Astronomy & Astrophysics manuscript no. IGWs_GP_TASOC
March 7, 2022

Photometric detection of internal gravity waves in upper main-sequence stars

III. Comparison of Gaussian processes and amplitude spectrum fitting

Dominic M. Bowman[1,2] and Trevor Z. Dorn-Wallenstein[3]

[1] Institute of Astronomy, KU Leuven, Celestijnenlaan 200D, B-3001 Leuven, Belgium
e-mail: dominic.bowman@kuleuven.be
[2] Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA
[3] University of Washington Astronomy Department, Physics and Astronomy Building, 3910 15th Ave NE, Seattle, WA 98105, USA

Received XX March 2022; accepted YYYY

ABSTRACT

Context. Recent studies of massive stars using high-precision space photometry have revealed that they commonly exhibit stochastic low-frequency (SLF) variability. This has been interpreted as being caused by internal gravity waves (IGWs) excited at the interface of convective and radiative regions within stellar interiors, such as the convective core or sub-surface convection zones, or caused by dynamic turbulence excited in the envelopes of main-sequence massive stars.
Aims. We aim to compare the properties of SLF variability in massive main-sequence stars observed by the Transiting Exoplanet Survey Satellite (TESS) mission determined by different fitting methods, and establish the robustness of how SLF variability correlates with a star's location in the Hertzsprung–Russell (HR) diagram. We also aim to quantify the impact of data quality on the inferred SLF morphologies using both fitting methodologies.
Methods. From a sample of previously observed and characterised massive stars observed by TESS, we compare the resultant parameters of SLF variability, in particular the characteristic frequency, obtained from fitting the amplitude spectrum of the light curve with those inferred from Gaussian process (GP) regression of the light curve.
Results. We find a notable difference in the characteristic frequency obtained from the amplitude spectrum fitting and the Gaussian process regression of the light curve for a minority of the considered sample. However, the trends among mass, age and the properties of SLF variability previously reported remain unaffected. We also find that the method of GP regression is more efficient in terms of computation time and, on average, more robust against the quality and noise properties of the input time series data in determining the properties of SLF variability.
Conclusions. GP regression is a useful and novel methodology to efficiently characterise SLF variability in massive stars compared to previous techniques used in the literature. We conclude that the correlation between a star's SLF variability, in particular the characteristic frequency, and its location in the HR diagram is robust for main-sequence massive stars. There also exists a distribution in the stochasticity of SLF variability in massive stars, which indicates that the coherency of oscillatory characteristics are also a function of mass and age in massive stars.

Key words. stars: early-type – stars: fundamental parameters – stars: massive – stars: rotation – stars: oscillations

1. Introduction

The treasure trove of time series photometric data assembled by space telescopes including CoRoT (Auvergne et al. 2009), Kepler/K2 (Borucki et al. 2010; Koch et al. 2010; Howell et al. 2014) and TESS (Ricker et al. 2015) in the last decade have revealed a remarkable range of diverse variability mechanisms in massive stars (Bowman 2020). A recent discovery is that essentially all early-type main-sequence stars show stochastic low-frequency variability (SLF) in time series photometry (Bowman et al. 2019a,b, 2020), in addition to other variability mechanisms including coherent heat-driven pulsation modes, rotational modulation, and the photometric signatures of binarity (Degroote et al. 2009a; Blomme et al. 2011; Pedersen et al. 2019; Burssens et al. 2020). Similarly, SLF variability has also been reported in later evolutionary stages of massive stars (see e.g., Bowman et al. 2019b; Dorn-Wallenstein et al. 2019, 2020; Nazé et al. 2021; Elliott et al. 2022), which show similar morphologies to main sequence stars. It currently remains unclear which of the several plausible mechanisms that can generate SLF variability dominate in particular parameter regimes. Hence the investigation of SLF variability has opened up a novel method to understanding the interiors of massive stars (Bowman 2020).

Coupled with spectroscopy, the analysis of time series photometry is able to extract detailed and precise constraints on the physics of stellar interiors thanks to asteroseismology (Aerts et al. 2010; Aerts 2021). In early-type stars, a heat-engine excitation mechanism operating in the iron opacity bump at 200 000 K gives rise to pressure (p) and gravity (g) modes (Dziembowski & Pamyatnykh 1993; Dziembowski et al. 1993; Miglio et al. 2007; Szewczuk & Daszyńska-Daszkiewicz 2017), which probe predominantly the envelope and near-core physics, respectively. The analysis of coherent pulsations has so far revealed a diversity in the amounts of core-boundary and envelope mixing profiles in late-B stars (Degroote et al. 2009b, 2010; Daszyńska-

Daszkiewicz et al. 2013; Moravveji et al. 2015, 2016; Szewczuk & Daszyńska-Daszkiewicz 2018; Szewczuk et al. 2021; Pedersen et al. 2021; Michielsen et al. 2021; Bowman & Michielsen 2021), and tight constraints on differential radial rotation profiles and (core) masses and ages for late-O and early-B stars (Aerts et al. 2003; Pamyatnykh et al. 2004; Dziembowski & Pamyatnykh 2008; Burssens et al. 2022).

Similarly to asteroseismology of coherent pulsation modes, the exploitation of SLF variability in time series photometry, which is seen across a wide range in mass, age and metallicity in massive stars, offers us the opportunity to improve our understanding of stellar structure and evolution (Bowman 2020). In their original methodology and characterisation study, Bowman et al. (2019a) used a sample of 35 early-type stars observed by the CoRoT mission to demonstrate how scaling laws based on solar granulation are inconsistent with the observed SLF variability[1] (Paper I). Based on a much larger sample of more than 160 massive stars observed by Kepler/K2 and TESS, which included metal-rich galactic stars and metal-poor stars in the Large Magellanic Cloud (LMC), Bowman et al. (2019b) demonstrated that a common morphology in the SLF variability is observed for essentially all stars with $M \gtrsim 3$ M$_\odot$. The SLF variability was determined to be dependent on the brightness of the star and hence its location in the Hertzsprung–Russell (HR) diagram. More recently, Bowman et al. (2020) used a sample of galactic O and B stars observed by the TESS mission and combined it with high resolution spectroscopy (Simón-Díaz et al. 2011; Burssens et al. 2020) to show that the morphology of SLF in massive stars is strongly dependent on the mass and age of the star (Paper II).

The detection of SLF variability in massive stars has inspired various theoretical works to explain this new type of observational signal. Motivated by the omnipresence of a convective core during the main-sequence evolution of a massive star, it was first Aerts & Rogers (2015) who postulated that SLF variability observed at the surface of three O-type stars could be explained by convectively driven internal gravity waves (IGWs) excited by core convection. This conclusion was supported based on the qualitatively similar frequency spectra observed for these stars with those predicted from 2D hydrodynamical simulations (Rogers et al. 2013; Rogers 2015; Rogers & McElwaine 2017). More recent 2D and 3D simulations of core convection with different numerical setups also show that IGWs excited by core convection are predicted to produce detectable velocity and temperature perturbations (Edelmann et al. 2019; Horst et al. 2020). Other mechanisms that may be responsible for SLF variability in massive stars are IGWs excited by sub-surface convection zones (Cantiello et al. 2009, 2021; Lecoanet et al. 2019, 2021), or the convective dynamics of stellar envelopes (Jiang et al. 2015, 2018; Schultz et al. 2022). However, the efficiency of convection in sub-surface zones is strongly metallicity dependent (Jermyn et al. 2022), which complicates the goal of finding a universal explanation of SLF variability in massive stars. Furthermore, in more evolved post-main sequence stars, the effect of optically thick and clumping in winds can also produce SLF variability (see e.g. Krtička & Feldmeier 2018; Krtička & Feldmeier 2021).

It remains an open question as to which excitation mechanism(s) dominates in specific parameter regimes of the HR diagram, as the properties of (sub-surface) convection, and the driving and propagation of IGWs depend on stellar properties such as mass, radius, metallicity and rotation. What is clear is that a mechanism that aims to unanimously explain SLF variability in massive stars must do so for stars that span a broad range of masses (i.e. $3 \lesssim M \lesssim 100$ M$_\odot$), ages from the zero-age main sequence (ZAMS) to far beyond the terminal-age main sequence (TAMS), rotation rates between essentially zero and close to the critical rotation rate, and for both metal-poor (i.e. $Z \leq 0.002$) and metal-rich stars (i.e. $Z \geq 0.014$). A mechanism must also explain the broad frequency range of SLF variability and how its properties correlate with mass and age (Bowman et al. 2019b, 2020).

In this paper, part III of the current series, we compare the difference in the measured properties of SLF variability when using different fitting methods: directly fitting the amplitude spectrum of the light curve as previously performed by Bowman et al. (2020), and Gaussian process (GP) regression of the light curve. Our goal is to demonstrate a more efficient method for extracting the characteristic frequency of the SLF variability for a large number of massive stars observed by the ongoing TESS mission. In section 2, we outline the approach to using GP regression and briefly describe the stellar sample. In section 3, we present the results of our comparison and we conclude in section 4.

2. Method

Given that SLF variability observed in massive stars is inherently a representation of a stochastic process, there is no reasonable parametric function to model it in the time domain. This has led previous studies to characterise SLF variability instead in the Fourier domain (Blomme et al. 2011; Bowman et al. 2019a,b, 2020; Dorn-Wallenstein et al. 2019, 2020; Nazé et al. 2021; Elliott et al. 2022). However, there are tools and methodologies, specifically GP regression (Rasmussen & Williams 2006), that yield a non-parametric model for stochastic signals in the time domain based on the statistical properties of time-series data.

The application of GP regression has been highly successful in modelling radial velocity time series data, the light curves of transiting exoplanet host stars, and stellar variability in evolved low-mass stars (Brewer & Stello 2009; Barclay et al. 2015; Grunblatt et al. 2017; Pereira et al. 2019). As pointed out by Pereira et al. (2019), previous applications of GP regression in astrophysical contexts have typically focussed on determining accurate non-parametric models for time series data without particularly paying close attention to the physical processes that cause the variability. However, GP regression can be used under certain conditions to estimate physical parameters that set the stochastic processes in time series data, such as modelling the signatures of granulation in the light curves of evolved low-mass stars (Pereira et al. 2019). In this work, we extend and test the application of GP regression to extract physical parameters of SLF variability in massive stars.

2.1. Previous studies parameterising SLF variability

Previous efforts in parameterising SLF variability in massive stars have used a semi-Lorentzian function to fit the amplitude spectrum of the light curve with a from similar to:

$$\alpha(\nu) = \frac{\alpha_0}{1 + \left(\frac{\nu}{\nu_{\rm char}}\right)^{\gamma}} + C_{\rm w} \,, \tag{1}$$

where α_0 represents the amplitude at a frequency of zero, γ is the logarithmic amplitude gradient, $\nu_{\rm char}$ is the characteristic frequency, and $C_{\rm w}$ is a frequency-independent (i.e. white) noise term (Bowman et al. 2019a).

[1] Sometimes referred to as a low-frequency power excess or 'red noise' in the literature.

Typically, to avoid an uncontrollable bias in the resultant fit when comparing various stars using Eqn. (1), the light curves of massive stars have undergone iterative pre-whitening to remove dominant periodicities a priori. This means that periodic variability caused by coherent pulsation modes or rotational modulation are removed from the light curve by means of subtracting a sinusoid with a given amplitude, frequency and phase (Bowman et al. 2019a). This is particularly important for B-type stars, such as β Cep and SPB stars, which are commonly multi-periodic pulsators in addition to having SLF variability (Bowman et al. 2019b; Burssens et al. 2020). On the other hand, the parameters of pulsation modes (i.e. frequencies, amplitudes and phases) are also covariant with the parameters of SLF variability. Thus a two-step process of pre-whitening pulsation modes prior to fitting the SLF variability will inevitably still include a bias. Investigating the impact of this bias is beyond the scope of the current paper but will be explored in future work.

