

# Modelling stochastic and quasi-periodic behaviour in stellar time-series: Gaussian process regression versus power-spectrum fitting

Niamh K. O’Sullivan<sup>★</sup> and Suzanne Aigrain

*Denys Wilkinson Building, Department of Physics, University of Oxford, Oxford OX1 3RH, UK*

Accepted 2024 April 15. Received 2024 March 21; in original form 2024 January 8

## ABSTRACT

As the hunt for an Earth-like exoplanets has intensified in recent years, so has the effort to characterize and model the stellar signals that can hide or mimic small planetary signals. Stellar variability arises from a number of sources, including granulation, supergranulation, oscillations, and activity, all of which result in quasi-periodic or stochastic behaviour in photometric and/or radial velocity observations. Traditionally, the characterization of these signals has mostly been done in the frequency domain. However, the recent development of scalable Gaussian process regression methods makes direct time-domain modelling of stochastic processes a feasible and arguably preferable alternative, obviating the need to estimate the power spectral density of the data before modelling it. In this paper, we compare the two approaches using a series of experiments on simulated data. We show that frequency-domain modelling can lead to inaccurate results, especially when the time-sampling is irregular. By contrast, Gaussian process regression results are often more precise, and systematically more accurate, in both the regular and irregular time-sampling regimes. While this work was motivated by the analysis of radial velocity and photometry observations of main-sequence stars in the context of planet searches, we note that our results may also have applications for the study of other types of astrophysical variability such as quasi-periodic oscillations in X-ray binaries and active galactic nuclei variability.

**Key words:** asteroseismology – methods: data analysis – techniques: radial velocities – Sun: granulation.

## 1. INTRODUCTION

Stochastic processes are ubiquitous in astrophysical time-series observations. Quasi-periodic oscillations (QPOs) arise in accreting systems, from active galactic nuclei (AGNs) to compact binaries, while magnetic activity, convection, and oscillations in stellar interiors all give rise to stochastic variability ranging from strongly periodic to entirely aperiodic. The resulting time-series are in general too complex to describe using parametric models, but their power spectral density (PSD) can often be modelled using simple analytic functions or sums thereof. It has thus become common practice, in many astrophysical subfields, to model time-series observations displaying stochastic behaviour by estimating their PSDs and fitting the latter in the frequency domain.

The functions used to model the PSD are often generalized forms of the Lorentzian or Cauchy distribution:

$$P(\nu) = \frac{\alpha}{1 + \left( \frac{\nu - \nu_0}{\gamma} \right)^2}^\beta, \quad (1)$$

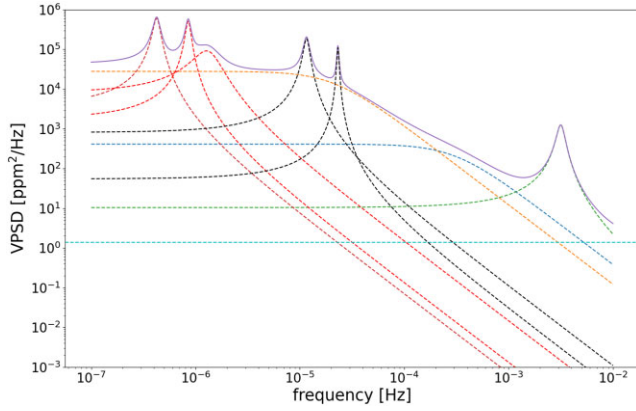
where  $\nu_0$  is the central frequency,  $\alpha$  is the amplitude,  $\beta$  is the slope, and  $\gamma$  is the width. (For a standard Cauchy or Lorentzian distribution normalized to unity,  $\beta = 2$  and  $\pi\gamma = \alpha^{-1}$ .) More complex models can be constructed by co-adding multiple components. The inferred parameters of the individual components are frequently given a physical interpretation, so it is important to understand how reliable these inferences are. Are there degeneracies between the parameters,

particularly when multiple components are included, and when the number of components is not known a priori? How robust are the parameter uncertainties derived from the fits?

While it is a very powerful and widespread approach, there are a number of drawbacks to modelling stochastic processes in the Fourier domain. The dynamic range of PSDs is generally very large, spanning many orders of magnitude in both axes, and the number of independent estimates of the power at different frequencies is a strong function of the frequency itself, which makes fitting analytic functions to the PSD notoriously challenging. Furthermore, many real-world astrophysical time-series have irregular time-sampling, which precludes using the fast Fourier transform (FFT) to estimate the PSD. The generalized Lomb–Scargle (GLS) periodogram is a popular method to approximate the PSD of irregularly sampled time-series data, but the results are highly sensitive to the window function of the observations (Zechmeister & Kürster 2009). When the sampling is almost regular (e.g. for regular sampling with gaps), it is common practice to interpolate over the missing data, but that carries its own problems.

An alternative approach, which circumvents some of these issues, is to model the data directly in the time domain, using a stochastic process model whose PSD can be written down analytically in terms of the model’s parameters. Widely used examples include (continuous or discrete) auto-regressive moving average models as well as certain types of Gaussian Process (GP) models (see e.g. Brewer & Stello 2009, Pereira et al 2019, Bowman & Dorn-Wallenstein 2022, Hey et al 2024). We refer the reader to Aigrain & Foreman-Mackey 2023 for a recent review of GP regression for astronomical time series. For a long time, the computational

<sup>★</sup> E-mail: [niamh.osullivan@physics.ox.ac.uk](mailto:niamh.osullivan@physics.ox.ac.uk)



**Figure 1.** Composite model of the PSD of the solar background, from A23. The full model (solid purple line) is the sum of quasi-periodic components corresponding to the envelope of the solar oscillations (green dashed line), and to the solar rotation period and its first two harmonics (red dashed line), aliases thereof (black dashed line), plus aperiodic components corresponding to granulation (blue dashed line) and supergranulation (orange dashed line), plus white (photon) noise (cyan dashed line).

cost of this approach was prohibitive for time-series consisting of  $\gtrsim 10^3$  observations, but recent developments have overcome this limitation (Foreman-Mackey et al. 2017; Foreman-Mackey 2018).

In this study, we use simulations to quantify the limitations of both approaches and to ascertain when one might be preferable over the other. This work was initially motivated by recent attempts to quantify the amplitudes and time-scales of different contributions to stellar ‘noise’ in radial velocity (RV) planet searches, particularly the study of Al Moulla et al. (2023, hereafter A23), who used data from the *HARPS* and *HARPS-N* solar telescopes to reconstruct and model the PSD of solar disc-integrated RV variability on time-scales ranging from minutes to months. Their final fit to the solar PSD is reproduced in Fig. 1, illustrating the complexity of the overall model, and the significant overlap between the time-scales of the different components. These results are likely to have a significant impact on both the theoretical modelling of stellar (super)granulation and the observing strategy of future Extreme Precision RV (EPRV) surveys. It is therefore important to understand how robust the fitted parameters are to the PSD-fitting methodology and to see if GP regression outperforms the ‘standard’ approach. We also note that similar types of PSD fitting are commonplace in other fields, including asteroseismology, QPOs, and AGN variability, where GP modelling is also starting to emerge as a viable alternative.

The remainder of this paper is structured as follows. We start by reviewing the common practice of fitting generalized Lorentzians to the PSD of solar variability—the so-called solar background in Section 2. In Section 3, we discuss GP kernels that can be used to fit the solar background in the time domain. In Section 4, we describe how we simulate time-series with known PSDs, and fit them in the frequency and in the time domain. We then evaluate and compare the precision and accuracy of the two approaches, for different intrinsic PSDs and time-sampling properties, in Section 5. We discuss the implications of our findings and outline our future work in Section 6.

## 2. MODELLING THE SOLAR BACKGROUND IN THE FREQUENCY DOMAIN

In this section, we review the practice of characterizing stellar variability on time-scales ranging from hours to months by fitting a sum of modified Lorentzian functions to the PSD of the observations. This approach was first developed by Harvey (1985), in the context of helioseismic RV observations, and has since become widespread, being routinely applied to both solar and stellar RV and photometric observations (see e.g. Goldreich, Murray & Kumar 1994; Aigrain, Favata & Gilmore 2004; A23).

### 2.1 The ‘solar background’

The term ‘solar background’ refers to the broad spectrum of solar variability in RV or photometry, which acts as a nuisance signal in the detection and study of  $p$ -mode oscillations (see e.g. Harvey 1985). It is modelled as the sum of up to five components, in order of decreasing frequency: the broad envelope of  $p$ -mode oscillations, caused by trapped sound waves in the stellar atmosphere (Goldreich et al. 1994; Kjeldsen & Bedding 1995), granulation, caused by convective upflows (Roudier et al. 1991; Rimmele et al. 1995; Lefebvre et al. 2008), supergranulation, the origin of which is still under debate (see Rieutord & Rincon 2010), the rotational modulation of active regions, caused by stellar spots rotating in and out of our field of view (Saar & Donahue 1997; Meunier, Desort & Lagrange 2010; Haywood et al. 2016), and long-term trends caused by the star’s activity cycle (Santos et al. 2010). Beyond helio (and astero)seismology, this background is also an important nuisance signal that needs to be understood and mitigated for exoplanet searches, as it can hide and mimic planet signals in both Doppler and photometric stellar time-series.

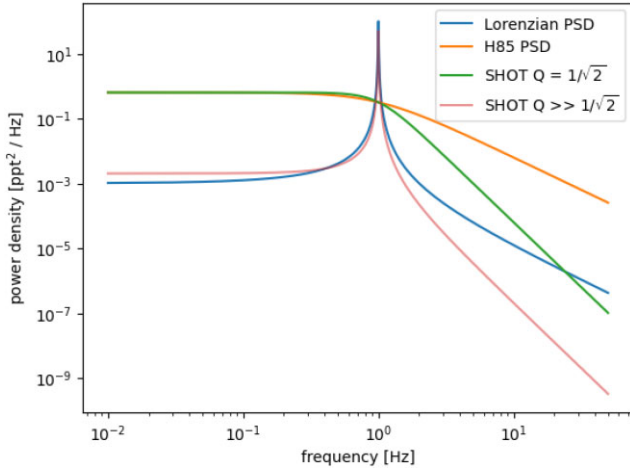
As illustrated in Fig. 1, the components of the solar background have time-scales ranging from minutes (in the case of  $p$ -modes), to hours (granulation), to days (supergranulation), to weeks or months (rotation), and in some cases decades (magnetic activity cycles). Carefully designed observing strategies can be used to minimize the impact of the higher frequency components on RV planet searches, e.g. by selecting exposure times that average out the  $p$ -modes and spacing out consecutive exposures by more than the characteristic time-scale of granulation (Dumusque et al. 2011). However, the supergranulation component is particularly problematic, as it has characteristic time-scales of 1–2 d (Meunier et al. 2015), which limits the frequency with which a given target can be monitored without being affected by correlated noise. Meunier (2021) provides a more detailed review of the effect of solar and stellar variability on RV planet searches.

The characterization of the components of the solar background has been traditionally done by fitting analytic functions to estimates of the PSDs of the observations. In the rest of this section, we give a brief overview of the functional forms used in the literature for modelling the different components and discuss some of the subtleties of the fitting process.

### 2.2 Aperiodic components

Harvey (1985) introduced the notion of modelling granulation and supergranulation signals in solar RVs as a random process with exponentially decaying autocovariance, with variance  $\sigma^2$  and characteristic decay time-scale  $\tau$ :

$$k_H(t, t') = \sigma^2 \exp\left(-\frac{|t - t'|}{\tau}\right). \quad (2)$$



**Figure 2.** Comparison between the different types of PSD functions used in this paper. The orange and green curves, respectively, show the aperiodic Lorentzian PSD of equation (3) and the steeper-sloped version of equation (5), while the blue and pink curves, respectively, show the periodic Lorentzian PSD of equation (9), and the high- $Q$  SHO-term PSD of equation (13).

The PSD of this process is

$$P_H(\nu) = \frac{2\sigma^2\tau}{1 + (2\pi\nu\tau)^2}. \quad (3)$$

This is often known as a Harvey function (hence the subscript ‘H’).

There is a difference of a factor 2 between the expression given by Harvey (1985) for the PSD of such a process and equation (3); we can check that the latter is correct by integrating it over the full frequency domain:

$$V_H = \int_{-\infty}^{+\infty} P_H(\nu) d\nu = \sigma^2, \quad (4)$$

satisfying Parseval’s theorem. This discrepancy means that some of the literature estimates of  $\sigma$  derived from PSD fits of this type may need to be adjusted by a factor  $\sqrt{2}$ .

In any case, later studies (e.g. Harvey et al. 1993; Lefebvre et al. 2008; Michel et al. 2009; Karoff et al. 2013; Kallinger et al. 2014) have shown that the power spectrum of granulation in particular is significantly steeper, and better approximated by a PSD of the form

$$P_H(\nu) = \frac{A_H}{1 + (2\pi\nu\tau)^4}. \quad (5)$$

This is the expression used in more recent work, including A23. The difference between equations (3) and (5) is illustrated in Fig. 2. The variance of this process is

$$V_H = \int_{-\infty}^{+\infty} P_H(\nu) d\nu = \frac{A_H}{2\sqrt{2}\tau}, \quad (6)$$

and thus depends on  $\tau$  as well as  $A_H$ .

Nordlund et al. (1997) provide a physical explanation for the apparently steeper power spectrum of granulation. They argue that the Harvey (1985) model corresponds to a turbulent cascade, which has power on all time-scales, whereas granulation is a convection phenomenon, for which there is a minimum characteristic spatial (and thus temporal) scale, and thus a steep decay in the power spectrum.

Note that some studies, e.g. Harvey et al. (1993), let the index of the power spectrum be a free parameter, usually denoted as  $c$ :

$$P_H(\nu) = \frac{A_H}{1 + (2\pi\nu\tau)^c}, \quad (7)$$

but this tends to result in fitted values in the range  $3.5 \leq c \leq 5.5$ , consistent with  $c = 4$  within the uncertainties.

