3	9f2-6f3
1f0+1f2	8f0-6f2	3f1-2f3	9f2-7f3
1f0+1f3	9f0-6f2	3f1-5f3	1f2-1f4
1f1+1f2	9f0-7f2	4f1-3f3	3f2-1f4
1f1+1f3	9f0-8f2	6f1-4f3	4f2-2f4
1f1+2f3	2f0-1f3	7f1-6f3	4f2-3f4
1f1+1f4	3f0-2f3	9f1-8f3	4f2-4f4
2f1+1f3	5f0-2f3	3f1-3f4	5f2-4f4
2f1+1f4	5f0-4f3	4f1-2f4	6f2-3f4
1f2+1f4	5f0-5f3	4f1-4f4	6f2-4f4
1f3+1f4	7f0-4f3	5f1-4f4	7f2-4f4
1f0-1f1	7f0-5f3	6f1-5f4	7f2-6f4
2f0-1f1	8f0-5f3	7f1-6f4	8f2-6f4
3f0-1f1	8f0-6f3	8f1-6f4	8f2-8f4
3f0-3f1	2f0-2f4	8f1-7f4	8f2-9f4
5f0-4f1	2f0-3f4	8f1-8f4	9f2-7f4
6f0-3f1	3f0-1f4	1f2-1f3	9f2-8f4
6f0-5f1	4f0-3f4	2f2-1f3	1f3-1f4
7f0-6f1	4f0-4f4	3f2-2f3	1f3-2f4
8f0-5f1	5f0-2f4	4f2-3f3	2f3-1f4
8f0-8f1	7f0-4f4	5f2-2f3	2f3-3f4
9f0-7f1	8f0-6f4	5f2-3f3	3f3-1f4
9f0-8f1	1f1-1f2	5f2-4f3	4f3-4f4
1f0-2f2	1f1-2f2	5f2-5f3	5f3-4f4
1f0-3f2	3f1-3f2	6f2-4f3	5f3-5f4
4f0-3f2	5f1-6f2	6f2-5f3	6f3-5f4
5f0-3f2	6f1-7f2	7f2-6f3	6f3-6f4
5f0-4f2	7f1-7f2	7f2-7f3	
5f0-5f2	8f1-8f2	8f2-6f3	