The fitting of the amplitude spectrum in previous studies has been achieved using a Markov chain Monte Carlo (MCMC) numerical framework (e.g. using the `python` EMCEE package; Foreman-Mackey et al. 2013, 2019c). It was used to sample the posterior probability distribution to determine the optimum parameters for Eqn. (1) for a given star. Yet such an approach is relatively computationally expensive and is also potentially quite sensitive to the explicit algorithm used to transform from the time domain into the frequency domain. For example, the choice of whether to use a (discrete) Fourier transform (e.g. Deeming 1975; Kurtz 1985) or (a variant of) a Lomb-Scargle periodogram (e.g. Scargle 1982; Press & Rybicki 1989) should ideally not impact the results. Yet, such a calculation is a numerical form of post-processing of the data set, and includes choices such as the frequency over-sampling factor, frequency range, normalisation constant (e.g. the number of data points and/or the data set length). Not all studies in the literature approach such tasks in a similar way — see discussion by Bowman & Michielsen (2021). Other forms of post-processing include gap-filling a data set to maximise the duty cycle and minimise the shot noise level in the amplitude spectrum. This may seem advantageous but implicitly assumes a model for the signal and noise components of the data and can affect subsequent (asteroseismic) inference. Because of these choices, it can be argued that fitting the light curve directly is the most robust strategy since it is the original data set. Relying on the light curve as the data set and avoiding transformations into the Fourier domain, a GP regression method avoids these potential discrepancies.

In this work, we investigated the efficacy of using GP regression for determining a non-parametric fit to the light curves of massive stars dominated by SLF variability. In so doing, we restricted our sample to the 30 O-type stars from the TESS sample previously studied by Bowman et al. (2020) that do not have significant coherent pulsation modes and thus do not require iterative pre-whitening prior to fitting Eqn. (1). Therefore, our comparison of fitting the amplitude spectrum and GP regression of the light curve was not influenced by any post-processing of the light curves, such as iterative pre-whitening.

2.2. Defining a GP kernel

A GP is a non-parametric model that describes correlated stochastic variability in a time series. Each data point in a time series is described as a correlated random variable with a mean and a variance, for which any finite collection of variables has a multi-variate Gaussian distribution (Rasmussen & Williams 2006). In the application of GP regression, a kernel to describe the correlation among data points is chosen, which is then applied to determine the best set of parameters to reproduce the observed time series (see e.g. Pereira et al. 2019). The quality of the fit is judged by a GP marginalized log likelihood function (see equation 1 of Foreman-Mackey 2018). This is typically the computational bottleneck, therefore the selection of an appropriate GP kernel is crucial (Foreman-Mackey et al. 2017).

In this work, we used the `python` CELERITE2 and EXOPLANET software packages (Foreman-Mackey et al. 2017; Foreman-Mackey 2018; Foreman-Mackey et al. 2021)[2] for performing GP regression of TESS light curves. Inspired by Pereira et al. (2019) in their application of GP regression to study the stochastic granulation variability seen in red giants, we utilised a GP kernel that corresponds to a stochastically driven damped harmonic oscillator, whose power spectral density is given by:

$$S(\omega) = \sqrt{\frac{2}{\pi}} \frac{S_0 \omega_0^4}{\left(\omega^2 - \omega_0^2\right)^2 + \frac{\omega^2 \omega_0^2}{Q^2}} \tag{2}$$

where ω_0 is the resonant frequency of the oscillator, S_0 is related to the amplitude of the variability, and Q is the quality factor. For very large values of Q, Eqn. (2) corresponds to an undamped simple harmonic oscillator. In the limit of $Q = 1/\sqrt{2}$ (see Pereira et al. 2019 and Foreman-Mackey et al. 2017 for the derivation; omitted here for brevity), the power spectral density reduces to

$$S(\omega) = \sqrt{\frac{2}{\pi}} \frac{S_0}{\left(\frac{\omega}{\omega_0}\right)^4 + 1} \, . \tag{3}$$

One can also parameterise the GP of the time series using a characteristic timescale, $\rho_{\rm char} = 2\pi/\omega_0$, the standard deviation, $\sigma = \sqrt{S_0 \omega_0 Q}$, and a characteristic damping timescale, $\tau_{\rm damp} = 2Q/\omega_0$. The form of Eqn. (3) is similar to the characteristic semi-Lorentzian function used to fit SLF variability in massive stars (cf. Eqn. (1)). Hence, it is applicable for modelling the SLF variability observed in massive stars.

In our framework, we used the kernel of a damped harmonic oscillator and input guesses for its parameters as follows: ω_0 as the characteristic angular frequency (i.e. $2\,\pi\,\nu_{\rm char}$); $Q = 1/\sqrt{2}$; and $\sigma = \alpha_0 \sqrt{\omega_0\, Q} \times (\pi/2)^{0.25}$ based on solving for σ from equating Eqns (1) and (3). We also emulated white noise in Fourier space by including a noise jitter term in the GP regression following Pereira et al. (2019), which is proportional to $\sqrt{N}$, where N is the number of data points in the light curve. The jitter term is a constant added in quadrature to the diagonal of the covariance matrix and has a kernel function of:

$$k(t_i, t_j) = C^2_{\rm jitter}\, \delta_{i,j} \, , \tag{4}$$

where i and j are matrix elements and $C^2_{\rm jitter}$ is the GP jitter term. Physically, the jitter term is used to represent uncorrelated noise in the data that cannot be captured by the GP regression, and hence represents a proxy of the measurement uncertainty in a light curve (Foreman-Mackey et al. 2017; Foreman-Mackey 2018; Foreman-Mackey et al. 2021).

[2] `https://github.com/exoplanet-dev/celerite2`

Table 1. Lower and upper limits of the four uniform logarithmic priors used in the GP regression.

Parameter	Lower Bound	Upper Bound
$\rho_{\rm char}$	$\ln\left(\frac{0.01}{\nu_{\rm char,guess}}\right)$	$\ln\left(\frac{100}{\nu_{\rm char,guess}}\right)$
$\tau_{\rm damp}$	$\ln\left(0.01\left(\frac{2Q_{\rm guess}}{\omega_{0,\rm guess}}\right)\right)$	$\ln\left(100\left(\frac{2Q_{\rm guess}}{\omega_{0,\rm guess}}\right)\right)$
σ	$\ln(\sigma) = -10$	$\ln(\sigma) = +10$
$C_{\rm jitter}$	$\ln(C_{\rm jitter}) = -15$	$\ln(C_{\rm jitter}) = +15$

2.3. Input priors

To run a GP regression for a given star, there are five main parameters to be determined: (i) the characteristic timescale, $\rho_{\rm char}$ (from which the characteristic frequency can be determined: $\omega_0 = 2\pi\nu_{\rm char}$, where $\nu_{\rm char} = 1/\rho_{\rm char}$); (ii) the characteristic damping time scale, $\tau_{\rm damp}$ (from which the characteristic damping frequency can be calculated: $\nu_{\rm damp} = 1/\tau_{\rm damp}$); (iii) the standard deviation of the variability process, σ; (iv) a jitter term to emulate uncorrelated noise in the observations, $C_{\rm jitter}$; and (v) the mean of the light curve. We take the $\nu_{\rm char}$ parameters from Bowman et al. (2020) multiplied by 2π to calculate $\omega_{0,\rm guess}$ as input guesses, and assign non-informative (i.e. uniform) priors in logarithmic space to all input parameters to the GP regression except the mean of the light curve, which is assigned a normal prior based on the mean of TESS input time series. We specified liberal lower and upper bounds for the four uniform priors as given in Table 1. An input guess of the quality factor of the GP regression model is set as $Q_{\rm guess} = 1/\sqrt{2}$, and is evaluated as $Q = \pi\tau_{\rm damp}/\rho_{\rm char}$. We note that our results are in fact not sensitive to the input guesses for these parameters as long as the priors are wide enough to contain the true value.

2.4. Parameter and uncertainty estimation

A GP regression is a non-parametric fit to the light curve, so it does not encode amplitude or phase information to automatically transform the resultant model into the Fourier domain. To make our results visually comparable to the amplitude spectrum fitting results of Bowman et al. (2020), we calculated the power spectral density (i.e. amplitude squared per unit frequency as a function of frequency) of the TESS light curve and the GP regression model. We also converted the fits from Bowman et al. (2020) into units of power spectral density for comparison purposes. Following Pereira et al. (2019), we normalised the power spectral density of both the GP regression model and the TESS data using Parseval's theorem. This is necessary to allow a consistent comparison of the power spectral density of the GP regression model including the jitter term and that of the TESS data in the Fourier domain. We also normalised the power spectral density of the TESS data and the GP regression model to have the same amplitude at zero frequency (i.e. α_0), such that we enforced that the data and both fitting techniques converge at unity at zero frequency. This allowed a more sensible comparison when visually comparing the two fitting methods in the Fourier domain, but does not impact the inferred parameters, such as $\nu_{\rm char}$.

To estimate the uncertainties of the GP regression parameters, we used the `python` PYMC3 software package (Salvatier et al. 2016)[3] and performed a Hamiltonian Monte-Carlo with a no U-turn sampler to determine parameter confidence intervals. In a standard MCMC approach, parameter chains make stochastic walks within the multi-dimension parameter space, whereas a no U-turn sampler employs a Hamiltonian description of the probability distribution to sample the posterior probability distribution for the model parameters following Bayes' theorem. This makes a Hamiltonian Monte-Carlo no U-turn samplers typically much more efficient in finding global minima (Salvatier et al. 2016).

In our implementation of the Hamiltonian Monte-Carlo with a no U-turn sampler, we used four parameter chains, and at each iteration each chain is used to construct a model subject to a GP marginalised log likelihood function within the PYMC3 model that also includes the log priors. A burn-in period of approximately 1000 iterations was typically needed, which were discarded before parameter and uncertainty estimation was performed using another 1000 iterations. Convergence of the parameter chains is confirmed using the parameter variance criteria from Gelman & Rubin (1992). The GP model is evaluated on the maximum a posteriori hyperparameters. From the posteriors, we used the median and the 94th percentile of the posterior probability density to estimate the lower and upper confidence intervals for each marginalised parameter distribution.

3. Results

We applied our GP regression framework to the exact same TESS light curves of 30 massive stars previously analysed by Bowman et al. (2020). For each star, we provide a summary figure in Appendix A, which contains the TESS observations and the best-fitting GP regression model in the top panel, and the power spectral density of the TESS observations, that of the best-fitting model from Bowman et al. (2020), and the GP regression model in the bottom panel. A full summary of our parameter results from the GP regression and their upper and lower confidence intervals is given in Table B.1.

An immediate major advantage of the GP regression methodology is that is provides additional useful inferred parameters of the SLF variability in massive stars, such as $\nu_{\rm damp}$ and Q, which are not provided by fitting Eqn. (1) to an amplitude spectrum. Therefore, in addition to the characteristic time scale of the SLF variability, we are able to make inferences on its coherency and quasi-periodicity.