For lower frequency components such as supergranulation (in RV) or faculae (in photometry, Karoff et al. 2013), even steeper PSD indices ( $c \sim 6$ ) are sometimes reported, but it is not clear how robust those estimates are, as the high-frequency tails of these components overlap strongly with the granulation component.

## 2.3 Periodic components

To model (quasi-)periodic components, e.g. the envelope of the  $p$ -modes, Harvey et al. (1993) and later Lefebvre et al. (2008) used a modified Lorentzian function of the form

$$P_L(\nu) = A_L \left( \frac{\nu}{\nu_0} \right)^b \left[ \frac{\Gamma^2}{(\nu - \nu_0)^2 + \Gamma^2} \right]^c, \quad (8)$$

where  $\nu_0$  is the central frequency and  $A_L$  is the power at  $\nu = \nu_0$ . To reduce the number of free parameters, many recent studies that use this type of periodic model, including A23, fix  $b = 0$  and  $c = 1$ , so that

$$P_L(\nu) = A_L \frac{\Gamma^2}{(\nu - \nu_0)^2 + \Gamma^2}, \quad (9)$$

in which case  $\Gamma$  is the half width at half-maximum. The variance of this process is

$$V_L(\nu) = \int_{-\infty}^{+\infty} P_L(\nu) d\nu = \pi A_L \Gamma. \quad (10)$$

Once again, it depends not only on the amplitude but also on the width (or damping time-scale).

Some other recent studies, e.g. Kallinger et al. (2014), use a Gaussian instead of a Lorentzian function to model the envelope of the  $p$ -modes:

$$P_L(\nu) = A_L \exp \left[ - \left( \frac{\nu - \nu_0}{\sigma} \right)^2 \right], \quad (11)$$

where in this case  $\sigma$  is the standard deviation of the Gaussian (not to be confused with the standard deviation of the process). This is more in line with the standard asteroseismology practice for measuring  $\nu_{\max}$ , the peak frequency of the stellar  $p$ -modes, which is then used in scaling relations to determine the stellar mean density (see e.g. Chaplin & Miglio 2013).

Neither of the above formulations have a particularly clear physical motivation; however, their use is justified empirically rather than theoretically.

## 2.4 PSD estimation and fitting

For regularly sampled observations, PSD estimation is straightforward using the Fast Fourier Transform (FFT). For irregularly sampled time-series, such as ground-based RV observations, the Lomb–Scargle (LS) periodogram (Lomb 1976; Scargle 1982), and later derivatives thereof are the PSD estimation method of choice in the astronomical community. It is worth noting that care is needed when selecting the normalization of the LS periodogram so that it gives results that are directly comparable with the FFT-estimated PSDs. See VanderPlas (2018) for a detailed discussion of the LS periodogram and its interpretation.

The fitting process itself is often challenging, owing to the wide dynamic range covered by the PSD in both frequency and power

density, the difficulty in assigning relative weights to the individual PSD samples, and the overlapping nature of the background components, which leads to significant degeneracies between their parameters. Here, we describe the procedure used by A23, which is fairly representative of previous studies. They estimate the PSD using the GLS periodogram (Zechmeister & Kürster 2009), which is then binned to a regular grid in log frequency space, to give the low-frequency components adequate weight in the fit. As previously mentioned, the slope of the individual background components is fixed to improve convergence and reduce degeneracies between the components. Any white noise in the original time-series translates to a constant lower envelope in the PSD, which is incorporated as a constant term in the fit. The fitting itself is done by minimizing the sum of squared residuals using the Levenberg–Marquart algorithm, which also provides estimates of uncertainties on the fitted parameters using the curvature of the metric around the optimum.

The model of A23 (illustrated in Fig. 1), includes aperiodic terms (equation 5) to represent granulation and supergranulation, and Lorentzian periodic terms (equation 9) to represent the envelope of the  $p$ -modes and signals associated with the rotational modulation of active regions. For the latter, they include terms at the fundamental and first two harmonics of the solar rotation period, as well as 1-d aliases of the fundamental and first harmonics. The need to explicitly include terms to model the aliases is another drawback of modelling the data in the frequency domain.

### 3. MODELING THE SOLAR BACKGROUND IN THE TIME DOMAIN USING GPS

GP regression is a natural alternative to PSD fitting when modelling stochastic or quasi-periodic processes in time-series data. Indeed, stochastic process models provided the original motivation for the choice of function used to fit the solar background PSD as described in Sections 2.2 and 2.3. It is thus natural to ask whether GP regression could be used to model the same signals directly in the time domain, side-stepping some of the implementation challenges alluded to in Section 2.4. In particular, the CELERITE package (Foreman-Mackey et al. 2017, hereafter F17; Foreman-Mackey 2018) provides an efficient implementation of GP regression for certain classes of covariance functions that are particularly appropriate to model stellar signals, and scale well to large data sets.

#### 3.1 The CELERITE simple harmonic oscillator term

A generic CELERITE model is constructed by adding or multiplying together basic building blocks (known as ‘terms’) with covariance:

$$k_{\text{cel}}(t, t') = a \exp(-c|t - t'|), \quad (12)$$

where  $a$  and  $c$  are complex numbers. These individual terms can then be combined additively or multiplicatively to construct more complex models.

In particular, the CELERITE package provides a family of kernels that approximate the behaviour of a harmonic oscillator and are particularly appropriate to model a wide range of stellar signals, from aperiodic to strongly periodic. This is implemented in the form of the ‘simple harmonic oscillator’ (SHO) term, controlled by three parameters: an amplitude  $S_0$ , a characteristic (undamped) angular

frequency  $\omega_0$ , and a quality factor  $Q$ . The PSD of the SHO term<sup>1</sup> is

$$P_{\text{SHO}}(\nu) = \frac{2S_0 \nu_0^4}{(\nu^2 - \nu_0^2)^2 + \nu^2 \nu_0^2 / Q^2}. \quad (13)$$

Around  $\nu \approx \nu_0$ , this PSD approximates a standard Lorentzian or Cauchy distribution, but it falls more steeply (with index 4) at high frequencies.

As of version 2 of the CELERITE package, an alternative parametrization is also available, where the user can specify the period  $\rho$ , damping time-scale  $\tau$ , and standard deviation  $\sigma$  of the process instead of the angular frequency, amplitude, and quality factor:

$$\rho \equiv 2\pi/\omega_0, \quad \tau \equiv 2Q/\omega_0, \quad \text{and} \quad \sigma \equiv \sqrt{S_0 \omega_0 Q}. \quad (14)$$

This alternative parametrization matches the variable names in Sections 2.2 and 2.3, making the relationship between the two classes of models more readily apparent.

#### 3.2 Aperiodic components

Clearly, it is straight-forward to reproduce the PSD of equation (3) using a CELERITE term with real-valued  $a$  and  $b$ . This is also the limiting behaviour of the SHO term when  $Q = 1/2$  and  $\omega_0 \rightarrow 0$ .

To reproduce the steeper slope at high frequencies used by more recent studies, however, a natural choice is the SHO term with  $Q = 1/\sqrt{2}$ . In this case, the PSD of the SHO term becomes

$$P_{1/\sqrt{2}}(\nu) = \frac{2S_0}{1 + (\nu/\nu_0)^4}. \quad (15)$$

This is exactly equivalent to equation (5) if  $\nu_0 \equiv 1/2\pi\tau$  and  $2S_0 \equiv A_H$ .

#### 3.3 Periodic terms

The PSD of a generic CELERITE term with real-valued  $a$  but complex-valued  $c$  is (adapting equation 11 of F17)

$$P_C(\nu) = \frac{a}{c} \left[ \frac{1}{1 + \left(\frac{2\pi\nu - d}{c}\right)^2} + \frac{1}{1 + \left(\frac{2\pi\nu + d}{c}\right)^2} \right]. \quad (16)$$

If we set  $a \equiv 2\pi A_L \Gamma$ ,  $c \equiv 2\pi \Gamma$ , and  $d \equiv 2\pi \nu_0$ , we obtain

$$P_C(\nu) = A_L \left[ \frac{\Gamma^2}{(\nu - \nu_0)^2 + \Gamma^2} + \frac{\Gamma^2}{(\nu + \nu_0)^2 + \Gamma^2} \right], \quad (17)$$

which is the sum of two of the Lorentzian functions in equation (9), with characteristic frequencies  $\pm \nu_0$ . This is a fairly close match to the periodic components used by A23, but not an exact one, because of the negative frequency term.

On the other hand, we know that  $p$ -modes are stochastically driven oscillations. Therefore, it makes sense to model them as such, using an SHO term with large  $Q$ , with PSD given by equation (13). If we set  $S_0 = A_L \Gamma^2/\nu_0^2$  and  $Q = \nu_0/2\Gamma$ , the resulting process will have the same variance and coherence time as the one defined by the Lorentzian PSD of equation (9). As Fig. 2 shows, the PSDs of these two processes look very similar in the vicinity of the peak at the characteristic frequency, but diverge away from it, particularly

<sup>1</sup>Note that equation (15) was obtained by substituting for  $\omega = 2\pi\nu$  in equation (20) of F17, then multiplying it by a factor  $\sqrt{2\pi}$ . This is to account for the fact that F17 defined the Fourier transform to be unitary in angular frequency units rather than in natural frequency units, resulting in a different normalization constant for the forward transform.



at high frequencies, where the SHO term gives rise to a steeper power-law slope, with index 4 rather than 2.

### 3.4 The CELERITE rotation term

The PSD of solar and stellar variability often displays peaks at the first harmonic of the rotation period as well as the fundamental, which need to be modelled explicitly to obtain a good fit. For example, A23 included periodic components at the first two harmonics of the solar rotation period (see Section 2.4) in their final fit. Foreman-Mackey et al. (2021) recommends modelling stellar rotation signals using the sum of two underdamped SHO terms: one at the fundamental period, and one at its first harmonic. This is implemented in version 2 of the CELERITE package as the ‘rotation term’.

The rotation term is controlled by five parameters: the standard deviation  $\sigma$  of the process, the period  $P$ , two parameters controlling the quality factors of the fundamental and first harmonics (these are specified in such a way that the fundamental has quality factor  $Q_1 > 1/2$  and the harmonic has  $Q_2 > Q_1$ ), and one parameter corresponding to the amplitude of the first harmonic term relative to the fundamental.

## 4. TEST METHODOLOGY

In this section, we describe how we simulate time-series data sets with known PSDs including either periodic terms, aperiodic terms, or a mixture of the two, and model them both in the frequency domain using analytic PSD functions, and in the time domain using a GP, to compare the two approaches directly. We restrict ourselves to relatively simple processes with one or two components but also test the impact of irregular time-sampling.

### 4.1 Simulating the time-series

We first simulated a set of regularly spaced time stamps lasting 288 d with 100 points per day, corresponding to a 5-min cadence. This corresponds to a Nyquist frequency of  $452.38 \text{ d}^{-1}$ . We then simulated observations with a known PSD by drawing samples from a CELERITE GP. For aperiodic components, we use an SHO term with  $Q = 1/\sqrt{2}$ , the PSD of which is given by equation (15). For periodic terms, we use an SHO term with large  $Q$ , the PSD of which is given by equation (13). An example of such a simulated time-series for an aperiodic component with  $Q = 1/\sqrt{2}$ ,  $S_0 = 2$  (in arbitrary units), and  $\omega_0 = 5 \text{ rad d}^{-1}$  (such as might be expected for a supergranulation signal) is shown in the left-hand panel of Fig. 3.

Finally, white noise with a standard deviation of  $\sigma_w = 0.5$  was added to the simulated time-series to represent measurement uncertainties.

We later investigate the impact of changing the integration time by binning the simulated time-series, and of the baseline by truncating them.

To produce time-series with irregular time-sampling, such as might be expected for ground-based RV observations, we used a representative set of time stamps from the publicly available data set obtained by the HARPS-*N* solar telescope between 2015 and 2018 (Dumusque et al. 2021). The HARPS-*N* solar telescope observes the Sun at 5-min cadence for several hours every day, with occasional interruptions caused by weather and technical issues. We take a 100-d subset of the HARPS-*N* time stamps and then simulated observations in the same way as described above for the regularly sampled case. As the HARPS-*N* data have a roughly 5-min cadence, we expect the Nyquist frequency to be similar to the regular time-sampling case.

An example of the resulting irregularly sampled time-series is shown in the right-hand panel of Fig. 3.

Note that all the variances, standard deviations, and amplitudes discussed in this section are in arbitrary units, but could represent normalized flux units (such as parts-per-thousand or parts-per-million) or relative RV units (such as  $\text{m s}^{-1}$  or  $\text{cm s}^{-1}$ ).

### 4.2 Estimating PSDs

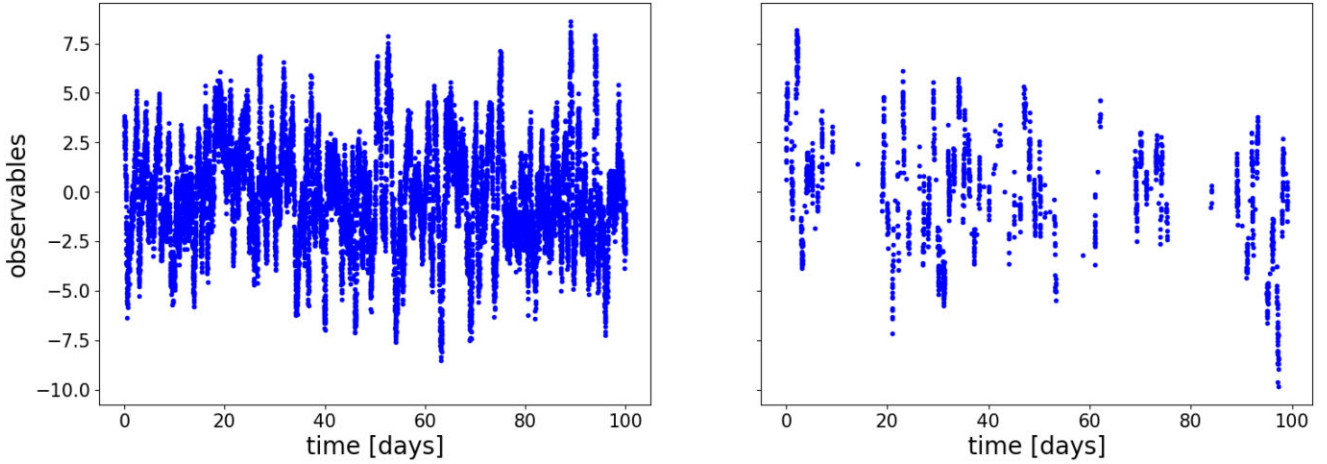
For regularly sampled time-series, we used the NUMPY implementation of the FFT to estimate the discrete Fourier transform (DFT) of the observations. The PSD is then given by the squared modulus of the DFT, divided by the frequency step  $1/T$  where  $T$  is the duration of the simulated time array. The left-hand panel of Fig. 4 shows the FFT-estimated PSD for the simulated time-series shown in the left-hand panel of Fig. 3.

For irregularly sampled time-series, we used the ASTROPY implementation of the GLS periodogram (Zechmeister & Kürster 2009) with the normalization parameter set to ‘PSD’ to estimate the PSD of the observations on a regular grid of frequencies between  $\nu = 0$  and  $\nu = 0.1$  cycles per min (corresponding to the approximate effective Nyquist frequency of the HARPS-*N* solar observations), with frequency step  $1/T$  where  $T$  was the duration of the observations. We found it was necessary to divide the output of the periodogram code by the frequency step and by the number of data points to obtain a properly normalized PSD. As the right-hand panel of Fig. 4 shows, the PSDs obtained in this manner for irregularly sampled time-series can be strikingly different to the true model used to generate the observations, especially at high frequencies. Naturally, these differences will translate to differences between the true and fitted parameters, highlighting the limitations of PSD fitting for irregularly sampled observations.

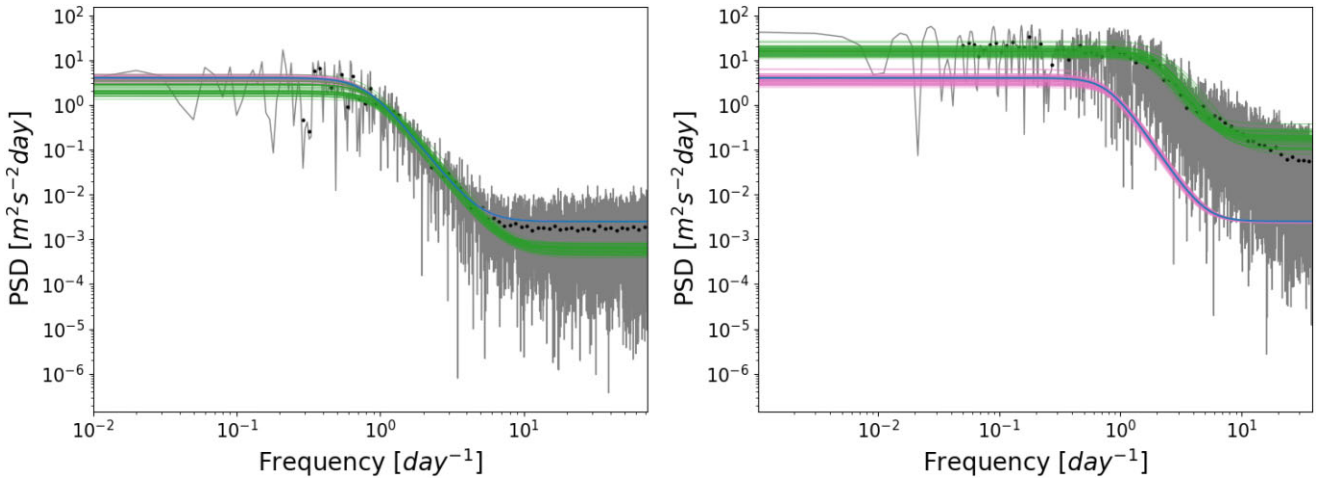
Although the LS or GLS periodograms are widely used to evaluate the PSD of irregularly sampled time-series, they do not explicitly account for the window function of the data, which can significantly alter the PDS estimate. A set of finite duration, irregularly sampled observations of a continuous function  $f(t)$  can be represented as the convolution of  $f(t)$  and a window function  $w(t)$ , which is equal to one during each observation and zero elsewhere. If the duration of the observations is short compared with the time-scale of variations in  $f(t)$ , the window function can be approximated by a sum of delta functions located at the times of the observations,  $\{t_n\}$ . The Fourier transform of the observations is then the convolution of the Fourier transform of  $f(t)$ ,  $F(\nu)$ , and of the Fourier transform of the window function,  $W(\nu)$ . We can gain some insight into the impact of the window function on the estimated PSD by computing the PSD of the window function, which (for instantaneous observations) is given by (see e.g. VanderPlas 2018)

$$P_W(\nu; \{t_n\}) = \left| \sum_{n=1}^N i \sin 2\pi \nu t_n + \cos 2\pi \nu t_n \right|^2. \quad (18)$$

To illustrate this impact, we show in Fig. 5 the PSD of a single aperiodic term estimated from data with both regular and irregular time-sampling, together with the PSD of the window function for both cases. This highlights how the window function dominates the irregularly sampled PSD, which differs substantially from that of the true PSD. Note that the peaks in the window function PSD (corresponding to periodicities in the observation times, e.g. 1 d) do not appear clearly in the PSD of the observations, because the PSD of the true signal, which it is convolved with, has a broad spectrum (if the true signal was strongly harmonic, this would not be the case).



**Figure 3.** Left-hand panel: Example of a simulated, noise-free, regularly sampled supergranulation time-series created using a CELERITE SHO term GP with  $Q = 1/\sqrt{2}$ ,  $S_0 = 2$ , and  $\nu_0 = \frac{5}{2\pi}$  cycles per day. Right-hand panel: Noisy, irregularly sampled time-series generated for the same process with time-stamps from the HARPS-N solar telescope.



**Figure 4.** PSD estimates and fits for a single aperiodic component with  $S_0 = 2$  and  $\nu_0 = 5/2\pi$ , with regular sampling (left-hand panel) and irregular sampling (right-hand panel). The blue line shows the true PSD of the process used to simulate the data, which is the same in both panels. The grey line shows the PSD estimated using the FFT (left-hand panel) and GLS (right-hand panel), respectively, while the black dots show the binned version used in the PSD fit. The green lines show posterior samples from the PSD-fitting MCMC chain, and the pink lines show the same but from the GP fit. Fitted values can be found in Tables 1 and 2.

Ideally one would seek to deconvolve the PSDs of the signal and the window function, but this is not possible in practice because the window function is zero for most values of  $t$  (VanderPlas 2018). As a result, most authors that fit PSD models for irregularly sampled data usually include a constant term to account for the extra white noise induced by the window function but do not explicitly account for its structure beyond that.

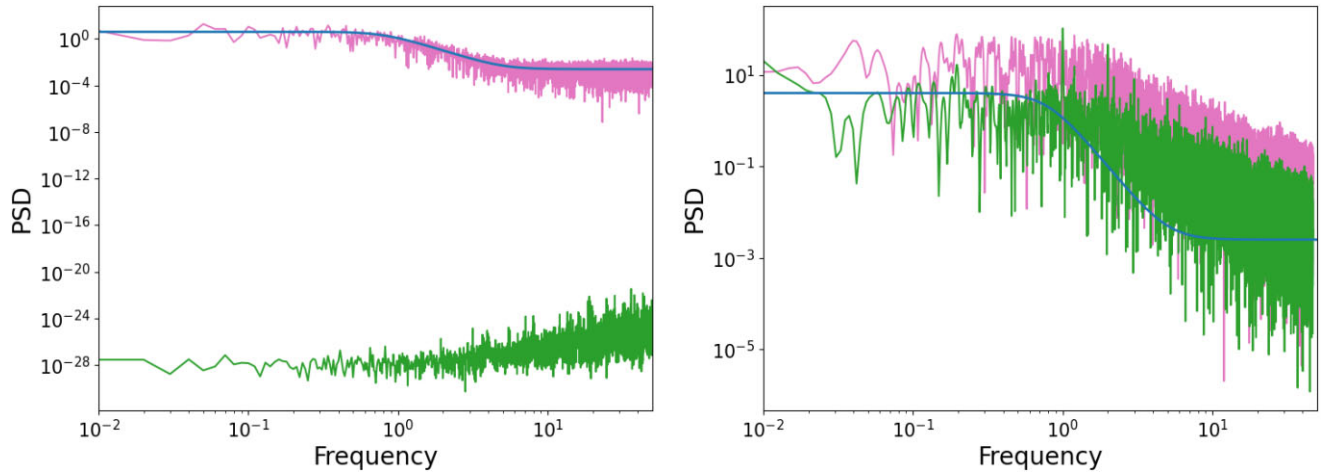
This is one reason why we might expect that modelling stochastic processes in the time domain should be particularly advantageous over doing so in the frequency domain for irregularly sampled observations.

### 4.3 Fitting PSDs

Once the PSD corresponding to a given observation has been estimated, we proceed to fit it, using a PSD model function corresponding to the true PSD of the GP used to generate the data in the first place (namely equation 15 for aperiodic components,

and equation 13 for periodic components). The parameters of the model for each component are  $\ln S_0$  and  $\ln \omega_0$ , plus  $\ln Q$  for the periodic components only. Fitting for the logarithm of the parameters improves convergence and ensures that they always remain positive. Additionally, a constant term,  $\sigma_w$ , (also fit in log-space) is included in the fit to account for white noise and any undersampled signals. The constant added to the PSD model is set to  $\sigma_w/2T$ , where  $T$  is the total duration of the time-series.

Following A23, we bin the PSD into 100 equally spaced bins in log frequency space before fitting, except when working with periodic terms, where we found it necessary to increase the number of bins to 200 in order to properly resolve the PSD peak. We use only bins containing a minimum of three samples (corresponding to frequencies above 0.28 cycles per day), and discard the PSD estimates at lower frequencies. We note that binning PSDs before fitting is normally done in part to ensure that the PSD samples used in the fit are approximately uncorrelated. This would require at least 10, rather than 3, samples per bin. However, this would



**Figure 5.** Impact of the window function on the PSD of a for a single aperiodic component with  $S_0 = 2$  and  $\nu_0 = 5/2\pi$ , with regular sampling (left-hand panel) and irregular sampling (right-hand panel). The PSD estimate of the window function is shown in green, the estimated PSD of the simulated data is shown in pink, while the true PSD is shown in blue.

translate to a minimum frequency of *geq1* cycle per day, which is very close to the typical time-scale for supergranulation in Sun-like stars, and would therefore preclude the extraction of any information about supergranulation time-scales from our simulated time-series. Our decision to use bins with at least  $N_{\min} = 3$  samples represents a compromise between maximizing the frequency range and minimizing correlation between bins. All bins we assigned equal weight in the fits. Again, this follows the procedure described in A23, where no mention is made of assigning weights or uncertainties to the binned PSD samples.

We perform the fit in log-space, i.e. we minimize the sum of the squared differences between the logarithm of the binned PSD and that of the model function, and assign the same weight to all the binned samples. We also carried out some tests where the fits were performed in linear space and the samples were weighted according to the standard deviation in each bin. However, we found that standard deviation of the PSD samples in each bin systematically underestimates the uncertainty on the binned PSD samples, resulting in poor fits and unrealistically small uncertainties on the fit parameters.

We use Markov Chain Monte Carlo (MCMC) to sample the joint posterior probability distribution over the parameters of the fit, using version 3 of the EMCEE package (Foreman-Mackey et al. 2013, 2019). The walkers are initialized in a tight Gaussian ball (with standard deviation 0.01 dex) around a local optimum found using the MINIMIZE function in SCIPY’s OPTIMIZE module. Log-uniform priors are used for all the parameters, within the interval  $[-11; 11]$  in all cases except for the white noise standard deviation  $\log \sigma_w$ , which we restrict to the interval  $[-2; 2]$ . The number of walkers is set to 4 times the number of parameters for each fit, and the MCMC chains are run for 10 000 steps.

We use EMCEE’s built-in functionality to estimate the autocorrelation length of the chains, in order to assess convergence and select an appropriate burn-in and thinning factor. Provided that the longest autocorrelation length estimate across all parameters is  $\tau_{\max} \leq 200$  steps, i.e. less than a 50th of the chain, we consider the autocorrelation time estimates to be reliable and the chains well-converged. We then discard the first  $3\tau_{\max}$  samples as part of the burn-in phase, and then the remainder of the chains by a factor  $\tau_{\max}/4$ . In the few cases where the chains do not converge in 10 000 steps (predominantly

when using the PSD-fitting method for multicomponent models), we arbitrarily set the burn-in time to 1000 steps and the thinning factor to 30, but note that the results should be treated with caution.

For illustrative purposes, we evaluated the PSD model for 50 random samples from the truncated, thinned MCMC chains, which are shown in green on the PSD plots in this paper (such as Fig. 4). We also use the CORNER.PY package (Foreman-Mackey 2016) to display one-dimensional (1D) and two-dimensional (2D) posteriors for each fit and compute the median estimates of each parameter and the associated  $1\sigma$  uncertainties. Figure 6 shows the corner plots obtained in this manner for the time-series shown in Fig. 3, for which the PSDs are shown in Fig. 4.

#### 4.4 GP modelling

We also modelled the simulated time-series directly in the time domain, using a GP model of the same form as the one used to generate them. The fit parameters, priors, and details of the optimization and posterior sampling procedures were identical to those described in Section 4.3, the only difference being that the figure of merit of the fit was the GP likelihood, rather than the  $\chi^2$  of the PSD fit.

The results of the GP fits are shown in pink on the PSD and corner plots (such as Figs 4 and 6), while the results of the PSD fits are shown in green on the same figures, and the true model in blue.