3.1. Comparison of Bowman et al. (2020) and GP regression

In the top- and bottom-left panels of Fig. 1, we show the pairwise comparison of the $\nu_{\rm char}$ values determined in this work using GP regression and those determined by Bowman et al. (2020), which are colour coded by the σ and $\nu_{\rm damp}$ GP regression parameters, respectively. On average, a fairly good agreement is seen between the two sets of $\nu_{\rm char}$ values. Yet a fraction of the sample stars have $\nu_{\rm char}$ values that differ by more than a factor of two. In some cases this leads to significant differences when considering the relatively small uncertainties on this parameter, but this varies from star to star. The discrepancies in some stars are not necessarily systematic with $\nu_{\rm char}$. However, we note that stars with a significant difference in $\nu_{\rm char}$ values between the two methods have lower σ and higher $\nu_{\rm damp}$ values (i.e. the SLF variability is relatively low amplitude and more stochastic). Also, in cases when a significant difference is found, the GP regression value for $\nu_{\rm char}$ is typically larger than that of Bowman et al. (2020). This is likely related to the fact that the GP regression is a fit to the light curve, whereas a fit of the amplitude spectrum is

[3] `https://github.com/pymc-devs/pymc`

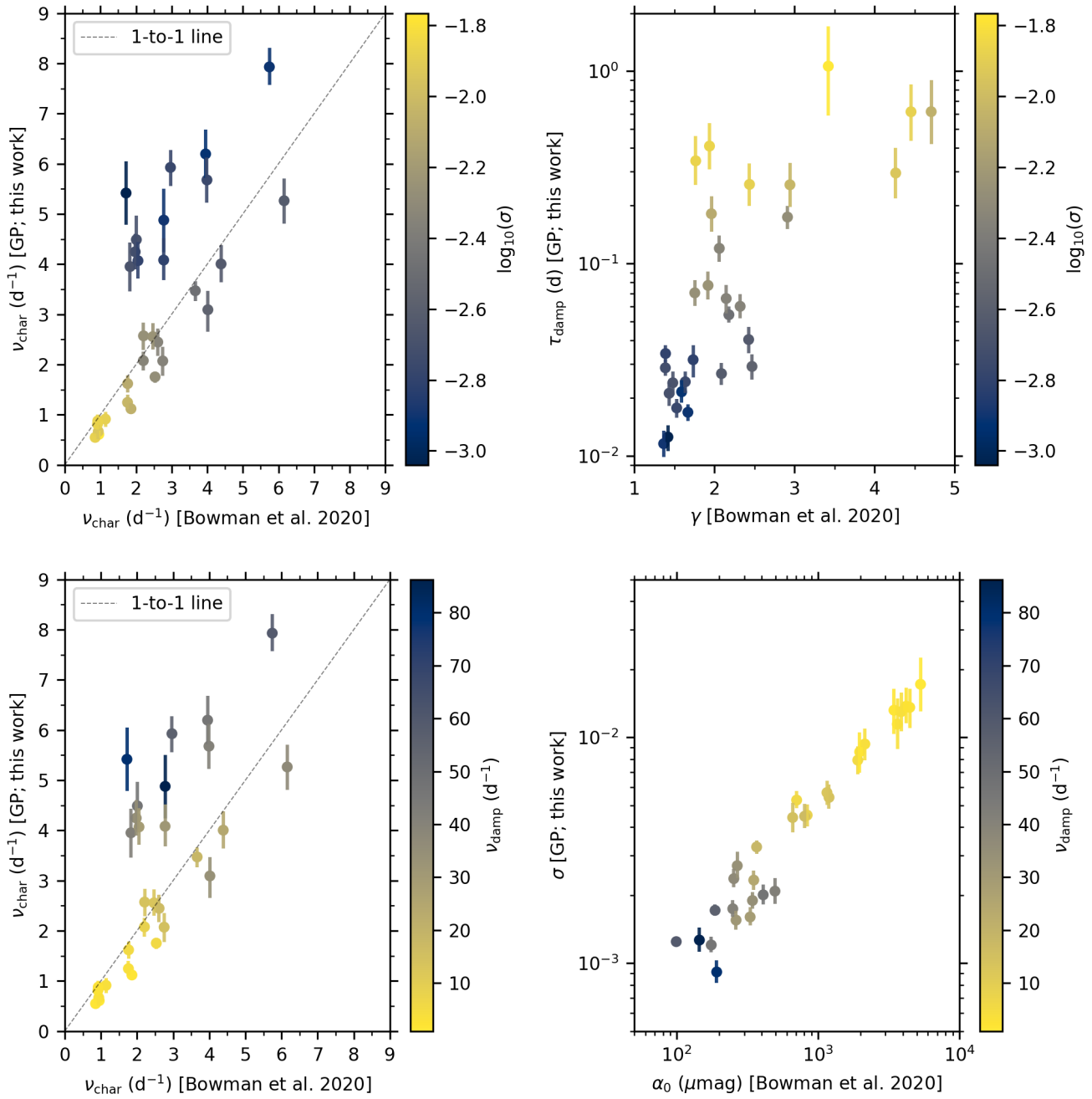

Fig. 1. Comparison of GP regression parameters determined in this work and equivalent parameters from Bowman et al. (2020). *Top-left panel:* $\nu_{\rm char}$ values colour coded by σ from the GP regression. *Top-right panel:* $\tau_{\rm damp}$ values from the GP regression and the steepness of the amplitude spectrum, γ, colour coded by σ from the GP regression. *Bottom-left panel:* $\nu_{\rm char}$ values colour coded by $\nu_{\rm damp}$ from the GP regression. *Bottom-right panel:* σ values from the GP regression and α_0 colour coded by $\nu_{\rm damp}$ from the GP regression.

somewhat biased to lower frequencies because of the amplitude suppression effect in the Fourier domain (Bowman 2017).

In the top-right panel of Fig. 1, we show the relationship between the steepness of the SLF variability in the amplitude spectrum, γ, as determined by Bowman et al. (2020) and the $\tau_{\rm damp}$ timescale colour-coded by σ as determined using GP regression in this work. Whilst there is considerable scatter between these two parameters, a positive correlation between γ and $\tau_{\rm damp}$ can clearly be seen. This correlation is caused by stars with SLF variability characterised with longer damping timescales, $\tau_{\rm damp}$, being less stochastic and more quasi-periodic (i.e. higher values of Q). In turn this leads to larger values of γ, and the majority of power in the SLF variability being concentrated near $\nu_{\rm char}$ as opposed to being spread over a wide range in frequency.

In the bottom-right panel of Fig. 1, we show the relationship between the amplitude of the SLF variability at zero frequency, α_0, as determined by Bowman et al. (2020) and the standard deviation of the GP regression, σ, in this work. A clear correlation between these parameters is seen, which demonstrates that they can serve as proxies of each other. We note that the relationship is almost linear in log-log space, with a linear regression of α_0 and σ yielding a Pearson correlation coefficient of $r = 0.97$ ($p < 10^{-10}$) and a gradient of 0.73 ± 0.03. This is expected given the equivalence of these parameters in both methodologies, but

the gradient of order unity and the extremely high degree of correlation is a useful sanity check.

Our comparison of the GP regression parameters and those of Bowman et al. (2020) leads us to define two sub-groups of stars: (i) those with high α_0 (i.e. high σ), low $\nu_{\rm char}$ (i.e. high $\tau_{\rm char}$), and low $\nu_{\rm damp}$ values; and conversely (ii) those with low σ, high $\nu_{\rm char}$, and high $\nu_{\rm damp}$ values. We term these the 'yellow' and 'blue' sub-groups, respectively, as indicated by the colour coding in Fig. 1. In other words, stars in the 'yellow' sub-group typically have high-amplitude, quasi-periodic low-frequency variability, and stars in the 'blue' group have low-amplitude, stochastic variability spanning a broad range in frequency, on average. These are not distinct populations and our sample in fact covers the transition between these two extremes. However, the definition of these two sub-groups serves to highlight the differences in the oscillatory properties of SLF variability among a sample of massive stars.

3.2. Correlation with mass and age in HR diagram

In Fig. 2 we show the spectroscopic HR diagram for the 30 O stars studied by Bowman et al. (2020), which we have remodelled using GP regression in this work. As discussed in section 3.1, there are some star-to-star differences in the determined $\nu_{\rm char}$ parameters between the results of Bowman et al. (2020) and from our GP regression. However, an important result from this work is that the dependence of $\nu_{\rm char}$ on the mass and age of a star remains unaffected. Therefore, as demonstrated by comparing the left and right panels of Fig. 2, our new modelling approach further strengthens the conclusion of Bowman et al. (2020) that the light curve of a massive star encodes valuable constraints on its mass and age. More specifically, younger stars near the ZAMS have larger $\nu_{\rm char}$ values of approximately 7 d^{-1}, on average, whereas older stars near the TAMS have smaller $\nu_{\rm char}$ values of approximately 1 d^{-1}.

Furthermore, we note that the 'yellow' sub-group of stars (cf. section 3.1) are typically higher mass and more evolved and the 'blue' sub-group of stars are closer to the ZAMS on average. This is a new result of our work. Not only does the value of $\nu_{\rm char}$ of a star's SLF variability probe its mass and age, but the degree of coherency and stochasticity of a massive star's light curve is also a function of age. More specifically, more massive and more evolved stars typically have larger Q values, as shown in Fig. 3. The larger the Q value in an oscillator, the narrower and sharper the peak of the resonant frequency becomes in terms of power spectral density because the system is more selective of what range of frequencies are driven to resonance. For example, systems with low quality factors (i.e. $Q < 1/\sqrt{2}$) are refereed to as 'overdamped' and do not oscillate since damping dominates over driving. After a perturbation such a system is returned asymptotically to its equilibrium steady state by means of an exponential decay. Whereas systems with high quality factors (i.e. $Q > 1/\sqrt{2}$) are referred to as 'underdamped', which means they continue to oscillate at their resonant frequency whilst experiencing a decaying signal amplitude. This is because an underdamped system has driving that is able to overcome damping. For a perfect oscillation with zero damping the quality factor is infinite. Hence, coherent pulsators driven by the heat-engine mechanism (see e.g. Bowman 2020), would have extremely high Q values.

Given the measured Q values of the stars shown in Fig. 3, we conclude that SLF variability observed in massive stars is a mix of overdamped and underdamped oscillatory behaviour. Interestingly, the 'yellow' sub-group of stars have the highest Q values, the longest damping timescales, and $\nu_{\rm char}$ values that are comparable to theoretically predicted pulsation gravity mode frequencies in massive stars (i.e. $P \lesssim 1$ d). Therefore, irrespective of the exact excitation mechanism(s) of their observed SLF variability, such stars satisfy the characteristics of being underdamped oscillators. This lends support to the conclusion that SLF variability is caused by quasi-periodic gravity waves in massive stars (Bowman et al. 2019a,b).

3.3. Periodic or stochastic?

Two prime examples of stars in the 'yellow' sub-group are HD 152424 (TIC 247267245) and HD 154368 (TIC 41792209), which have $Q = 2.0^{+1.3}_{-0.8}$ and $Q = 2.2^{+1.0}_{-0.7}$, respectively (see Table B.1). We note that there are relatively large uncertainties for the Q values, the largest of all the stars, but firmly place the stars in the 'underdamped' regime. These stars have almost 'perfect' GP regression fits to their light curves (see Appendix A), and have light curves that are clearly dominated by astrophysical signal and not instrumental noise. Stars in the 'yellow' sub-group are those that deserve special attention for the purpose of identifying and extracting quasi-coherent pulsation modes on top of the SLF variability for the purpose of asteroseismology. However, such an attempt would require long-term light curves spanning dozens of TESS sectors to resolve individual modes. This is beyond the scope of our current paper but is the subject of future work (Bowman et al. in prep).

We also perform a second modelling setup and perform GP regression for all stars in which we fix the quality factor to be $Q = 1/\sqrt{2}$. As a consequence, the damping time scale, $\tau_{\rm damp}$, is no longer included as a parameter. This allows for a better understanding of the measured Q values when it is an inferred parameter and of the power spectral density predictions derived from the GP regression. We provide the parameters from the fixed-Q GP regression in Table B.2. Fixing Q as a parameter forces the shape of the power spectral density prediction from the GP regression to be similar to that of Eqn. 1. This has the potential to explain the differences in $\nu_{\rm char}$ values between those from Bowman et al. (2020) and from the non-fixed Q-value GP regression. However, the $\nu_{\rm char}$ values from the GP regressions with fixed and non-fixed Q values are within 1σ of each other, which demonstrates that we are somewhat insensitive to Q or there is degeneracy in these parameters.

3.4. Testing the impact of poorer quality observations

We have demonstrated that the best-fitting parameters to characterise the SLF variability in massive stars is sometimes dependent on the choice of methodology, but this depends on a particular star. More specifically, it depends on the quality of the observations in terms of which of either astrophysical signal or instrumental noise dominates. For the 'yellow' sub-group of stars, the astrophysical signal (i.e. SLF variability) strongly dominates over the noise. Whereas, some stars in the 'blue' subgroup have SLF variability that is comparable to the white noise, and only detected because of the high precision and long duration of TESS light curves. With poorer quality and/or noisier light curves, the SLF variability with peak-to-peak amplitudes of order mmag of stars in the 'blue' subgroup may not be detected at all.