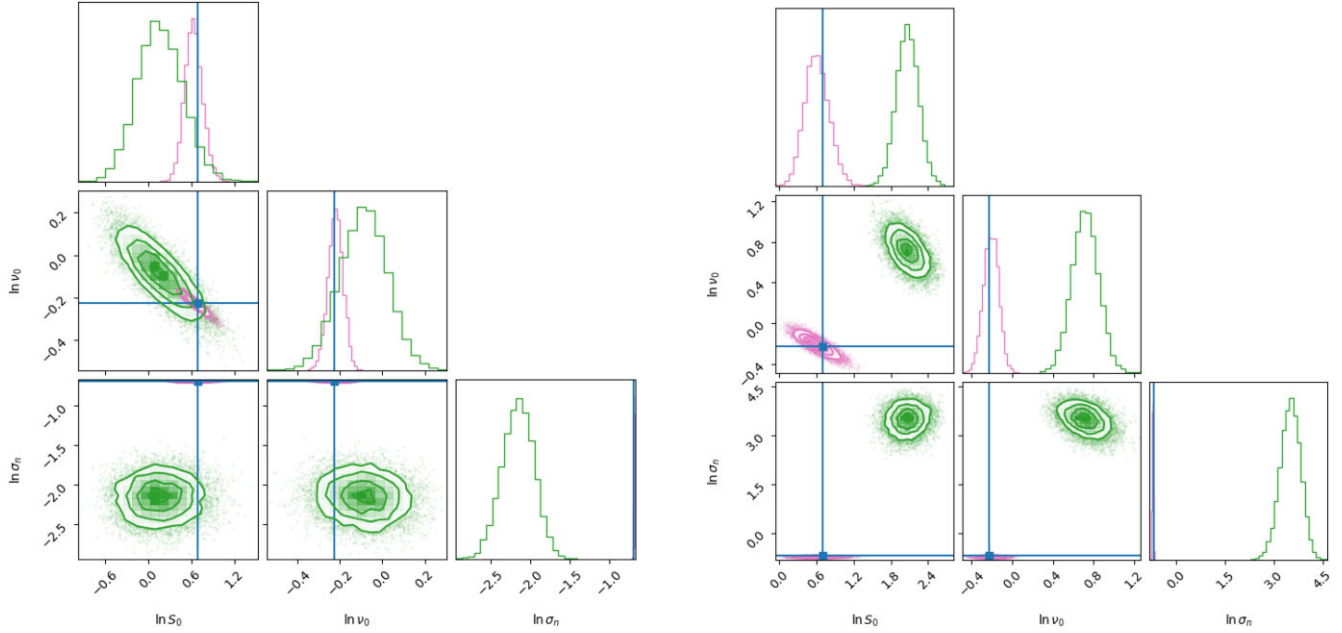
### 5. RESULTS

We are now ready to compare the results obtained for these simulated time-series with both the PSD-fitting and the GP-modelling approach, varying the type of model used and the time-sampling.

#### 5.1 Single-component case

We start with the simplest case, namely the single, aperiodic, (super)granulation-like component illustrated in Figs 3, 4, and 6. The fitted parameters for all the tests described in this section are reported in Tables 1 and 2.

Even in the regularly sampled case, the parameter estimates derived via PSD fitting already deviate from the true values at the  $2\sigma$ – $3\sigma$  level. This is not unexpected for the white noise component: The



**Figure 6.** MCMC posterior distribution plots for the case of a single aperiodic component with  $S_0 = 2$  and  $\nu_0 = 5/2\pi$ , with regular sampling (left-hand panel) and irregular sampling (right-hand panel). These correspond to the time-series shown in Fig. 3 and to the PSDs shown in Fig. 4. The 1D posterior distributions for each parameter, marginalized over all the other parameters, are shown by the histograms in the diagonal panels, in green for the PSD fits and in pink for the GP fits, with the true parameter values indicated by the blue lines. The 2D posteriors are shown in the off-diagonal panels using the same colour-coding, with  $1\sigma$ ,  $2\sigma$ , and  $3\sigma$  contours. Fitted values can be found in Tables 1 and 2.

**Table 1.** Injected and fitted parameters for all cases with regular time-sampling. Parameters more than  $2\sigma$  away from the true value are highlighted in grey.

Case		Parameters						
Aperiodic	Inputted	$\ln S_0$	$\ln \nu_0$	$\ln \sigma_w$				
		0.69	-0.23	-0.69				
	Fitted PSD	$0.16 \pm 0.12$	$-0.22 \pm 0.04$	$-2.15 \pm 0.21$				
	Fitted GP	$0.63 \pm 0.12$	$-0.22 \pm 0.04$	$-0.69 \pm -0.01$				
Periodic	Inputted	$\ln S_0$	$\ln \nu_0$	$\ln Q$	$\ln \sigma_w$			
		-2.30	1.16	2.30	-0.69			
	Fitted PSD	$-2.34 \pm 0.14$	$1.15 \pm 0.02$	$2.49^{+0.75}_{-0.44}$	$-2.65 \pm 0.17$			
	Fitted GP	$-2.27 \pm 0.04$	$1.15 \pm 0.00$	$2.35 \pm 0.12$	$-0.69 \pm 0.01$			
Rotational	Inputted	$\ln \sigma_0$	$\ln P$	$\ln Q_0$	$\ln dQ$	$\ln f$	$\ln \sigma_w$	
		3.91	-0.69	3.00	2.71	-0.22	-0.69	
	Fitted PSD	$4.55^{+1.97}_{-0.72}$	$-0.70 \pm 0.01$	$2.96^{+0.65}_{-0.48}$	$5.06^{+4.30}_{-4.97}$	$-1.48^{+1.35}_{-4.18}$	$-2.22 \pm 0.25$	
	Fitted GP	$3.89 \pm 0.07$	$-0.69 \pm 0.00$	$3.03 \pm 0.13$	$2.22^{+0.58}_{-0.85}$	$-0.10^{+0.08}_{-0.17}$	$-0.72 \pm 0.01$	
Periodic + Aperiodic	Inputted	$\ln S_1$	$\ln \nu_1$	$\ln Q_1$	$\ln S_2$	$\ln \nu_2$	$\ln \sigma_w$	
		-4.61	1.16	2.30	0	-0.23	-0.69	
	Fitted PSD	$-4.47 \pm 0.32$	$1.15 \pm 0.12$	$2.19^{+0.76}_{-0.46}$	$0.13^{+2.73}_{-0.66}$	$-0.31^{+0.27}_{-0.82}$	$-2.69 \pm 0.14$	
	Fitted GP	$-4.59 \pm 0.08$	$1.16 \pm 0.01$	$2.30 \pm 0.12$	$0.01 \pm 0.13$	$-0.26 \pm 0.05$	$-0.69 \pm 0.01$	
Aperiodic + Aperiodic	Inputted	$\ln S_1$	$\ln \nu_1$	$\ln S_2$	$\ln \nu_2$	$\ln \sigma_w$		
		4.61	0.87	9.21	-1.14	-0.69		
	Fitted PSD	$4.73 \pm 0.53$	$0.84 \pm 0.15$	$9.31 \pm 1.14$	$-1.31^{+0.38}_{-0.33}$	$-2.13^{+0.52}_{-0.69}$		
	Fitted GP	$4.69 \pm 0.14$	$0.84 \pm 0.04$	$9.51 \pm 0.23$	$-1.32 \pm 0.09$	$-0.77 \pm 0.01$		

finite duration of each observation and of the total time-series result in additional white noise in the PSD coming from the window function. However, it is concerning to see a significant deviation for the other parameters, even in this idealized case. The discrepancy between injected and recovered parameters is even starker in the irregular case, as expected due to the more complex window function. The PSD fits are also quite sensitive to somewhat arbitrary modelling choices such as the number of PSD bins used.

We then explored the effect of both increasing the exposure times (i.e. binning the time-series) and reducing the duration (truncating

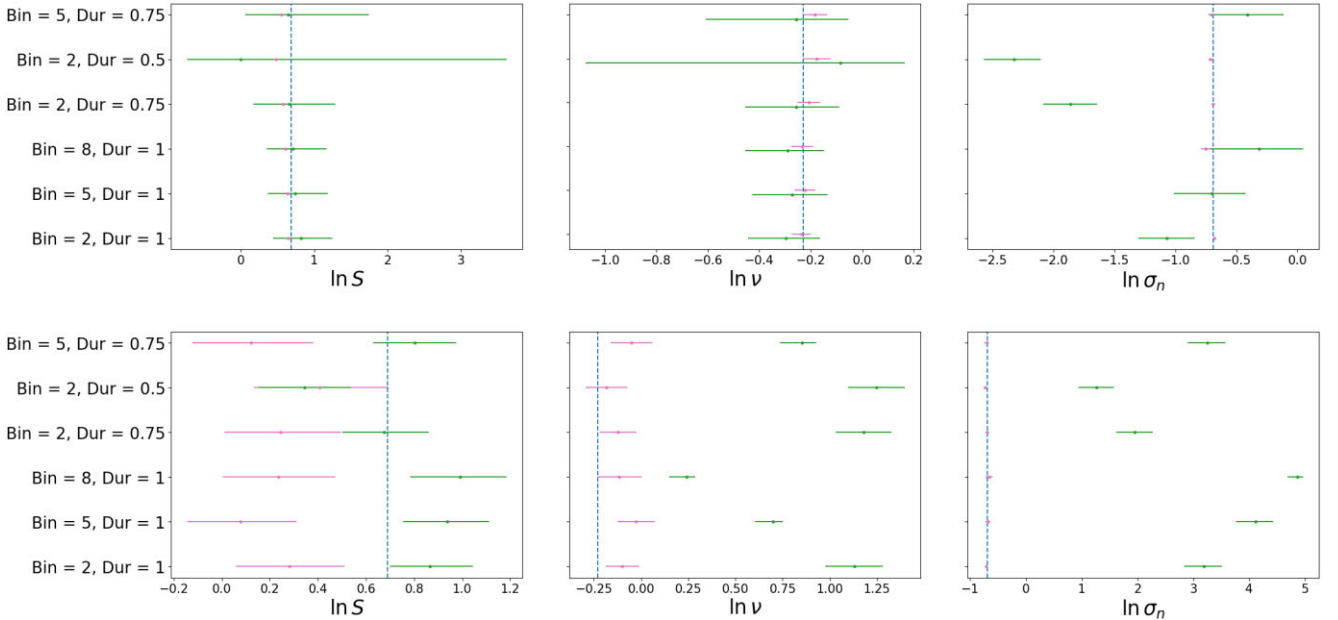
the time-series) to test the sensitivity of the results to these factor. As shown in Fig. 7, this degraded the accuracy of the results further, especially in the irregular case, where all parameters were overestimated, even when the ‘knee’ of the PSD appears well sampled by eye. In the regular case, the PSD fits, while generally consistent with the true value, have larger uncertainties associated with them.

By contrast, the GP-modelling results are both more precise and consistent with the truth in all these cases, though the precision worsens (as expected) when the time-series is binned, truncated, or



**Table 2.** Injected and fitted parameters for cases with irregular time-sampling. Parameters more than  $2\sigma$  away from the true value are highlighted in grey.

Case		Parameters					
<b>Aperiodic</b>	Inputted	$\ln S_0$	$\ln \nu_0$	$\ln \sigma_w$			
		0.69	-0.23	-0.69			
	Fitted PSD	$2.07 \pm 0.18$	$0.72 \pm 0.13$	$3.51 \pm 0.32$			
	Fitted GP	$0.59 \pm 0.20$	$-0.21 \pm 0.08$	$-0.73 \pm 0.02$			
<b>Periodic</b>	Inputted	$\ln S_0$	$\ln \nu_0$	$\ln Q$	$\ln \sigma_w$		
		-2.30	1.16	2.30	-0.69		
	Fitted PSD	$-0.69 \pm 0.13$	$1.39 \pm 0.07$	$0.61 \pm 0.23$	$3.47 \pm 0.25$		
	Fitted GP	$-2.18 \pm 0.11$	$1.16 \pm 0.01$	$2.01 \pm 0.16$	$-0.68 \pm 0.02$		
<b>Rotational</b>	Inputted	$\ln \sigma_0$	$\ln P$	$\ln Q_0$	$\ln dQ$	$\ln f$	$\ln \sigma_w$
		3.91	-0.69	3.00	2.71	-0.22	-0.69
	Fitted PSD	$4.49 \pm 0.11$	$-1.35 \pm 0.10$	$-5.45^{+1.50}_{-2.57}$	$-0.47^{+0.41}_{-0.62}$	$-1.61^{+0.76}_{-0.43}$	$0.18^{+3.34}_{-3.41}$
	Fitted GP	$4.00 \pm 0.07$	$-0.68 \pm 0.01$	$2.63 \pm 0.14$	$-4.79^{+4.26}_{-4.18}$	$-0.66 \pm 0.25$	$-0.78 \pm 0.04$
<b>Periodic + Aperiodic</b>	Inputted	$\ln S_1$	$\ln \nu_1$	$\ln Q_1$	$\ln S_2$	$\ln \nu_2$	$\ln \sigma_w$
		-4.61	1.16	2.30	0	-0.23	-0.69
	Fitted PSD	$-4.67^{+6.92}_{-4.47}$	$1.50 \pm 0.19$	$-5.17^{+6.51}_{-4.28}$	$1.10 \pm 0.14$	$1.22 \pm 0.09$	$2.55 \pm 0.25$
	Fitted GP	$-4.76 \pm 0.24$	$1.17 \pm 0.01$	$2.38 \pm 0.25$	$0.39 \pm 0.26$	$-0.54 \pm 0.16$	$-0.69 \pm 0.02$
<b>Aperiodic + Aperiodic</b>	Inputted	$\ln S_1$	$\ln \nu_1$	$\ln S_2$	$\ln \nu_2$	$\ln \sigma_w$	
		4.61	0.87	9.21	-1.14	-0.69	
	Fitted PSD	$8.39^{+0.28}_{-0.57}$	$1.68^{+0.06}_{-0.10}$	$7.65^{+0.98}_{-11.59}$	$-0.28^{+0.94}_{-8.33}$	$0.69^{+3.43}_{-3.85}$	
	Fitted GP	$4.01 \pm 0.29$	$1.01 \pm 0.08$	$8.76 \pm 0.27$	$-1.11 \pm 0.11$	$-0.80 \pm 0.06$	


**Figure 7.** Recovered versus true parameters for a single aperiodic component with regular time-sampling (top row) and irregular sampling (bottom row), as the binning is increased and the duration of the time-series is truncated. Green symbols show the results of the PSD fits, while pink symbols show those of the GP fits. The true parameter values are shown by the vertical blue lines.

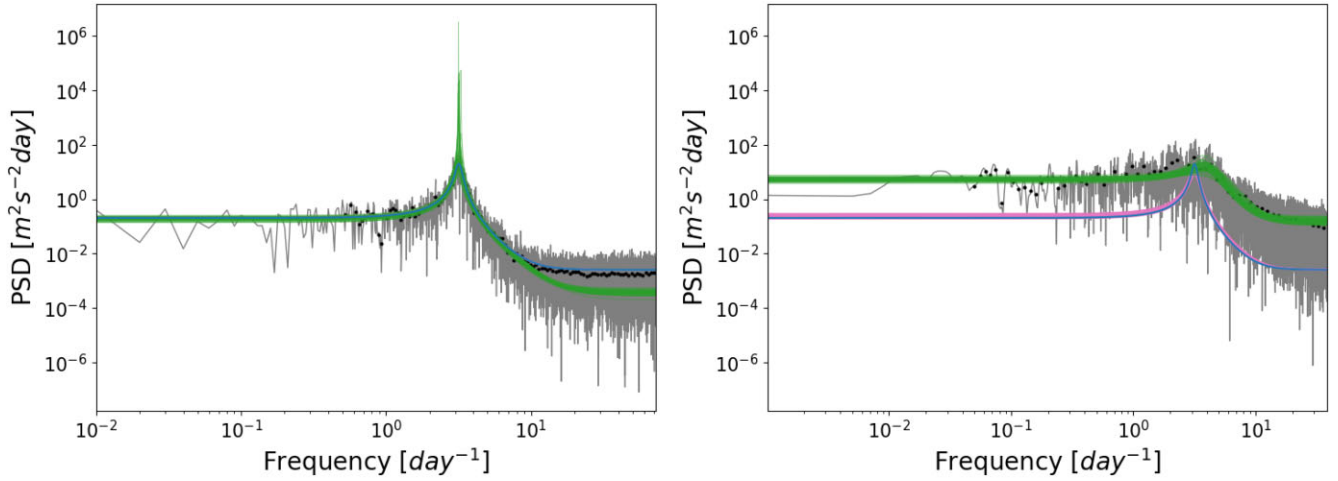
irregularly sampled. We repeated this test multiple times with a range of  $\omega_0$  and  $S_0$  values and found that the results were substantially the same.

We then performed the same set of tests for periodic models, with essentially similar results as shown in Fig. 8. The corner plots associated with this set of tests can be found in Fig. 8.

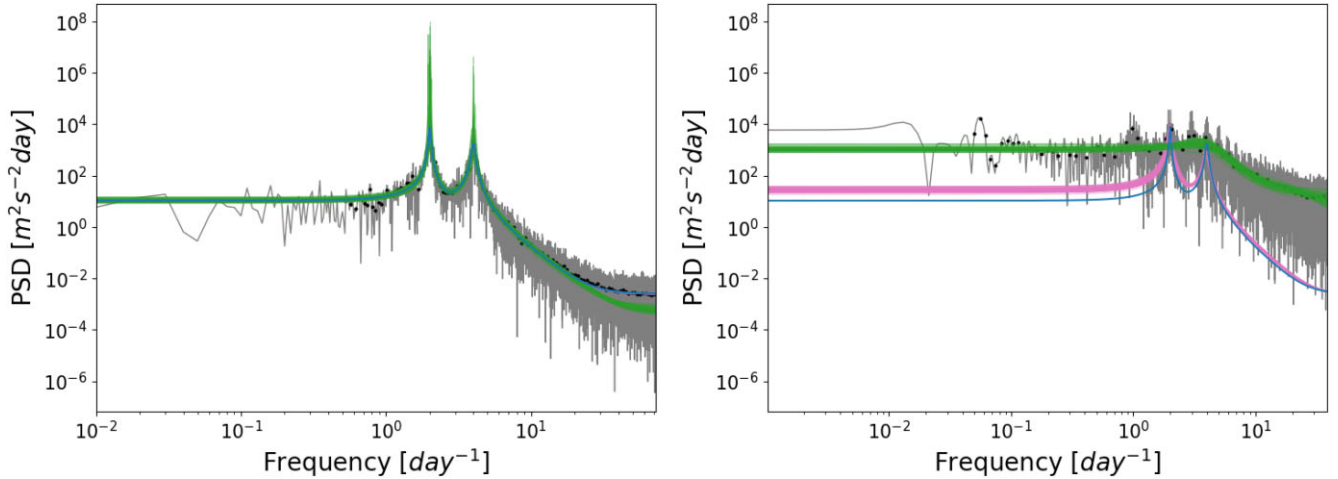
Finally, we considered a case consisting of two periodic terms, with frequencies differing by a factor 2, as implemented in the CELERITE rotation term mentioned in Section 3.4. The results are shown in Fig. 9. While both PSD and GP fitting were successful in the regularly sampled case (left-hand panel), the PSD-fitting ap-

proach entirely fails to recover the true shape of the PSD in the irregular case. By contrast, the GP fit still gives tight and accurate constraints on the underlying PSD and its parameters.

In all cases, the GP-modelling substantially outperformed the PSD-fitting method. Not only do the GP fits result in smaller uncertainties, but they are also substantially more accurate, and generally consistent with the ground truth, unlike the PSD-fitting results. We also note that, as the time-sampling was degraded, MCMC convergence also became increasingly slow in the PSD-fitting case.



**Figure 8.** PSD estimates and fits for a single periodic component with  $Q = 10$ ,  $S_0 = 0.1$ , and  $\nu_0 = 20/2\pi$ , with regular sampling (left-hand panel) and irregular sampling (right-hand panel). The colour coding is the same as Fig. 4. Fitted values can be found in Tables 1 and 2. The associated corner plots can be found in Fig. A1 and A2 respectively.



**Figure 9.** PSD estimates for the CELERITE rotation term model in the regular case (left-hand panel) and irregular case (right-hand panel), using the same colour convention as in Fig. 6. The injected parameters are as follows:  $\sigma = 50$ ,  $P = 0.5$ ,  $Q_0 = 20$ ,  $\delta Q = 15$ , and  $f = 0.8$ . Fitted values can be found in Tables 1 and 2. The associated corner plots can be found in Fig. A3 and A4 respectively.

## 5.2 Multiple components

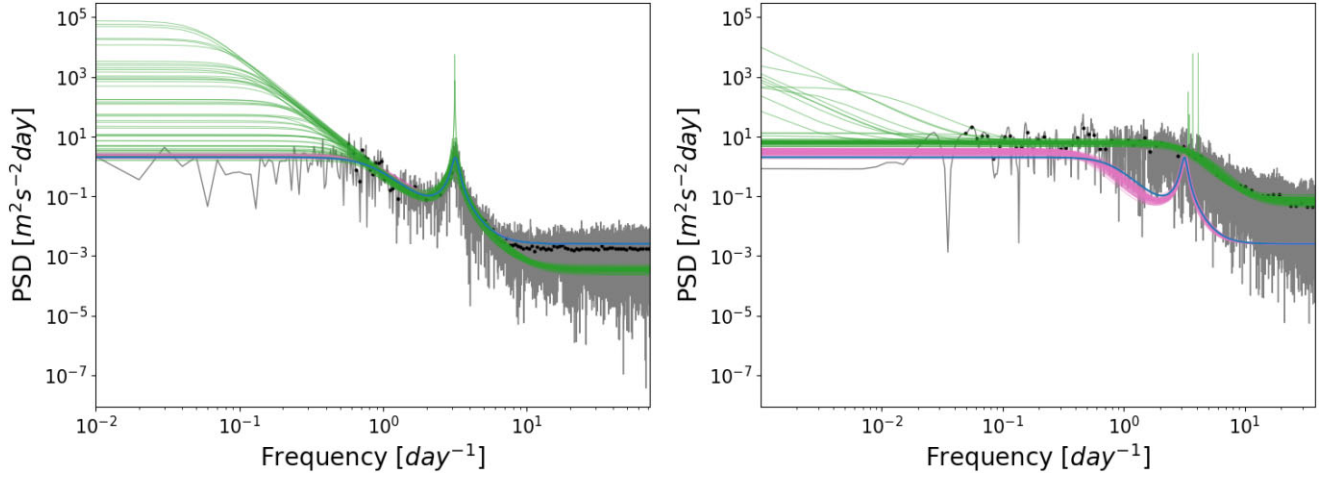
We next considered two-component models, to test our ability to disentangle between the components and recover their parameters correctly. We ran two such tests: one with one aperiodic and one periodic component, and one with two aperiodic components. All cases were run multiple times to ensure the repeatability of results. The PSDs plots for these tests can be found in the body of the text in Fig. 10, while the corner plots can be found in Appendix A. It should be noted that in all cases shown below where the MCMC did not converge is due to highly multimodal posterior distributions.

### 5.2.1 Model comparison

We started by considering a low-frequency, high-amplitude term combined with a higher frequency, lower amplitude, high- $Q$  periodic term. In the context of asteroseismology, the aperiodic term could represent granulation, and the periodic term the envelope of the  $p$ -modes. The results of the PSD fit for these tests are shown in

Fig. 10, and the corresponding corner plots in Appendix A. In both the regular and the irregular time-sampling cases, the GP once more outperformed the PSD-fitting method. It is worth noting that, in the regular time-sampling case, the PSD-fitting method struggles to constrain the quality factor of the oscillations (this is partly due to the binning of the PSD prior to fitting). Meanwhile, in the irregular case, the oscillations peak is entirely absent from the PSD estimate which naturally leads to poor PSD-fitting results (and the MCMC fails to converge), while the GP fits are much less severely affected: The uncertainties are increased but the results are still consistent with the ground truth.

An important question is whether the PSD- and GP-fitting methods can correctly identify the number and nature of the components present in the data. To check whether this is the case we repeated the fits of our two-component time-series and PSDs using different models, including: (i) one periodic component and one aperiodic component (true model), (ii) one aperiodic component only, (iii) two aperiodic components, and (iv) two periodic components. In each case, we compared the Bayesian information criterion (BIC) of the



**Figure 10.** PSD and plots for two-component models with one periodic and one aperiodic component with regular (left-hand panel) and irregular (right-hand panel) sampling, using the same colour convention as in Fig. 4. The injected parameters are as follows:  $S_1 = 0.01$ ,  $\nu_1 = 20/2\pi$ ,  $Q_1 = 10$ ,  $S_2 = 1$ ,  $\nu_2 = 5/2\pi$ , and  $\sigma_w = 0.5$ . Fitted values can be found in Tables 1 and 2. The associated corner plots can be found in Fig. A5 and A6 respectively.

**Table 3.**  $\Delta$ BIC results for multiple component cases in the irregular time-sampling regime. Positive values mean the model is preferred over the true model, negative results mean the opposite. The best-fitting model is highlighted in grey.

Case	$\Delta$ BIC p + ap	$\Delta$ BIC ap	$\Delta$ BIC ap + ap	$\Delta$ BIC rot
PSD p + ap	0	40.01	8.75	
GP p + ap	0	−45.21	−56.50	
PSD ap + ap		1.96	0	
GP ap + ap		−3.14	0	
PSD rot	−48.44	28.92	−67.91	0
GP rot	−184.53	−715.01	−729.66	0

best-fitting model. The results of this model comparison are reported in Table 3. Both the single-aperiodic component model (case ii) and the two-aperiodic component model (case iv) were preferred when using PSD fitting, although the single aperiodic component more so, while the true model (case i) is preferred when using the GP.

We then considered two aperiodic terms with more similar decay times, representative of supergranulation and granulation. The results of these tests are shown in Fig. 11, with the corner plots found in Appendix A. Here, the ‘true’ parameters are the same order of magnitude as the ones reported by A23 for granulation and supergranulation in the solar case. It can be seen that the GP fits result in time-scales consistent with the ground truth for both granulation and supergranulation, but the PSD-fitting approach does not, especially for the granulation-like component. This implies that the time-scales reported by A23 should be treated with caution since they were derived using a PSD-fitting procedure very similar to that used in this work. We also used the BIC to test whether a two-component model was preferred over a single-component case. The results are reported in Table 3. The single-component model is favoured in the PSD case; however, the preference is not statistically significant.

We then proceeded to reduce the frequency separation between the two components, to check when each method stops being able to distinguish between them. The results of these tests are shown in Table 4, where the BIC results for models with either two aperiodic components or one aperiodic component are shown.

When using a GP, the true model is preferred, and thus the two components are correctly identified, though the preference over a single-component model is statistically significant only down to a frequency ratio of  $\frac{\nu_1}{\nu_2} \sim 4$ . By contrast, the PSD fit could only separate the two components for the largest separation. It should be noted that the preference for a two aperiodic model in the irregular GP case with  $\nu_1/\nu_2 = 3$  is not statistically significant, and the result may be different depending on the initial time-series sample.