On the other hand, for the sample at large, there is fairly good agreement between the results in this work and the alternative

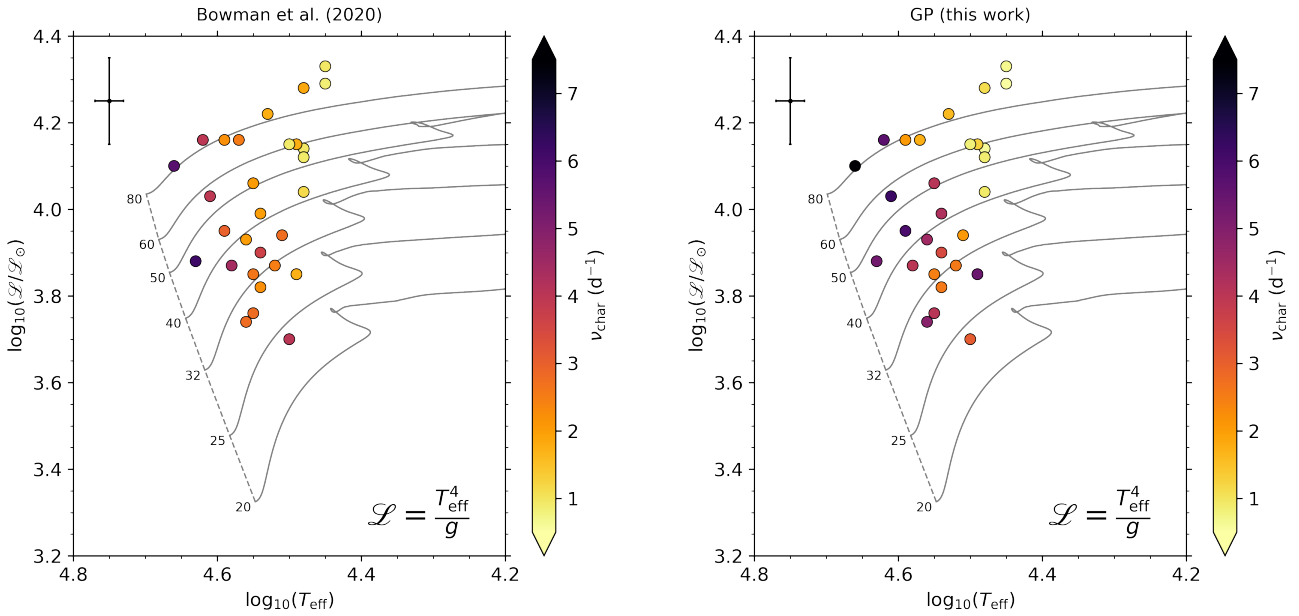

Fig. 2. Spectroscopic HR diagram of 30 O stars determined by Bowman et al. (2020) to be dominated by SLF variability in their TESS light curves (*left panel*) that we re-model using GP regression in this work (*right panel*), which are colour-coded by their corresponding $\nu_{\rm char}$ values. Evolutionary tracks from Burssens et al. (2020) are shown as solid grey lines and labelled in units of $M_\odot$, and the dashed grey line represents the ZAMS. A typical spectroscopic error bar for the sample is shown in the top-left corner.

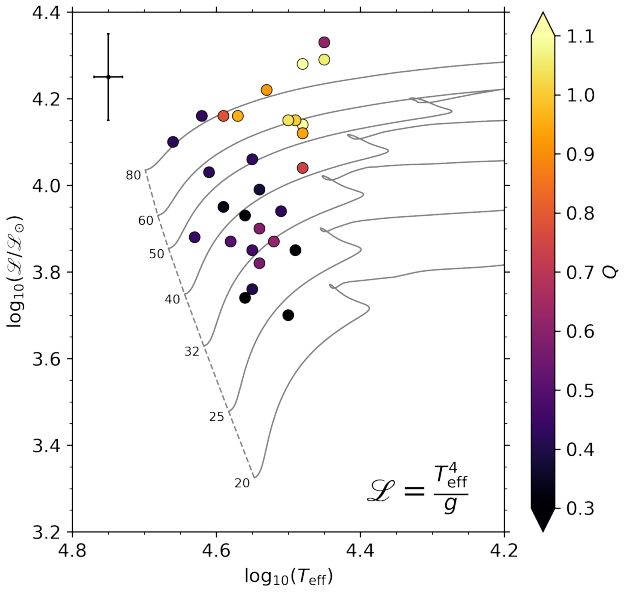

Fig. 3. Spectroscopic HR diagram of the same O stars as in Fig. 2 but colour-coded by the quality factor, Q, of the variability as determined in our GP regression.

approach taken previously by Bowman et al. (2020). This leads us to wonder which of the two approaches is more sensitive to the quality of the observations and perhaps less robust considering the notable difference in $\nu_{\rm char}$ values for some stars. For example, not all stars in our sample are the same brightness, such that the underlying random and instrumental noise contributions in their light curves are different. Perhaps the GP regression of a light curve in the time domain is more sensitive to this compared to the amplitude spectrum fitting in the Fourier domain.

As a first test, we artificially normalised the standard deviation of each TESS light curve to be unity and repeated our GP regression framework. The relationship between the original and 'scaled' $\nu_{\rm char}$ values is shown in Fig. 4, which clearly shows that all stars fall on the 1-to-1 line. Fractional differences of less than 1% are observed for all 30 stars in their original and 'scaled' $\nu_{\rm char}$ values, but these are (much) smaller than the formal uncertainties. Therefore, the GP regression methodology is not dependent on the standard deviation of the input SLF variability in a relative sense as long as it dominates over the white noise within the light curve. In other words, it is not the case that stars with higher amplitude SLF variability have more robust results in terms of their $\nu_{\rm char}$ values.

To further test the sensitivity of the GP regression to the quality of the input time series, we degraded the quality of the TESS light curves of the 30 massive stars by adding additional frequency independent (i.e. white) noise. This increased the value of $C_{\rm w}$ (cf. Eqn. 1) in the amplitude spectrum and consequently also the contribution of the jitter term in the GP regression. For each star in our sample, we created four additional light curves which had 0.5, 1, 2 and 3 times the standard deviation of the original light curve added to them as white noise to emulate increasingly noisier data sets. The latter of which can be considered an extremely and unrealistic noise-dominated light curve, but we include it for completeness. We then re-modelled these noisy light curves using our GP regression methodology and compare the results.

In Fig. 5, we show examples of the results of the GP regression applied to the 'yellow' (left) and 'blue' (right) subgroup members, HD 112244 (TIC 406050497) and HD 96715 (TIC 306491594), respectively. The GP regression typically returns the same $\nu_{\rm char}$ value for all noisy light curves of a star within the respective uncertainties. Whereas, there is consider-

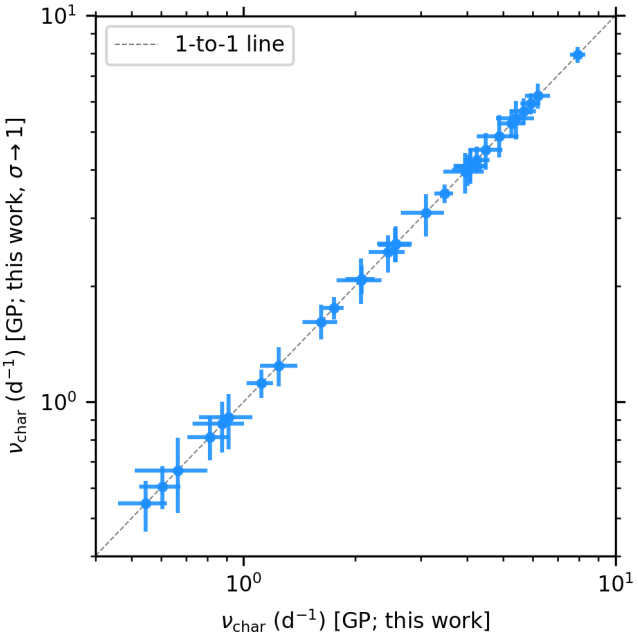

Fig. 4. Demonstration of negligible difference between $\nu_{\rm char}$ values if the standard deviation of the input light curve, σ, is normalised to unity before applying the GP regression.

ably larger scatter in $\nu_{\rm char}$ values determined from the amplitude spectrum fitting methodology of Bowman et al. (2020) to the same artificially noise-enhanced data sets. The values of $\nu_{\rm char}$ of the latter are significantly different from each other because of the typically much smaller uncertainties returned by the method of amplitude spectrum fitting. In almost all cases, and as demonstrated by the example stars in Fig. 5, the $\nu_{\rm char}$ returned by the amplitude spectrum fitting methodology is pushed to lower values when the white noise contribution is larger. Whereas, the $\nu_{\rm char}$ value from the GP regression stays the same within its uncertainties, hence not significantly different among the noise-enhanced data sets given its relatively large formal uncertainties.

Given the consistency of the $\nu_{\rm char}$ values of the GP regression, regardless of the underlying signal-to-noise ratio of SLF variability over the white noise, we conclude that it is the more robust methodology. The reason for this is because of the fundamental difference between the two fitting methods. The GP regression is a non-parametric fit of the light curve, in which the implemented jitter term effectively handles the included white noise. On the other hand, a fit of the amplitude spectrum is implicitly dependent on the Fourier noise floor and the spectral window, which impact the determination of the correlated $\nu_{\rm char}$ parameter (Bowman et al. 2019a). Therefore, large differences in the $C_{\rm w}$ term in Eqn. (1) can lead to significantly different $\nu_{\rm char}$ values, especially given the small uncertainties typically returned from this method.

4. Discussion and Conclusions

The vast majority of massive stars have SLF variability in time series photometric observations, if not all when sufficiently high enough quality data are available, which probes the physical properties of the star, such as mass and age. In this work, we have tested the use of GP regression applied to the time series photometric observations for a sample of 30 massive stars observed by the TESS mission and previously studied by Bowman et al. (2020). We use the same input data and compare our GP regression results, paying particular attention to the characteristic frequency of the SLF variability, $\nu_{\rm char}$, in comparison to the previous MCMC fitting of the amplitude spectrum by Bowman et al. (2020). We find that the majority of stars have the same $\nu_{\rm char}$ value within the uncertainties. This is encouraging and demonstrates that the previously found correlation of $\nu_{\rm char}$ with mass and age is robust given that the GP regression method is a fit to the light curve, and the method of Bowman et al. (2020) is a fit to the amplitude spectrum. A minority of stars have different and systematically higher values of $\nu_{\rm char}$ by up to a factor two when using the GP regression compared to amplitude spectrum fitting. The cause of this is the difference in the two fitting methodologies. Specifically, the power spectral density of the GP regression model is able to, in some cases, better match the shape of the SLF variability in the Fourier domain compared to the more rigid shape imposed by the standard semi-Lorentzian profile given in Eqn. (1).

We have demonstrated that GP regression is a viable method for characterising SLF variability in time series photometry of massive stars. Our work lends further support to the conclusion of Bowman et al. (2020) that SLF variability probes the mass and age of a massive star, demonstrated by the correlation between $\nu_{\rm char}$ and a star's location in the HR diagram, as shown in Fig. 2. Therefore, we conclude that the absence of coherent, heat-driven pulsation modes is not necessarily a prerequisite for asteroseismology of massive stars, given the probing power of $\nu_{\rm char}$ on bulk properties like mass and age. Furthermore, we have shown that a fraction of massive stars with SLF variability exhibit signatures of underdamped oscillatory behaviour, suggesting the presence of quasi-periodic pulsations. Although a range of quality factors exist in the SLF variability of massive stars, a transition from overdamped to underdamped oscillatory behaviour is found to coincide with a transition from the ZAMS to the TAMS, as shown in Fig. 3. We have also demonstrated that a GP regression is more robust against noisier input data time series, especially in the cases where the SLF variability is comparable in amplitude to the white noise.