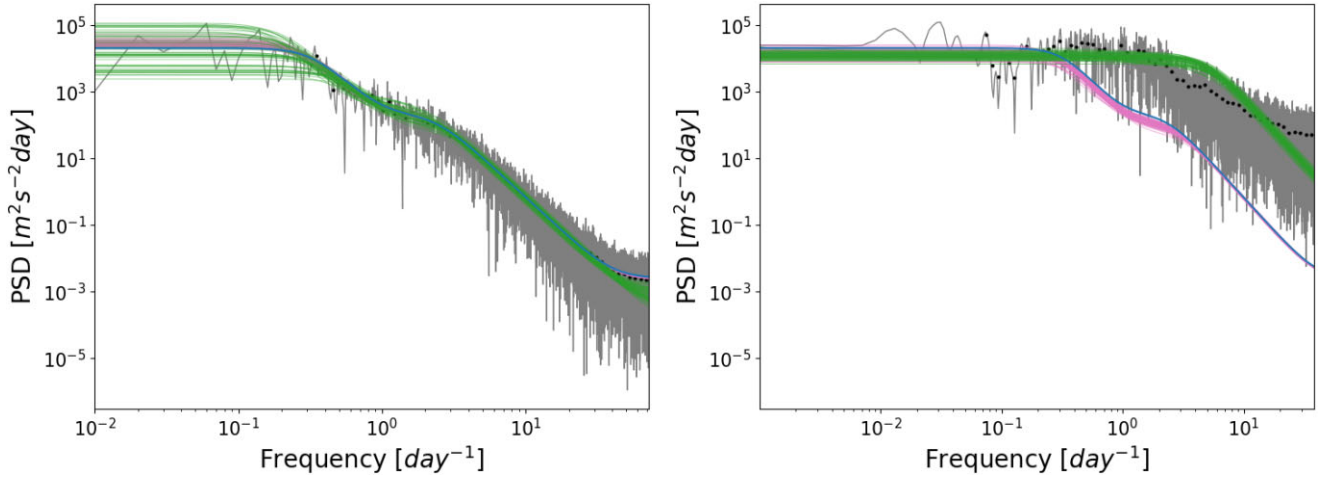
We also used the BIC test for the signal-component CELERITE rotation term to see if two-component models were preferred over the signal-component model in this case. The results are reported in Table 3. A single aperiodic component is preferred in the PSD case.

### 5.2.2 Parameter accuracy

Tables 1 and 2 show the recovered parameters compared with the true values for all five cases studied, for both regular and irregular time-sampling. Cells for which a derived parameter was more than  $2\sigma$  away from the true value are highlighted in grey. In every case discussed above, the GP approach outperformed the PSD-fitting approach. The GP-derived parameters are always close to the true values. However, in a few cases, the uncertainties appear underestimated, so that the  $2\sigma$  credible intervals do not quite overlap with the true value. This tendency to underestimate uncertainties is particularly pronounced for the white noise parameter and the quality factor of any periodic terms, but is also sometimes observed for the characteristic frequencies. The performance of the PSD-fitting method degrades very noticeable when using irregular sampling, whereas the GP fitting remains fairly robust, except in the case with two aperiodic components, where both methods failed to recover the granulation frequency and amplitude within  $2\sigma$ , though the GP values are still of the correct order of magnitude.

## 6. DISCUSSION AND CONCLUSIONS

Having compared the behaviour of PSD fitting and GP regression on simulated data sets with a known ground truth, we now proceed to



**Figure 11.** PSD and plots for two-component models with two aperiodic components with regular (left-hand panel) and irregular (right-hand panel) sampling, using the same colour convention as in Fig. 4. The injected parameters are as follows:  $S_1 = 100$ ,  $\nu_1 = 15/2\pi$ ,  $S_2 = 10,000$ ,  $\nu_2 = 2/2\pi$ , and  $\sigma_w = 0.5$ . Fitted values can be found in Tables 1 and 2. The associated corner plots can be found in Fig. A7 and A8 respectively.

**Table 4.** BIC values for two aperiodic components with regular and irregular time-sampling as the two components are brought closer together. The only initial parameters that are changed are  $\nu_1$  and  $\nu_2$ . Positive values mean a single-aperiodic component model is preferred over the true model, these cases are highlighted in grey, negative results mean the opposite.

$\nu_1 \backslash \nu_2$	Fit	$\Delta\text{BIC ap}$
7.5	PSD reg	−2.39
7.5	GP reg	−234.49
7.5	PSD irreg	−1.54
7.5	GP irreg	−50.79
4.29	PSD reg	8.46
4.29	GP reg	−35.29
4.29	PSD irreg	11.28
4.29	GP irreg	−22.06
3	PSD reg	14.32
3	GP reg	11.94
3	PSD irreg	2.02
3	GP irreg	−2.46
2	PSD reg	16.95
2	GP reg	17.05
2	PSD irreg	8.26
2	GP irreg	13.88

discuss the implications of our results, both for RV planet searches (the original motivation for this work) and more widely, and outline perspectives for future work.

### 6.1 GP regression versus PSD fitting

In this work, we have shown that the analytic functions commonly used to model stochastic variability in the light curves and RV observations of the Sun and Sun-like stars can be reproduced, either exactly or approximately, using GP models. We then showed, q2 using a series of tests on simulated data, that GP regression in the time domain systematically leads to better results than estimating and modelling the PSD in the frequency domain.

While the GP method leads to more accurate and precise results when applied to regularly sampled data, the difference in performance becomes even more pronounced in the irregular sampling case, where the PSD-fitting results deviate dramatically from the

truth, while the GP results are generally consistent with it within  $2\sigma$ . The main reason for this is that GP regression does not rely on PSD estimates, which by definition assume a specific model for the signal whose PSD is being estimated. In the case of the GLS periodogram, this model consists of a single sinusoidal signal, plus white noise. When the time-sampling is close to regular, individual sines and cosines with frequencies between  $1/T$  and  $1/2NT$  are approximately mutually orthogonal, and the periodogram closely approximates the true PSD, even if the signal is more complex than a single sinusoid. However, when the sampling becomes sparser, this orthogonality breaks down, and the periodogram is dominated by the window function. This is particularly problematic for stochastic signals, which have a broad intrinsic PSD, and thus deviate strongly from the simple model implicitly assumed by the GLS periodogram. The GP regression approach avoids this problem by obviating the need for a PSD estimate altogether.

It should be noted that the computational cost of GP regression is significantly larger than that of the PSD-fitting approach. For example, for the regularly sampled, single-aperiodic component case on a 2020 Mac Mini with an Apple M1 chip, the MCMC sampler completed 580 iterations per second for the PSD fitting versus 92.5 iterations per second for the GP-fitting case, a difference of a factor  $>6$ . None the less, as the evaluation of the GP likelihood scales linearly with the number of data points for CELERITE models, and the posterior sampling is easy to parallelize, the computational cost is an obstacle that can be overcome, even for fairly large data sets.

Therefore, we recommend that GP regression in the time domain should be used, wherever possible, in preference to PSD fitting, to model stochastic processes in astrophysical time-series data sets. One limitation of the GP-fitting method as implemented in this work, which emerged during our tests, is that it has a tendency to underestimate the uncertainties on the parameters slightly (by a factor of order 2). At this stage, the reason for this tendency is unknown, but we recommend treating these uncertainties with a degree of caution until this behaviour is better understood.

### 6.2 Implications for RV planet searches

The advantages of GP regression over PSD fitting are particularly relevant to ground-based RV observations, which are always irregularly



sampled. Indeed, GP regression is already widely used to model and mitigate activity signals in RV data sets (see e.g. Haywood et al. 2014; Rajpaul et al. 2015; Barragán et al. 2022). The results presented in this paper, and in particular the finding that GPs can robustly identify multicomponent signals in irregularly sampled data, indicate that GP regression may also have a role to play in the characterization and mitigation of other stellar signals such as supergranulation.

Besides the aforementioned GP methods, numerous other approaches to model activity signals at the spectral or line-profile level have been developed in recent years (see e.g. Collier Cameron et al. 2021; Cretignier, Dumusque & Pepe 2022; Cretignier et al. 2023; Liang, Winn & Melchior 2024). Using a combination of these approaches, it is widely expected that future EPRV surveys aiming to detect Earth analogues, such as the Terra Hunting Experiment (Hall et al. 2018), which will benefit from unprecedented time-sampling and high signal-to-noise ratio, will succeed in mitigating activity signals in Sun-like stars from several m/s down to the sub-m/s level.

The next major challenge is thus supergranulation (Meunier & Lagrange 2020) for which no effective correction methods have been developed yet. Indeed, Lakeland et al. (2024) recently analysed the RV variability of the quiet Sun as observed by *HARPS-N* (after accounting for the effects of active regions using SDO data, following the method developed by Haywood et al. 2016 and Milbourne et al. 2019). During the low-activity part of the Solar cycle, they found that the quiet photosphere contribution to RV variability is as important as, if not larger than, that from active regions.

A natural continuation of this work will thus be to use the GP methodology we have proposed here to revisit recent studies aiming to quantify the time-scales and amplitudes of variability due to (super)granulation on the Sun, including A23 and Lakeland et al. (2024), as well as in other stars (see e.g. Sulis et al. 2023, who compared granulation signals in photometry and RV using simultaneous observations with CHEOPS and *VLT/ESPRESSO*). While A23 and Sulis et al. (2023) used PSD fitting, Lakeland et al. (2024) used structure functions, a non-parametric approach that, like GP regression, can be applied directly to irregularly sampled data, but does not explicitly use a generative model, and it would be valuable to gain a better understanding of the relationship between structure functions and GP models. One potentially interesting avenue for future progress may also be to search for and characterize any non-Gaussianity in stellar RV signals.

### 6.3 Asteroseismology and planetary transits

The PSD models discussed in Section 2 first arose in the context of helioseismology, and are widely used in asteroseismology, whether using RV observations or photometry. While the ‘background’ in this context is merely a nuisance signal to the oscillations, it is conceivable that there may be some advantage to be gained by modelling this background directly in the time domain using GPs rather than by evaluating and fitting a PSD. This may be particularly helpful for later-type main-sequence stars, for which the oscillation frequencies are higher and the amplitudes lower than earlier type or evolved stars (see e.g. Chaplin & Miglio 2013). As we have shown in this work, multicomponent GP models can be used to detect (quasi-)periodic signals in the presence of correlated noise, even when no clear peak is visible in the PSD of the data. Such models, consisting of one or more aperiodic terms to represent (super)granulation and a single high quality-factor term to represent the envelope of the  $p$ -modes, may help push the detection threshold for  $p$ -modes beyond early- $K$  spectral types, which is the current limit.

This type of model has already been used to model granulation signals and oscillations simultaneously in order to characterize planetary transits around red giants (Grunblatt et al. 2017), and may likely prove important for precise measurements of planetary radii from transits (Barros et al. 2020), e.g. in the context of the PLATO space mission (Rauer et al. 2016).

#### 6.3.1 QPOs and accreting systems

Accreting systems on all scales, from compact binaries to AGNs display QPOs (see e.g. Reig 2011; Ingram & Motta 2019). QPOs are time-variable quasi-periodic signals observed at X-ray, ultraviolet, and optical wavelengths, and are usually modelled in the frequency domain as a sum of Lorentzian functions (Belloni, Psaltis & van der Klis 2002; van Straaten et al. 2002). The parameters of these models, including the frequency and amplitude of any quasi-periodic components as well as the frequency at which aperiodic components display breaks in their PSDs, are used to interpret QPOs and understand their origin and diversity. There is every reason to expect that standard QPO modelling procedures are subject to the same limitations as outlined in this work for stellar signals in RV or photometry, and that GP regression could be a more robust alternative, particularly for shorter or more sparsely sampled time-series. Indeed, GP regression has begun to be used more frequently in the context of QPOs and AGN variability (see Aigrain & Foreman-Mackey 2023, and references therein).

### ACKNOWLEDGEMENTS

This publication is part of a project that has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (Grant agreement No. 865624). NKOS thanks the LSSTC Data Science Fellowship Program, which is funded by LSSTC, NSF Cybertraining Grant number 1829740, the Brinson Foundation, and the Moore Foundation; her participation in the program has benefited this work. This work made use of NUMPY (Harris et al. 2020), MATPLOTLIB (Hunter 2007), and PANDAS (Pandas Development Team 2020) libraries. This work made use of Astropy:<sup>2</sup> a community-developed core Python package and an ecosystem of tools and resources for astronomy (Astropy Collaboration 2013, 2018, 2022).

### DATA AVAILABILITY

The codes used in this manuscript are readily available at <https://github.com/NiamhOSullivan>.

### REFERENCES

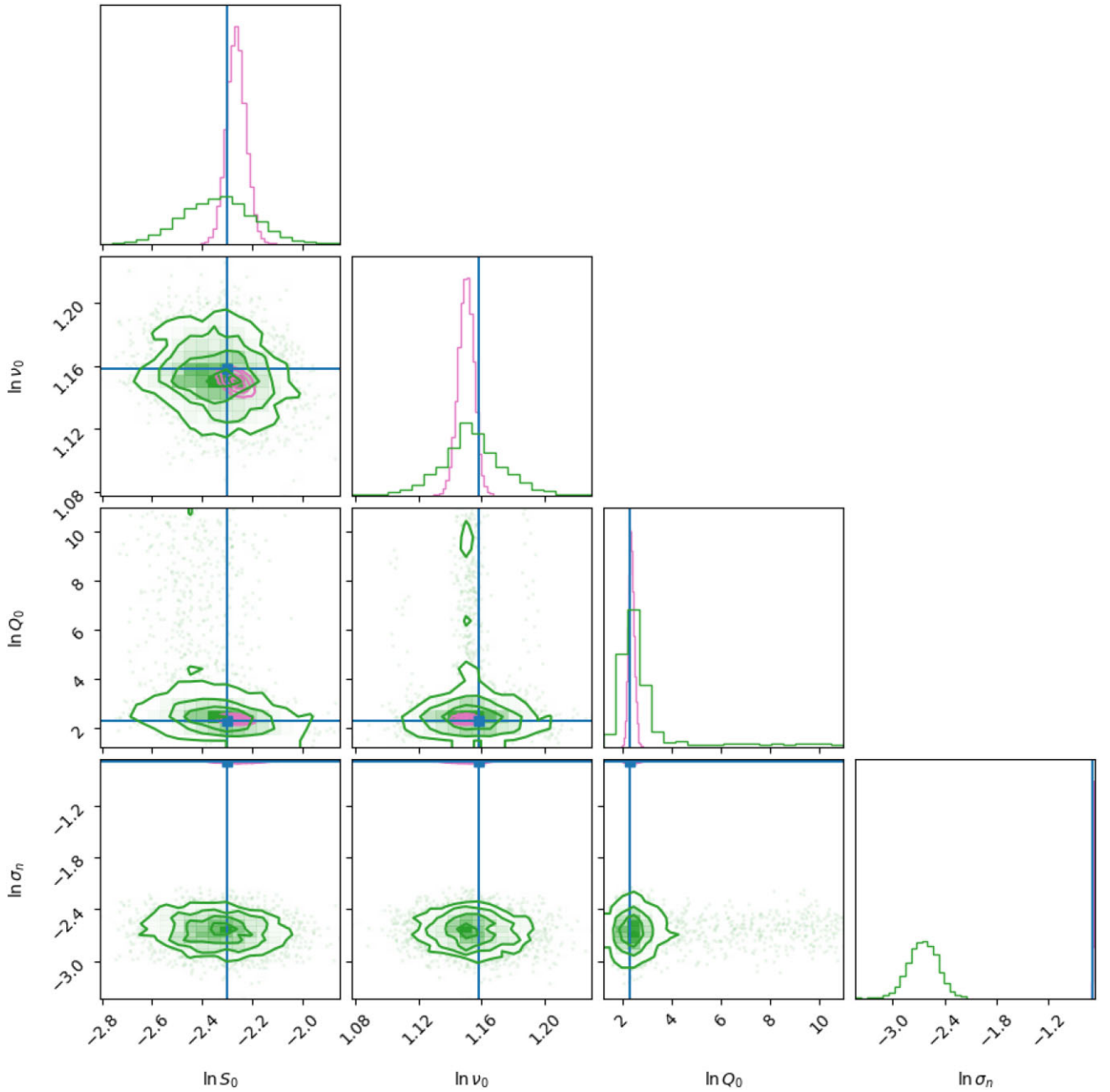
- Aigrain S., Favata F., Gilmore G., 2004, *A&A*, 414, 1139
- Aigrain S., Foreman-Mackey D., 2023, *ARA&A*, 61, 329
- Al Moulla K., Dumusque X., Figueira P., Lo Curto G., Santos N. C., Wildi F., 2023, *A&A*, 669, A39(A23)
- Astropy Collaboration, 2013, *A&A*, 558, A33
- Astropy Collaboration, 2018, *AJ*, 156, 123
- Astropy Collaboration, 2022, *ApJ*, 935, 167
- Barragán O., Aigrain S., Rajpaul V. M., Zicher N., 2022, *MNRAS*, 509, 866
- Barros S. C. C., Demangeon O., Díaz R. F., Cabrera J., Santos N. C., Faria J. P., Pereira F., 2020, *A&A*, 634, A75
- Belloni T., Psaltis D., van der Klis M., 2002, *ApJ*, 572, 392