In the future, we will apply our methodology to additional stars observed by TESS to extract the morphology of their SLF variability. This will allow us to further populate Fig. 2 and study the origin of SLF variability and its correlation to stellar properties such as mass and age, but also macroturbulence as shown by Bowman et al. (2020). The TESS mission is ongoing and currently providing high precision light curves of thousands of massive stars, which also span much longer time spans than those considered in this initial study. In our next study, we will expand our method to apply it to a much larger sample and use it infer masses and radii of the massive stars.

Acknowledgements. The authors acknowledge that this work was partially conducted on the traditional land of the first people of Seattle, the Duwamish People past and present and honours with gratitude the land itself and the Duwamish Tribe. DMB gratefully acknowledges funding from the Research Foundation Flanders (FWO) by means of a senior postdoctoral fellowship (grant agreement No. 1286521N) and an FWO long stay travel grant (agreement No. V411621N). DMB is also grateful to the staff and scientists at the Kavli Institute for Theoretical Physics, University of California, Santa Barbara, and to the organisers of the 'Probes of Transport in Stars' program, which hosted him during when part of this work was conducted. This research was supported in part by the National Science Foundation (NSF) under Grant Number NSF PHY-1748958. TZD-W gratefully acknowledges support from Grant Number NSF AST-1714285. The authors thank the TESS science team for the excellent data, which were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support to MAST for these data is provided by the NASA Office of Space Science via grant NAG5-7584 and by other grants and contracts. Funding for the TESS mission

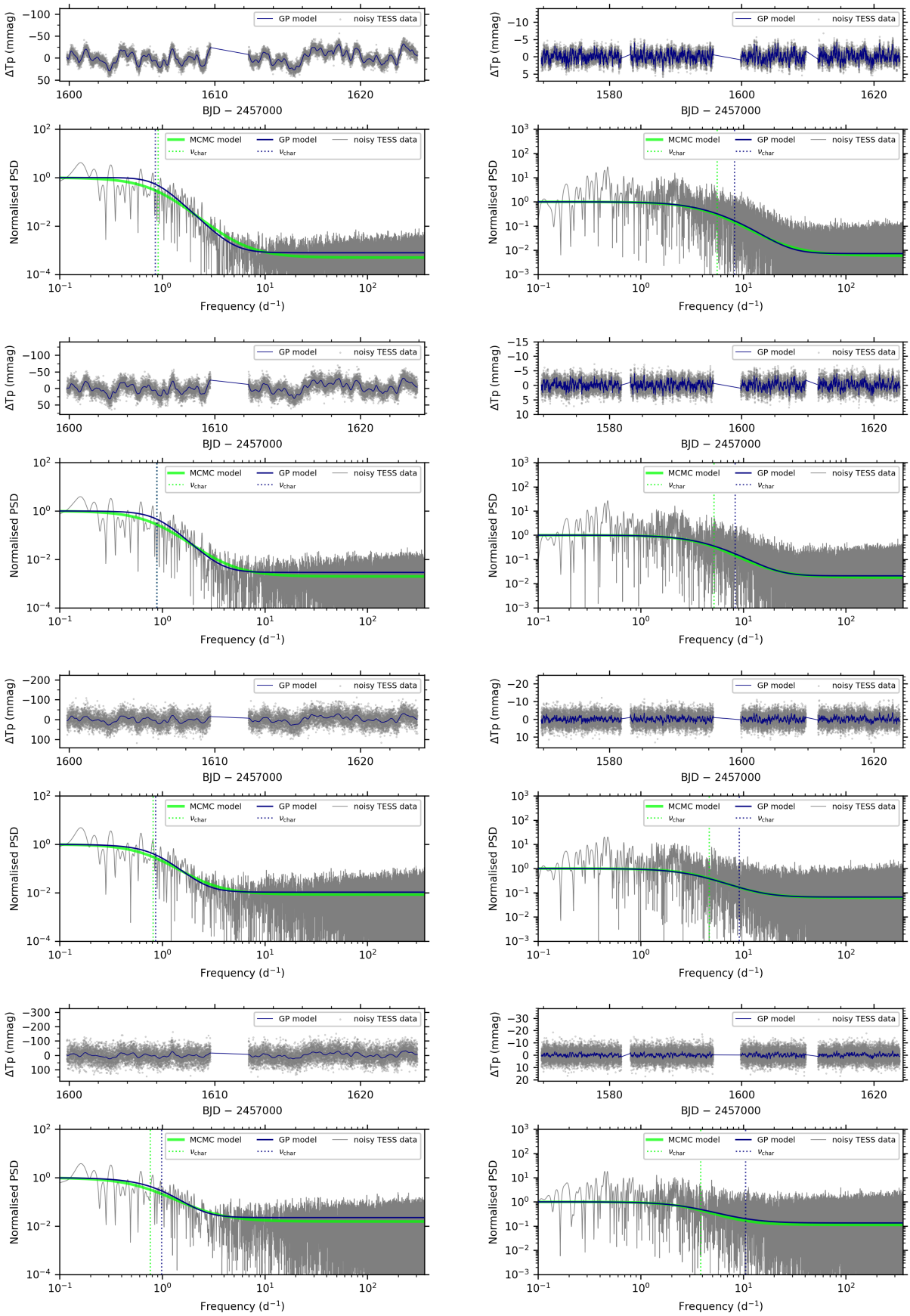

Fig. 5. Results for the noise-enhanced light curves of HD 112244 (TIC 406050497) and HD 96715 (TIC 306491594), which are typical 'yellow' and 'blue' subgroup members, are shown in the left and right columns, respectively. From top to bottom, increasingly larger amounts of white noise has been added. The $\nu_{\rm char}$ values from the GP regression are more consistent, on average, than those from fitting the amplitude spectrum.

is provided by the NASA Explorer Program. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France; the SAO/NASA Astrophysics Data System; and the VizieR catalog access tool, CDS, Strasbourg, France. This research has made use of the following `python` software packages: `MATPLOTLIB` (Hunter 2007), `Seaborn` (Waskom 2021), `Numpy` (Oliphant 2006; van der Walt et al. 2011; Harris et al. 2020), `astropy` (Astropy Collaboration et al. 2013, 2018), `SciPy` (Oliphant 2007; Virtanen et al. 2020), `Pandas` (Wes McKinney 2010), `celerite2` (Foreman-Mackey et al. 2017; Foreman-Mackey 2018; Foreman-Mackey et al. 2019a), `exoplanet` (Foreman-Mackey et al. 2019b, 2021), `arviz` (Kumar et al. 2019), and `pymc3` (Salvatier et al. 2016).

References

Aerts, C. 2021, Reviews of Modern Physics, 93, 015001
Aerts, C., Christensen-Dalsgaard, J., & Kurtz, D. W. 2010, Asteroseismology (Springer)
Aerts, C. & Rogers, T. M. 2015, ApJ, 806, L33
Aerts, C., Thoul, A., Daszyńska, J., et al. 2003, Science, 300, 1926
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Auvergne, M., Bodin, P., Boisnard, L., et al. 2009, A&A, 506, 411
Barclay, T., Endl, M., Huber, D., et al. 2015, ApJ, 800, 46
Blomme, R., Mahy, L., Catala, C., et al. 2011, A&A, 533, A4
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977
Bowman, D. M. 2017, Amplitude Modulation of Pulsation Modes in Delta Scuti Stars (Springer International Publishing)
Bowman, D. M. 2020, Frontiers in Astronomy and Space Sciences, 7, 70
Bowman, D. M., Aerts, C., Johnston, C., et al. 2019a, A&A, 621, A135
Bowman, D. M., Burssens, S., Pedersen, M. G., et al. 2019b, Nature Astronomy, 3, 760
Bowman, D. M., Burssens, S., Simón-Díaz, S., et al. 2020, A&A, 640, A36
Bowman, D. M. & Michielsen, M. 2021, A&A, 656, A158
Brewer, B. J. & Stello, D. 2009, MNRAS, 395, 2226
Burssens, S., Bowman, D. M., Michielsen, M., et al. 2022, submitted
Burssens, S., Simón-Díaz, S., Bowman, D. M., et al. 2020, A&A, 639, A81
Cantiello, M., Langer, N., Brott, I., et al. 2009, A&A, 499, 279
Cantiello, M., Lecoanet, D., Jermyn, A. S., & Grassitelli, L. 2021, ApJ, 915, 112
Daszyńska-Daszkiewicz, J., Szewczuk, W., & Walczak, P. 2013, MNRAS, 431, 3396
Deeming, T. J. 1975, Ap&SS, 36, 137
Degroote, P., Aerts, C., Baglin, A., et al. 2010, Nature, 464, 259
Degroote, P., Aerts, C., Ollivier, M., et al. 2009a, A&A, 506, 471
Degroote, P., Briquet, M., Catala, C., et al. 2009b, A&A, 506, 111
Dorn-Wallenstein, T. Z., Levesque, E. M., & Davenport, J. R. A. 2019, ApJ, 878, 155
Dorn-Wallenstein, T. Z., Levesque, E. M., Neugent, K. F., et al. 2020, ApJ, 902, 24
Dziembowski, W. A., Moskalik, P., & Pamyatnykh, A. A. 1993, MNRAS, 265, 588
Dziembowski, W. A. & Pamyatnykh, A. A. 1993, MNRAS, 262, 204
Dziembowski, W. A. & Pamyatnykh, A. A. 2008, MNRAS, 385, 2061
Edelmann, P. V. F., Ratnasingam, R. P., Pedersen, M. G., et al. 2019, ApJ, 876, 4
Elliott, A., Richardson, N. D., Pablo, H., et al. 2022, MNRAS, 509, 4246
Foreman-Mackey, D. 2018, Research Notes of the American Astronomical Society, 2, 31
Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, AJ, 154, 220
Foreman-Mackey, D., Agol, E., Angus, R., et al. 2019a, dfm/celerite: celerite v0.3.1
Foreman-Mackey, D., Czekala, I., Agol, E., Luger, R., & Barclay, T. 2019b, dfm/exoplanet: exoplanet v0.2.4
Foreman-Mackey, D., Farr, W., Sinha, M., et al. 2019c, The Journal of Open Source Software, 4, 1864
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Foreman-Mackey, D., Luger, R., Agol, E., et al. 2021, The Journal of Open Source Software, 6, 3285
Gelman, A. & Rubin, D. B. 1992, Statistical Science, 7, 457
Grunblatt, S. K., Huber, D., Gaidos, E., et al. 2017, AJ, 154, 254
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Horst, L., Edelmann, P. V. F., Andrássy, R., et al. 2020, A&A, 641, A18
Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398
Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90
Jermyn, A. S., Anders, E. H., & Cantiello, M. 2022, arXiv e-prints, arXiv:2201.10567
Jiang, Y.-F., Cantiello, M., Bildsten, L., Quataert, E., & Blaes, O. 2015, ApJ, 813, 74
Jiang, Y.-F., Cantiello, M., Bildsten, L., et al. 2018, Nature, 561, 498
Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, ApJ, 713, L79
Krtička, J. & Feldmeier, A. 2018, A&A, 617, A121
Krtička, J. & Feldmeier, A. 2021, A&A, 648, A79
Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. 2019, Journal of Open Source Software, 4, 1143
Kurtz, D. W. 1985, MNRAS, 213, 773
Lecoanet, D., Cantiello, M., Anders, E. H., et al. 2021, MNRAS, 508, 132
Lecoanet, D., Cantiello, M., Quataert, E., et al. 2019, ApJ, 886, L15
Michielsen, M., Aerts, C., & Bowman, D. M. 2021, A&A, 650, A175
Miglio, A., Montalbán, J., & Dupret, M.-A. 2007, MNRAS, 375, L21
Moravveji, E., Aerts, C., Pápics, P. I., Triana, S. A., & Vandoren, B. 2015, A&A, 580, A27
Moravveji, E., Townsend, R. H. D., Aerts, C., & Mathis, S. 2016, ApJ, 823, 130
Nazé, Y., Rauw, G., & Gosset, E. 2021, MNRAS, 502, 5038
Oliphant, T. E. 2006, A guide to NumPy, Vol. 1 (Trelgol Publishing USA)
Oliphant, T. E. 2007, Computing in Science Engineering, 9, 10
Pamyatnykh, A. A., Handler, G., & Dziembowski, W. A. 2004, MNRAS, 350, 1022
Pedersen, M. G., Aerts, C., Pápics, P. I., et al. 2021, Nature Astronomy, 5, 715
Pedersen, M. G., Chowdhury, S., Johnston, C., et al. 2019, ApJ, 872, L9
Pereira, F., Campante, T. L., Cunha, M. S., et al. 2019, MNRAS, 489, 5764
Press, W. H. & Rybicki, G. B. 1989, ApJ, 338, 277
Rasmussen, C. E. & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning (Massachusetts Institute of Technology)
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
Rogers, T. M. 2015, ApJ, 815, L30
Rogers, T. M., Lin, D. N. C., McElwaine, J. N., & Lau, H. H. B. 2013, ApJ, 772, 21
Rogers, T. M. & McElwaine, J. N. 2017, ApJ, 848, L1
Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. 2016, PeerJ Computer Science, 2, e55
Scargle, J. D. 1982, ApJ, 263, 835
Schultz, W. C., Bildsten, L., & Jiang, Y.-F. 2022, ApJ, 924, L11
Simón-Díaz, S., Castro, N., Garcia, M., Herrero, A., & Markova, N. 2011, Bulletin de la Societe Royale des Sciences de Liege, 80, 514
Szewczuk, W. & Daszyńska-Daszkiewicz, J. 2017, MNRAS, 469, 13
Szewczuk, W. & Daszyńska-Daszkiewicz, J. 2018, MNRAS, 478, 2243
Szewczuk, W., Walczak, P., & Daszyńska-Daszkiewicz, J. 2021, MNRAS, 503, 5894
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science Engineering, 13, 22
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261
Waskom, M. L. 2021, Journal of Open Source Software, 6, 3021
Wes McKinney. 2010, in Proceedings of the 9th Python in Science Conference, ed. Stéfan van der Walt & Jarrod Millman, 56 – 61

Appendix A: Fitted amplitude spectra figures

Summary figures for the 30 massive stars studied in this work containing the TESS light curve fitted using a GP regression, the normalised power spectral density of the GP regression model, TESS data and the corresponding best-fit model determined by Bowman et al. (2020) are given in Figs. A.1, A.2, and A.3.

Appendix B: Extended data tables

The full parameter list of GP regression parameters including uncertainties determined by a Hamiltonian no U-turn MC sampler are given in Table B.1 for the 30 massive stars studied in this work. In Table B.2, we provide the GP regression parameters for when a fixed value of $Q = 1/\sqrt{2}$ is used.

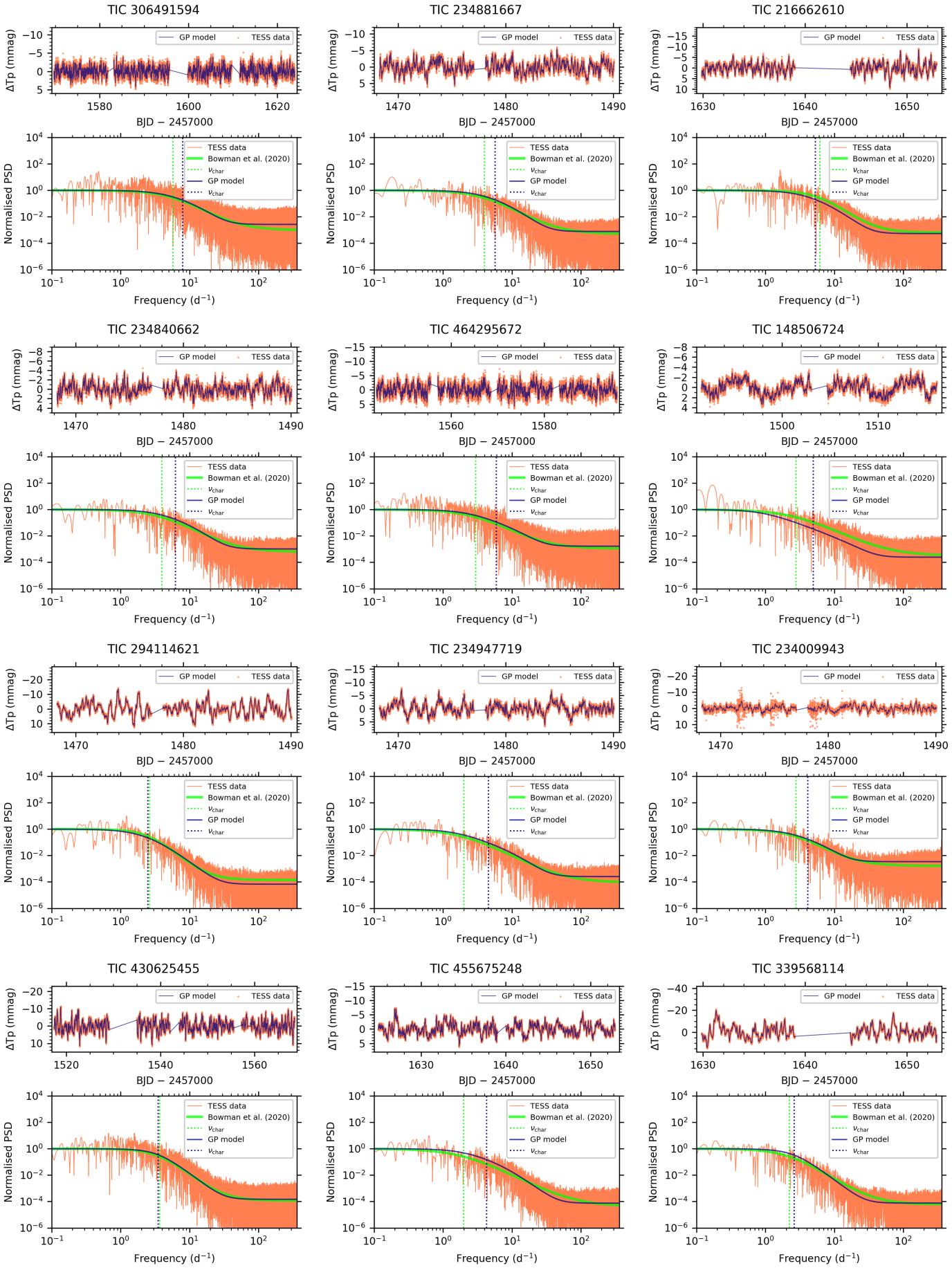

Fig. A.1. Summary figures for massive stars analysed in this work.

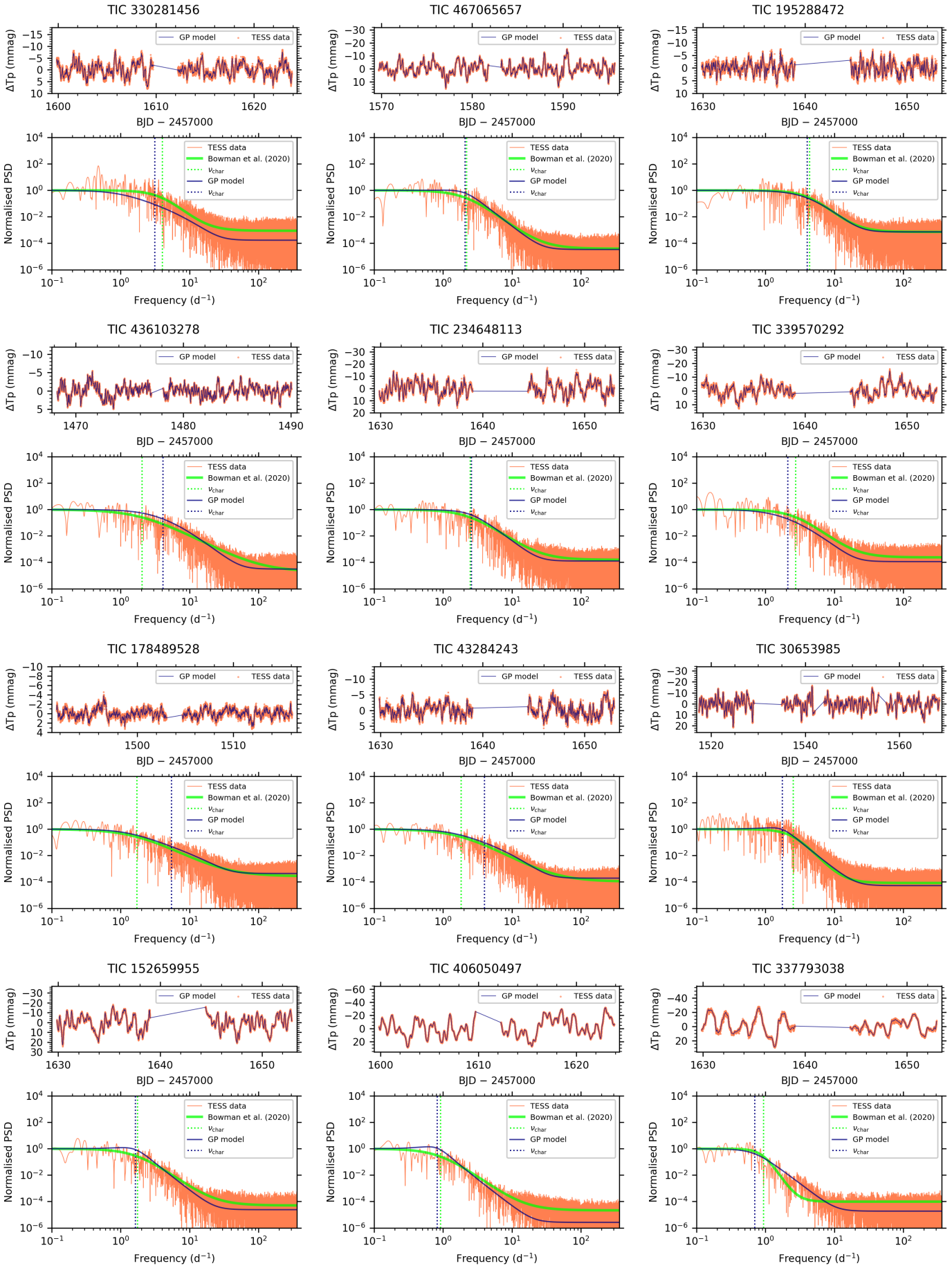

Fig. A.2. Summary figures for massive stars analysed in this work.