<sup>2</sup><http://www.astropy.org>

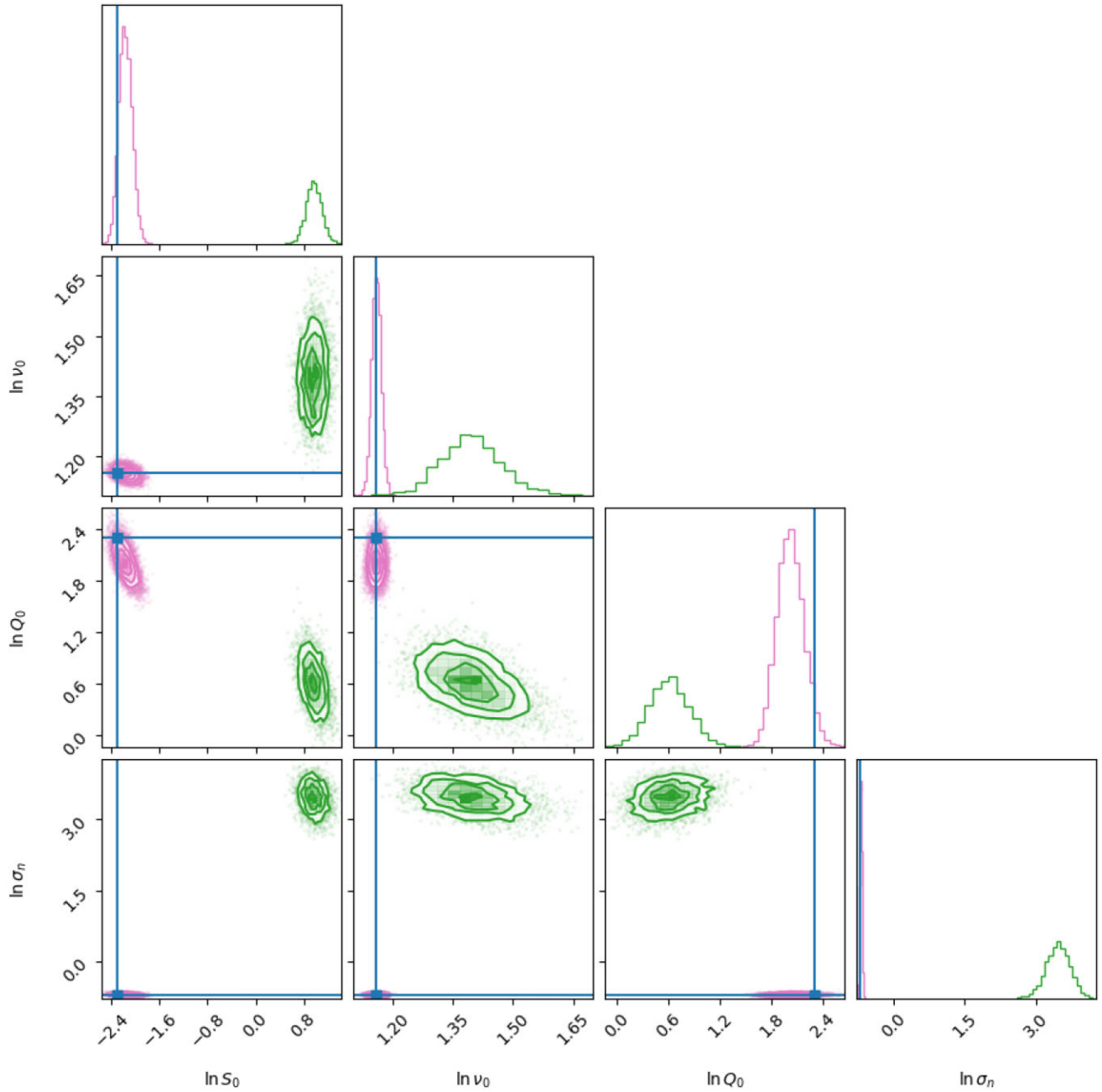
- Bowman D. M., Dorn-Wallenstein T. Z., 2022, *A&A*, 668, A134
- Brewer B. J., Stello D., 2009, *MNRAS*, 359, 2226
- Chaplin W. J., Miglio A., 2013, *ARA&A*, 51, 353
- Collier Cameron A. et al., 2021, *MNRAS*, 505, 1699
- Cretignier M., Dumusque X., Aigrain S., Pepe F., 2023, *A&A*, 678, A2
- Cretignier M., Dumusque X., Pepe F., 2022, *A&A*, 659, A68
- Dumusque X. et al., 2021, *A&A*, 648, A103
- Dumusque X., Udry S., Lovis C., Santos N. C., Monteiro M. J. P. F. G., 2011, *A&A*, 525, A140
- Foreman-Mackey D. et al., 2019, *J. Open Source Softw.*, 4, 1864
- Foreman-Mackey D. et al., 2021, *J. Open Source Softw.*, 6, 3285
- Foreman-Mackey D., 2016, *J. Open Source Softw.*, 1, 24
- Foreman-Mackey D., 2018, *Res. Notes Am. Astron. Soc.*, 2, 31
- Foreman-Mackey D., Agol E., Ambikasaran S., Angus R., 2017, *AJ*, 154, 220
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Goldreich P., Murray N., Kumar P., 1994, *ApJ*, 424, 466
- Grunblatt S. K. et al., 2017, *AJ*, 154, 254
- Hall R. D., Thompson S. J., Handley W., Queloz D., 2018, *MNRAS*, 479, 2968
- Harris C. R. et al., 2020, *Nature*, 585, 357
- Harvey J. W., Duvall T. L. J., Jefferies S. M., Pomerantz M. A., 1993, in Brown T. M., ed., ASP Conf. Ser. Vol. 42, GONG 1992. Seismic Investigation of the Sun and Stars. Astron. Soc. Pac., San Francisco, p. 111
- Harvey J., 1985, in Rolfe E., Battrick B., eds, Future Missions in Solar, Heliospheric and Space Plasma Physics, Vol. 235. ESA Special Publication, Noordwijk, p. 199
- Haywood R. D. et al., 2014, *MNRAS*, 443, 2517
- Haywood R. D. et al., 2016, *MNRAS*, 457, 3637
- Hey D., Huber D., Ong J., Stello D., Foreman-Mackey D., 2024, arXiv e-prints, p. arXiv: 2403.02489
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Ingram A. R., Motta S. E., 2019, *New Astron. Rev.*, 85, 101524
- Kallinger T. et al., 2014, *A&A*, 570, A41
- Karoff C. et al., 2013, *ApJ*, 767, 34
- Kjeldsen H., Bedding T. R., 1995, *A&A*, 293, 87
- Lakeland B. S. et al., 2024, *MNRAS*, 527, 7681
- Lefebvre S., García R. A., Jiménez-Reyes S. J., Turck-Chièze S., Mathur S., 2008, *A&A*, 490, 1143
- Liang Y., Winn J. N., Melchior P., 2024, *AJ*, 167, 23
- Lomb N. R., 1976, *Ap&SS*, 39, 447
- Meunier N., 2021, preprint (arXiv:2104.06072)
- Meunier N., Desort M., Lagrange A. M., 2010, *A&A*, 512, A39
- Meunier N., Lagrange A. M., 2020, *A&A*, 642, A157
- Meunier N., Lagrange A. M., Borgniet S., Rieutord M., 2015, *A&A*, 583, A118
- Michel E., Samadi R., Baudin F., Barban C., Appourchaux T., Auvergne M., 2009, *A&A*, 495, 979
- Milbourne T. W. et al., 2019, *ApJ*, 874, 107
- Nordlund A., Spruit H. C., Ludwig H. G., Trampedach R., 1997, *A&A*, 328, 229
- Pandas Development Team, 2020, *pandas-dev/pandas: Pandas*. Zenodo
- Pereira F. et al., 2019, *MNRAS*, 489, 5764
- Rajpaul V., Aigrain S., Osborne M. A., Reece S., Roberts S., 2015, *MNRAS*, 452, 2269
- Rauer H., Aerts C., Cabrera J., PLATO Team, 2016, *Astron. Nachr.*, 337, 961
- Reig P., 2011, *Ap&SS*, 332, 1
- Rieutord M., Rincon F., 2010, *Living Rev. Sol. Phys.*, 7, 2
- Rimmele T. R., Goode P. R., Harold E., Stebbins R. T., 1995, *ApJ*, 444, L119
- Roudier T., Vigneau J., Espagnet O., Muller R., Mein P., Malherbe J. M., 1991, *A&A*, 248, 245
- Saar S. H., Donahue R. A., 1997, *ApJ*, 485, 319
- Santos N. C., Gomes da Silva J., Lovis C., Melo C., 2010, *A&A*, 511, A54
- Scargle J. D., 1982, *ApJ*, 263, 835
- Sulis S. et al., 2023, *A&A*, 670, A24
- van Straaten S., van der Klis M., di Salvo T., Belloni T., 2002, *ApJ*, 568, 912
- VanderPlas J. T., 2018, *ApJS*, 236, 16
- Zechmeister M., Kürster M., 2009, *A&A*, 496, 577

## APPENDIX A: CORNER PLOTS

Corner plots of the parameters of the PSDs are shown in Figs 8, 9, 10, and 11.

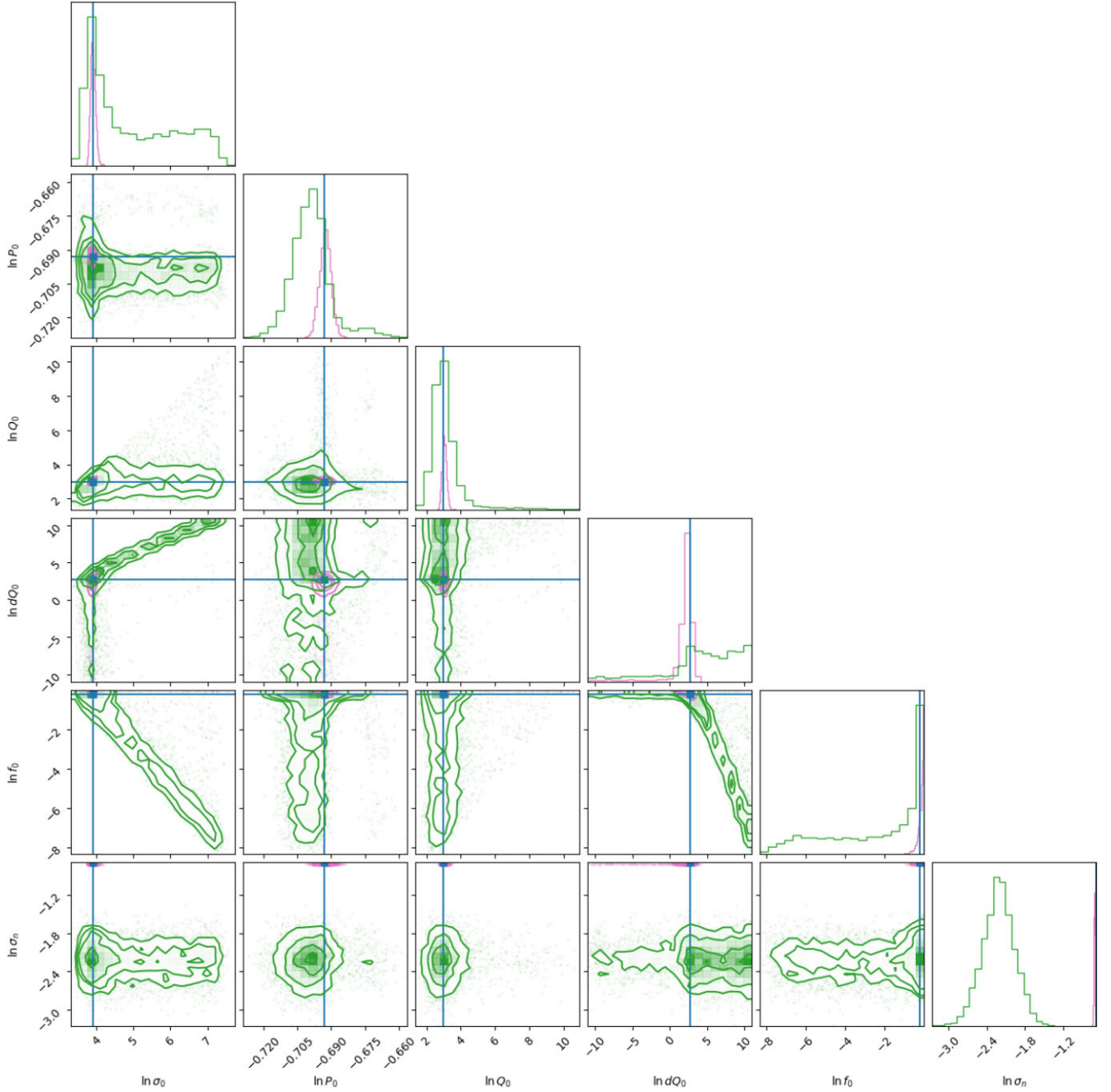


**Figure A1.** Corner plots for the case of a single periodic component with  $Q = 10$ ,  $S_0 = 0.1$ , and  $\nu_0 = 20/2\pi$ , with regular sampling. PSD fits are shown in Fig. 8. The GP MCMC posteriors are in pink, while the FFT/GLS posteriors are in green.