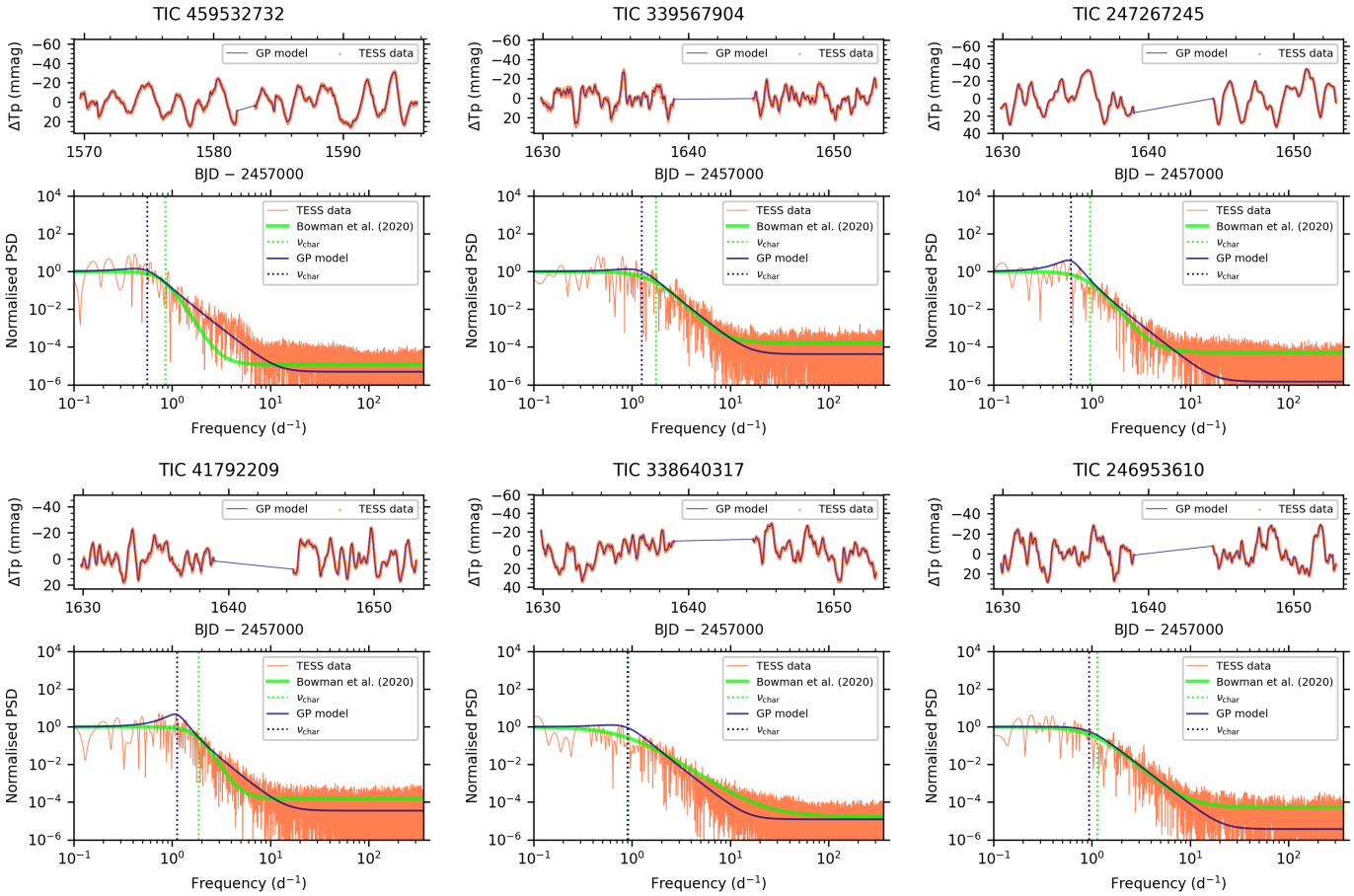

Fig. A.3. Summary figures for massive stars analysed in this work.

Table B.1. GP regression parameters for sample of massive stars studied in this work.

Name	TIC	Bowman et al. (2020)	GP regression (this work)						
		$\nu_{\rm char}$ ($\rm d^{-1}$)	$\nu_{\rm char}$ ($\rm d^{-1}$)	$\rho_{\rm char}$ (d)	$\tau_{\rm damp}$ (d)	$\nu_{\rm damp}$ ($\rm d^{-1}$)	σ (mag)	Q	ln(jitter)
HD 96715	306491594	5.74308 ± 0.00368	$7.94^{+0.38}_{-0.39}$	$0.126^{+0.006}_{-0.006}$	$0.017^{+0.002}_{-0.002}$	$59.4^{+6.1}_{-5.9}$	$0.00124^{+0.00006}_{-0.00006}$	$0.42^{+0.04}_{-0.04}$	$-7.619^{+0.008}_{-0.008}$
HD 46223	234881667	3.99087 ± 0.00135	$5.67^{+0.45}_{-0.44}$	$0.176^{+0.014}_{-0.013}$	$0.024^{+0.003}_{-0.003}$	$41.2^{+5.6}_{-5.4}$	$0.00174^{+0.00015}_{-0.00014}$	$0.43^{+0.06}_{-0.06}$	$-7.705^{+0.012}_{-0.011}$
HD 155913	216662610	6.15743 ± 0.00115	$5.26^{+0.46}_{-0.46}$	$0.190^{+0.018}_{-0.015}$	$0.027^{+0.003}_{-0.004}$	$37.3^{+4.7}_{-4.9}$	$0.00237^{+0.00022}_{-0.00020}$	$0.44^{+0.06}_{-0.07}$	$-7.671^{+0.014}_{-0.012}$
HD 46150	234840662	3.95351 ± 0.00181	$6.19^{+0.48}_{-0.50}$	$0.161^{+0.013}_{-0.012}$	$0.022^{+0.003}_{-0.003}$	$46.4^{+6.3}_{-6.4}$	$0.00120^{+0.00010}_{-0.00009}$	$0.42^{+0.06}_{-0.06}$	$-7.962^{+0.012}_{-0.012}$
HD 90273	464295672	2.96658 ± 0.00147	$5.92^{+0.40}_{-0.37}$	$0.169^{+0.011}_{-0.011}$	$0.018^{+0.002}_{-0.002}$	$56.2^{+6.6}_{-6.2}$	$0.00171^{+0.00010}_{-0.00010}$	$0.33^{+0.03}_{-0.04}$	$-7.249^{+0.008}_{-0.008}$
HD 53975	148506724	2.77860 ± 0.00194	$4.88^{+0.65}_{-0.66}$	$0.205^{+0.029}_{-0.027}$	$0.012^{+0.002}_{-0.002}$	$85.9^{+13.8}_{-13.1}$	$0.00126^{+0.00017}_{-0.00015}$	$0.18^{+0.03}_{-0.03}$	$-8.117^{+0.012}_{-0.012}$
HD 41997	294114621	2.61022 ± 0.00018	$2.44^{+0.26}_{-0.28}$	$0.409^{+0.048}_{-0.044}$	$0.060^{+0.009}_{-0.008}$	$16.7^{+2.5}_{-2.3}$	$0.00446^{+0.00058}_{-0.00052}$	$0.46^{+0.09}_{-0.07}$	$-7.613^{+0.011}_{-0.011}$
HD 46573	234947719	2.00862 ± 0.00051	$4.48^{+0.47}_{-0.48}$	$0.223^{+0.024}_{-0.023}$	$0.021^{+0.003}_{-0.003}$	$47.2^{+6.3}_{-6.2}$	$0.00201^{+0.00023}_{-0.00021}$	$0.30^{+0.05}_{-0.05}$	$-7.832^{+0.013}_{-0.011}$
HD 48279	234009943	2.77586 ± 0.00089	$4.09^{+0.43}_{-0.43}$	$0.244^{+0.029}_{-0.023}$	$0.032^{+0.007}_{-0.006}$	$31.6^{+6.1}_{-5.9}$	$0.00155^{+0.00017}_{-0.00015}$	$0.41^{+0.07}_{-0.08}$	$-6.891^{+0.012}_{-0.011}$
HD 74920	430625455	3.66392 ± 0.00053	$3.47^{+0.20}_{-0.20}$	$0.288^{+0.016}_{-0.017}$	$0.054^{+0.005}_{-0.005}$	$18.4^{+1.7}_{-1.7}$	$0.00327^{+0.00021}_{-0.00022}$	$0.59^{+0.06}_{-0.06}$	$-7.902^{+0.009}_{-0.008}$
HD 135591	455675248	1.97700 ± 0.00066	$4.24^{+0.31}_{-0.36}$	$0.236^{+0.020}_{-0.018}$	$0.029^{+0.003}_{-0.003}$	$34.8^{+3.4}_{-3.2}$	$0.00189^{+0.00016}_{-0.00016}$	$0.38^{+0.05}_{-0.05}$	$-8.588^{+0.010}_{-0.012}$
HD 326331	339568114	2.20987 ± 0.00017	$2.57^{+0.27}_{-0.26}$	$0.389^{+0.042}_{-0.039}$	$0.070^{+0.011}_{-0.011}$	$14.2^{+2.4}_{-2.1}$	$0.00541^{+0.00069}_{-0.00059}$	$0.57^{+0.12}_{-0.10}$	$-7.599^{+0.013}_{-0.013}$
HD 123056	330281456	4.01672 ± 0.00065	$3.09^{+0.40}_{-0.40}$	$0.324^{+0.047}_{-0.037}$	$0.029^{+0.005}_{-0.004}$	$34.2^{+5.2}_{-5.2}$	$0.00270^{+0.00040}_{-0.00032}$	$0.28^{+0.05}_{-0.05}$	$-7.567^{+0.011}_{-0.012}$
HD 97253	467065657	2.20426 ± 0.00019	$2.08^{+0.16}_{-0.17}$	$0.481^{+0.040}_{-0.037}$	$0.120^{+0.019}_{-0.019}$	$8.3^{+1.4}_{-1.3}$	$0.00451^{+0.00051}_{-0.00043}$	$0.78^{+0.15}_{-0.13}$	$-8.122^{+0.010}_{-0.011}$
HD 156738	195288472	4.38608 ± 0.00054	$4.01^{+0.36}_{-0.37}$	$0.249^{+0.024}_{-0.022}$	$0.040^{+0.006}_{-0.006}$	$24.8^{+3.9}_{-3.8}$	$0.00233^{+0.00024}_{-0.00023}$	$0.51^{+0.09}_{-0.09}$	$-7.476^{+0.013}_{-0.012}$
HD 36861	436103278	2.05434 ± 0.00077	$4.07^{+0.36}_{-0.35}$	$0.246^{+0.023}_{-0.020}$	$0.034^{+0.003}_{-0.003}$	$29.3^{+2.9}_{-2.9}$	$0.00160^{+0.00016}_{-0.00014}$	$0.44^{+0.06}_{-0.06}$	$-9.249^{+0.013}_{-0.012}$
HD 150574	234648113	2.47023 ± 0.00017	$2.56^{+0.25}_{-0.27}$	$0.391^{+0.040}_{-0.040}$	$0.077^{+0.014}_{-0.012}$	$13.1^{+2.2}_{-2.2}$	$0.00568^{+0.00074}_{-0.00060}$	$0.61^{+0.12}_{-0.12}$	$-7.340^{+0.013}_{-0.012}$
HD 152247	339570292	2.75008 ± 0.00027	$2.07^{+0.29}_{-0.28}$	$0.483^{+0.075}_{-0.060}$	$0.066^{+0.011}_{-0.011}$	$15.2^{+2.6}_{-2.5}$	$0.00442^{+0.00071}_{-0.00066}$	$0.43^{+0.10}_{-0.09}$	$-7.368^{+0.012}_{-0.013}$
HD 55879	178489528	1.71966 ± 0.00108	$5.41^{+0.64}_{-0.62}$	$0.185^{+0.023}_{-0.020}$	$0.013^{+0.002}_{-0.002}$	$79.7^{+12.0}_{-12.8}$	$0.00091^{+0.00011}_{-0.00009}$	$0.21^{+0.03}_{-0.03}$	$-8.286^{+0.011}_{-0.012}$
HD 154643	43284243	1.83226 ± 0.00041	$3.96^{+0.48}_{-0.47}$	$0.253^{+0.033}_{-0.029}$	$0.024^{+0.003}_{-0.003}$	$41.5^{+6.2}_{-5.4}$	$0.00208^{+0.00029}_{-0.00023}$	$0.30^{+0.05}_{-0.05}$	$-7.992^{+0.013}_{-0.013}$
CPD-47 2963	30653985	2.53356 ± 0.00019	$1.75^{+0.12}_{-0.11}$	$0.571^{+0.038}_{-0.036}$	$0.174^{+0.026}_{-0.023}$	$5.7^{+0.8}_{-0.8}$	$0.00527^{+0.00049}_{-0.00043}$	$0.96^{+0.15}_{-0.14}$	$-7.841^{+0.008}_{-0.008}$
HD 156154	152659955	1.76469 ± 0.00008	$1.62^{+0.17}_{-0.18}$	$0.618^{+0.075}_{-0.059}$	$0.181^{+0.043}_{-0.035}$	$5.5^{+1.3}_{-1.1}$	$0.00790^{+0.00120}_{-0.00104}$	$0.92^{+0.25}_{-0.21}$	$-7.811^{+0.013}_{-0.012}$
HD 112244	406050497	0.91870 ± 0.00003	$0.81^{+0.12}_{-0.10}$	$1.236^{+0.176}_{-0.156}$	$0.408^{+0.142}_{-0.097}$	$2.4^{+0.7}_{-0.6}$	$0.01374^{+0.00285}_{-0.00249}$	$1.03^{+0.39}_{-0.28}$	$-8.012^{+0.012}_{-0.010}$
HD 151804	337793038	0.93978 ± 0.00001	$0.67^{+0.13}_{-0.15}$	$1.499^{+0.390}_{-0.275}$	$0.296^{+0.103}_{-0.084}$	$3.4^{+1.1}_{-1.0}$	$0.01138^{+0.00323}_{-0.00261}$	$0.62^{+0.25}_{-0.22}$	$-7.006^{+0.013}_{-0.012}$
HD 303492	459532732	0.85247 ± 0.00001	$0.54^{+0.08}_{-0.09}$	$1.836^{+0.327}_{-0.261}$	$0.613^{+0.233}_{-0.174}$	$1.6^{+0.5}_{-0.5}$	$0.01316^{+0.00329}_{-0.00266}$	$1.05^{+0.41}_{-0.35}$	$-7.540^{+0.011}_{-0.010}$
HD 152249	339567904	1.76168 ± 0.00005	$1.24^{+0.15}_{-0.15}$	$0.804^{+0.101}_{-0.090}$	$0.256^{+0.073}_{-0.060}$	$3.9^{+0.9}_{-1.0}$	$0.00931^{+0.00170}_{-0.00142}$	$1.00^{+0.30}_{-0.27}$	$-7.287^{+0.012}_{-0.013}$
HD 152424	247267245	0.95329 ± 0.00001	$0.61^{+0.07}_{-0.08}$	$1.652^{+0.229}_{-0.188}$	$1.058^{+0.662}_{-0.426}$	$0.9^{+0.4}_{-0.5}$	$0.01715^{+0.00550}_{-0.00383}$	$2.00^{+1.31}_{-0.84}$	$-8.352^{+0.012}_{-0.012}$
HD 154368	41792209	1.85794 ± 0.00004	$1.12^{+0.08}_{-0.11}$	$0.894^{+0.086}_{-0.067}$	$0.612^{+0.275}_{-0.199}$	$1.6^{+0.6}_{-0.6}$	$0.00857^{+0.00188}_{-0.00145}$	$2.15^{+1.02}_{-0.71}$	$-7.775^{+0.012}_{-0.012}$
HD 152003	338640317	0.92626 ± 0.00003	$0.88^{+0.13}_{-0.13}$	$1.139^{+0.195}_{-0.160}$	$0.341^{+0.117}_{-0.087}$	$2.9^{+0.9}_{-0.9}$	$0.01354^{+0.00297}_{-0.00256}$	$0.94^{+0.38}_{-0.29}$	$-7.339^{+0.013}_{-0.011}$
HD 152147	246953610	1.13960 ± 0.00002	$0.91^{+0.13}_{-0.16}$	$1.095^{+0.200}_{-0.155}$	$0.259^{+0.070}_{-0.067}$	$3.9^{+1.0}_{-1.0}$	$0.01296^{+0.00294}_{-0.00233}$	$0.74^{+0.25}_{-0.21}$	$-7.884^{+0.012}_{-0.014}$

Table B.2. GP regression parameters for sample of massive stars studied in this work when fixed $Q = 1/\sqrt{2}$.

Name	TIC	Bowman et al. (2020)	GP regression (this work)			
		$\nu_{\rm char}$ ($\rm d^{-1}$)	$\nu_{\rm char}$ ($\rm d^{-1}$)	$\rho_{\rm char}$ (d)	σ (mag)	ln(jitter)
HD 96715	306491594	5.74308 ± 0.00368	$7.97^{+0.30}_{-0.28}$	$0.125^{+0.005}_{-0.005}$	$0.00124^{+0.00005}_{-0.00005}$	$-7.613^{+0.008}_{-0.008}$
HD 46223	234881667	3.99087 ± 0.00135	$5.98^{+0.35}_{-0.34}$	$0.167^{+0.010}_{-0.009}$	$0.00173^{+0.00012}_{-0.00012}$	$-7.702^{+0.012}_{-0.011}$
HD 155913	216662610	6.15743 ± 0.00115	$5.62^{+0.35}_{-0.35}$	$0.178^{+0.011}_{-0.011}$	$0.00235^{+0.00020}_{-0.00017}$	$-7.667^{+0.012}_{-0.013}$
HD 46150	234840662	3.95351 ± 0.00181	$6.44^{+0.35}_{-0.37}$	$0.155^{+0.009}_{-0.008}$	$0.00120^{+0.00009}_{-0.00008}$	$-7.958^{+0.012}_{-0.012}$
HD 90273	464295672	2.96658 ± 0.00147	$5.91^{+0.25}_{-0.24}$	$0.169^{+0.007}_{-0.007}$	$0.00171^{+0.00008}_{-0.00008}$	$-7.242^{+0.007}_{-0.008}$
HD 53975	148506724	2.77860 ± 0.00194	$5.25^{+0.31}_{-0.30}$	$0.191^{+0.011}_{-0.012}$	$0.00125^{+0.00009}_{-0.00008}$	$-8.103^{+0.012}_{-0.011}$
HD 41997	294114621	2.61022 ± 0.00018	$2.70^{+0.20}_{-0.21}$	$0.371^{+0.031}_{-0.026}$	$0.00443^{+0.00047}_{-0.00041}$	$-7.611^{+0.012}_{-0.011}$
HD 46573	234947719	2.00862 ± 0.00051	$4.97^{+0.30}_{-0.30}$	$0.201^{+0.013}_{-0.012}$	$0.00199^{+0.00016}_{-0.00014}$	$-7.825^{+0.012}_{-0.012}$
HD 48279	234009943	2.77586 ± 0.00089	$4.15^{+0.33}_{-0.34}$	$0.241^{+0.018}_{-0.020}$	$0.00154^{+0.00013}_{-0.00013}$	$-6.889^{+0.012}_{-0.011}$
HD 74920	430625455	3.66392 ± 0.00053	$3.61^{+0.17}_{-0.17}$	$0.277^{+0.013}_{-0.013}$	$0.00326^{+0.00022}_{-0.00019}$	$-7.901^{+0.008}_{-0.009}$
HD 135591	455675248	1.97700 ± 0.00066	$4.81^{+0.25}_{-0.24}$	$0.208^{+0.011}_{-0.010}$	$0.00189^{+0.00014}_{-0.00012}$	$-8.584^{+0.011}_{-0.011}$
HD 326331	339568114	2.20987 ± 0.00017	$2.71^{+0.24}_{-0.22}$	$0.369^{+0.033}_{-0.030}$	$0.00540^{+0.00060}_{-0.00063}$	$-7.598^{+0.012}_{-0.013}$
HD 123056	330281456	4.01672 ± 0.00065	$3.50^{+0.26}_{-0.23}$	$0.286^{+0.020}_{-0.020}$	$0.00267^{+0.00025}_{-0.00023}$	$-7.561^{+0.011}_{-0.011}$
HD 97253	467065657	2.20426 ± 0.00019	$2.03^{+0.16}_{-0.15}$	$0.493^{+0.039}_{-0.037}$	$0.00449^{+0.00056}_{-0.00045}$	$-8.122^{+0.011}_{-0.011}$
HD 156738	195288472	4.38608 ± 0.00054	$4.19^{+0.29}_{-0.30}$	$0.239^{+0.017}_{-0.016}$	$0.00232^{+0.00022}_{-0.00020}$	$-7.474^{+0.012}_{-0.012}$
HD 36861	436103278	2.05434 ± 0.00077	$4.59^{+0.26}_{-0.26}$	$0.218^{+0.013}_{-0.012}$	$0.00159^{+0.00013}_{-0.00013}$	$-9.247^{+0.013}_{-0.012}$
HD 150574	234648113	2.47023 ± 0.00017	$2.64^{+0.23}_{-0.21}$	$0.379^{+0.032}_{-0.031}$	$0.00565^{+0.00075}_{-0.00054}$	$-7.339^{+0.013}_{-0.012}$
HD 152247	339570292	2.75008 ± 0.00027	$2.31^{+0.20}_{-0.20}$	$0.434^{+0.038}_{-0.037}$	$0.00437^{+0.00055}_{-0.00048}$	$-7.366^{+0.013}_{-0.012}$
HD 55879	178489528	1.71966 ± 0.00108	$5.70^{+0.32}_{-0.35}$	$0.176^{+0.010}_{-0.010}$	$0.00091^{+0.00006}_{-0.00006}$	$-8.274^{+0.012}_{-0.011}$
HD 154643	43284243	1.83226 ± 0.00041	$4.52^{+0.30}_{-0.30}$	$0.221^{+0.015}_{-0.015}$	$0.00206^{+0.00019}_{-0.00016}$	$-7.986^{+0.013}_{-0.013}$
CPD-47 2963	30653985	2.53356 ± 0.00019	$1.62^{+0.11}_{-0.11}$	$0.616^{+0.042}_{-0.041}$	$0.00526^{+0.00053}_{-0.00048}$	$-7.842^{+0.008}_{-0.008}$
HD 156154	152659955	1.76469 ± 0.00008	$1.51^{+0.16}_{-0.17}$	$0.664^{+0.078}_{-0.067}$	$0.00788^{+0.00138}_{-0.00108}$	$-7.811^{+0.012}_{-0.013}$
HD 112244	406050497	0.91870 ± 0.00003	$0.72^{+0.11}_{-0.10}$	$1.388^{+0.221}_{-0.180}$	$0.01379^{+0.00310}_{-0.00260}$	$-8.012^{+0.011}_{-0.012}$
HD 151804	337793038	0.93978 ± 0.00001	$0.70^{+0.11}_{-0.12}$	$1.431^{+0.253}_{-0.219}$	$0.01105^{+0.00286}_{-0.00219}$	$-7.005^{+0.013}_{-0.012}$
HD 303492	459532732	0.85247 ± 0.00001	$0.49^{+0.08}_{-0.08}$	$2.031^{+0.356}_{-0.305}$	$0.01288^{+0.00330}_{-0.00283}$	$-7.540^{+0.011}_{-0.010}$
HD 152249	339567904	1.76168 ± 0.00005	$1.14^{+0.15}_{-0.13}$	$0.879^{+0.112}_{-0.105}$	$0.00923^{+0.00178}_{-0.00149}$	$-7.287^{+0.013}_{-0.012}$
HD 152424	247267245	0.95329 ± 0.00001	$0.43^{+0.09}_{-0.09}$	$2.308^{+0.581}_{-0.430}$	$0.01751^{+0.00672}_{-0.00459}$	$-8.352^{+0.011}_{-0.013}$
HD 154368	41792209	1.85794 ± 0.00004	$0.83^{+0.11}_{-0.12}$	$1.201^{+0.178}_{-0.165}$	$0.00855^{+0.00190}_{-0.00162}$	$-7.776^{+0.013}_{-0.013}$
HD 152003	338640317	0.92626 ± 0.00003	$0.81^{+0.11}_{-0.12}$	$1.241^{+0.200}_{-0.169}$	$0.01349^{+0.00319}_{-0.00253}$	$-7.339^{+0.012}_{-0.012}$
HD 152147	246953610	1.13960 ± 0.00002	$0.90^{+0.12}_{-0.13}$	$1.106^{+0.157}_{-0.151}$	$0.01283^{+0.00269}_{-0.00258}$	$-7.885^{+0.012}_{-0.012}$