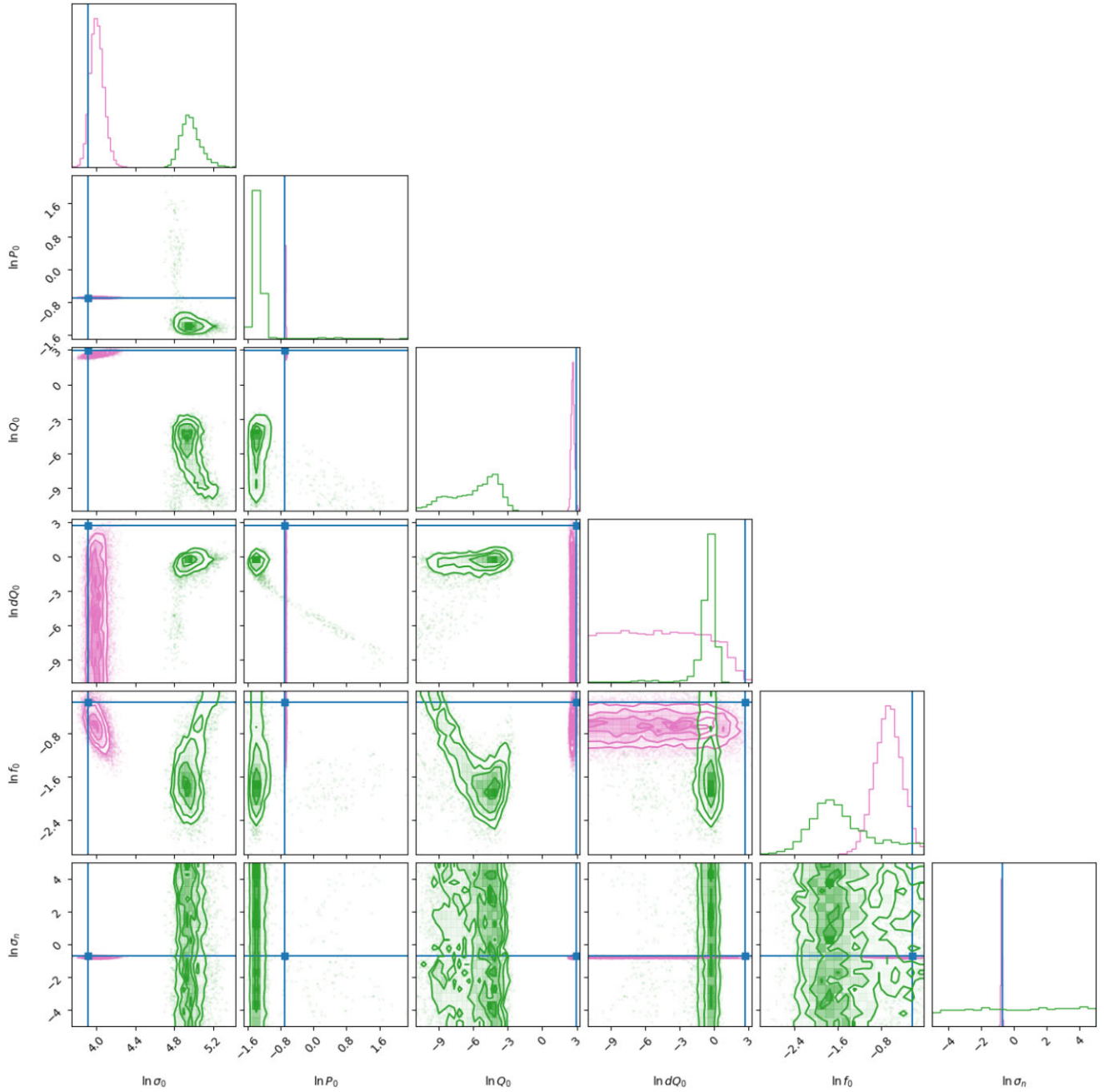


**Figure A2.** Corner plot for the case of a single periodic component with  $Q = 10$ ,  $S_0 = 0.1$ , and  $\nu_0 = 20/2\pi$ , with irregular sampling. PSD fits are shown in Fig. 8. The GP MCMC posteriors are in pink, while the FFT/GLS posteriors are in green.

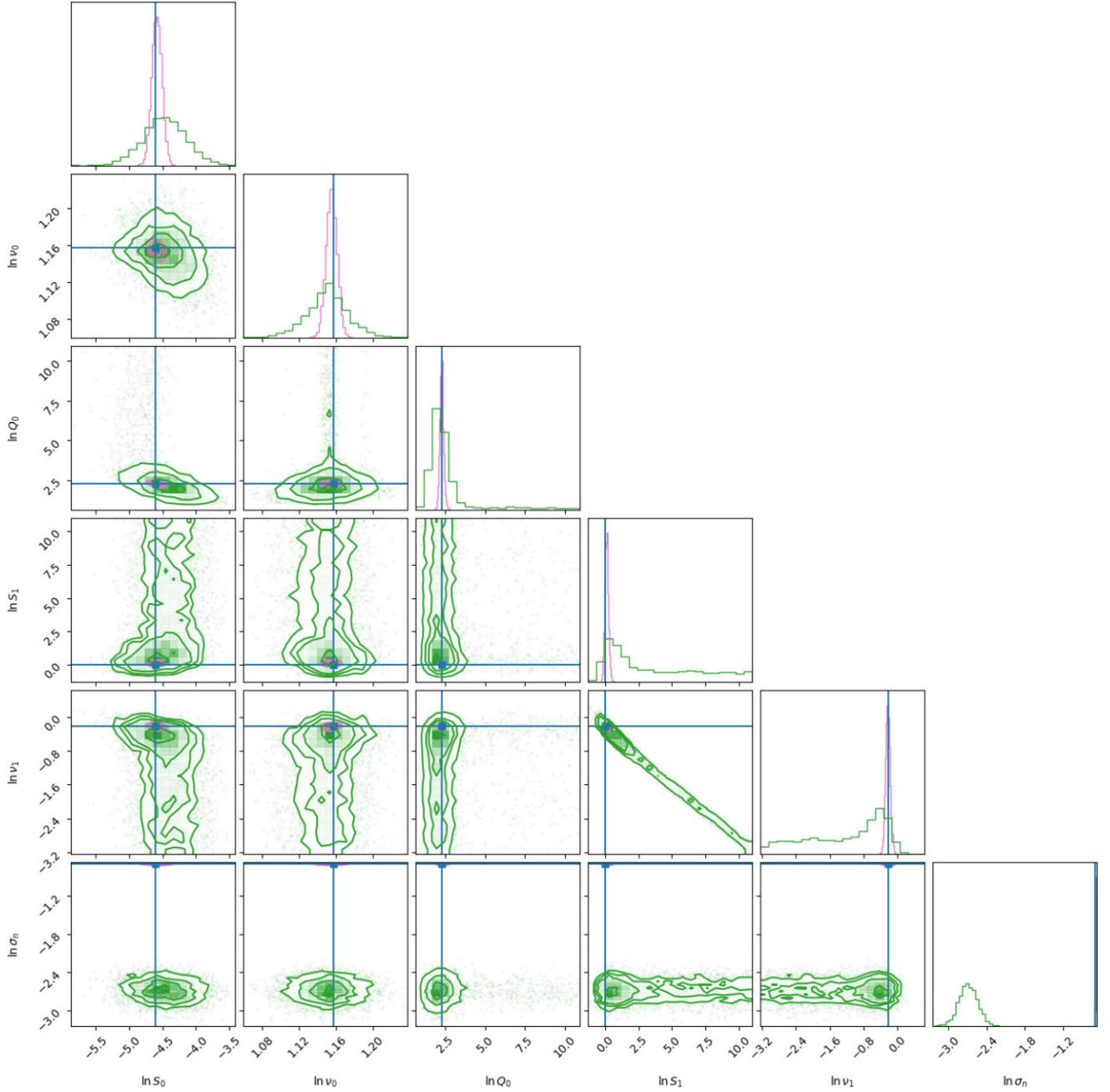




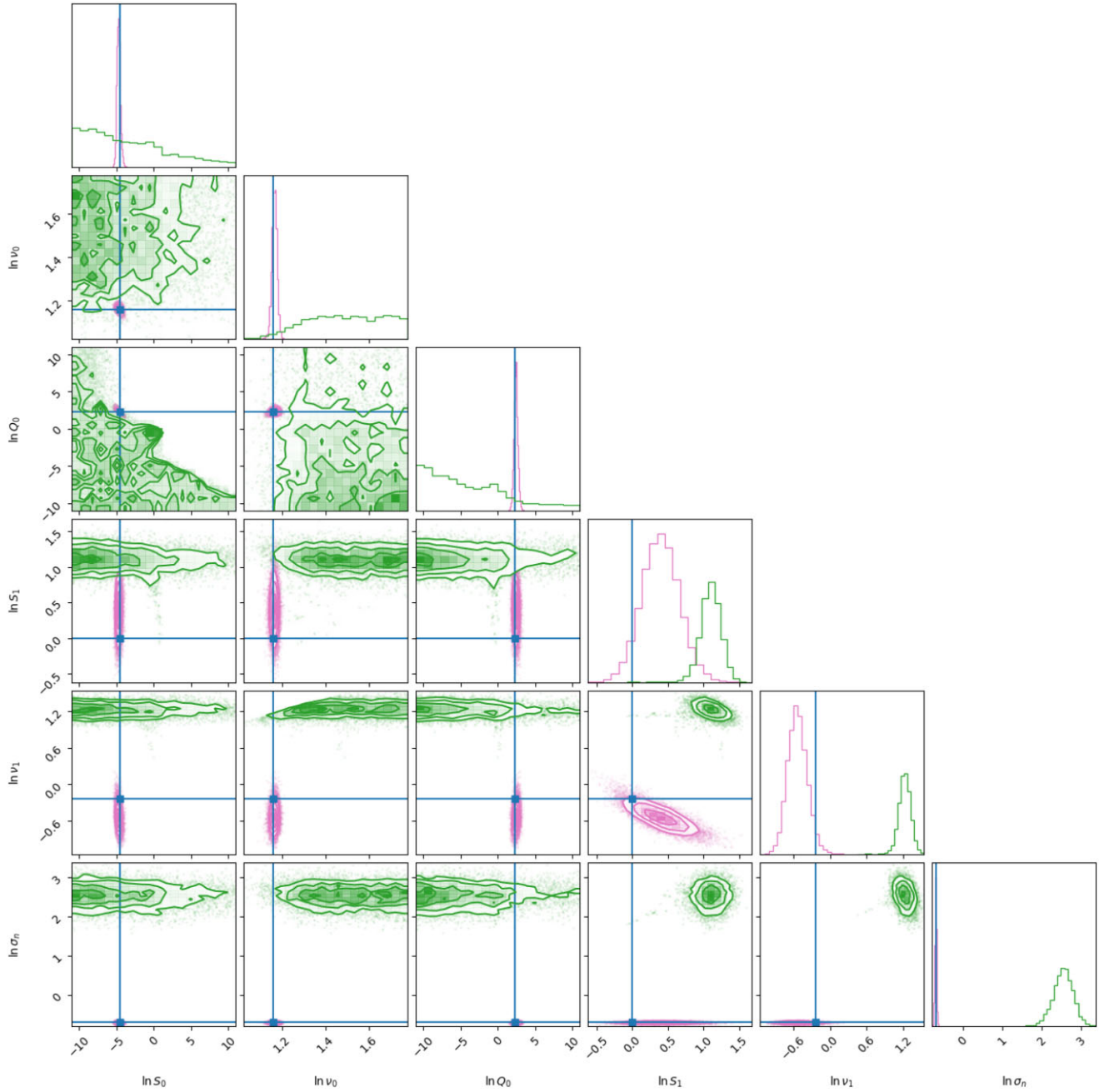
**Figure A3.** Corner plots for the CELERITE rotation term model in the regular time-sampling case. The injected parameters are as follows:  $\sigma = 50$ ,  $P = 0.5$ ,  $Q_0 = 20$ ,  $\delta Q = 15$ , and  $f = 0.8$ . The GP MCMC posteriors are in pink, while the FFT/GLS posteriors are in green.



**Figure A4.** Corner plots for the CELERITE rotation term model in the irregular time-sampling case. The injected parameters are as follows:  $\sigma = 50$ ,  $P = 0.5$ ,  $Q_0 = 20$ ,  $\delta Q = 15$ , and  $f = 0.8$ . The GP MCMC posteriors are in pink, while the FFT/GLS posteriors are in green.

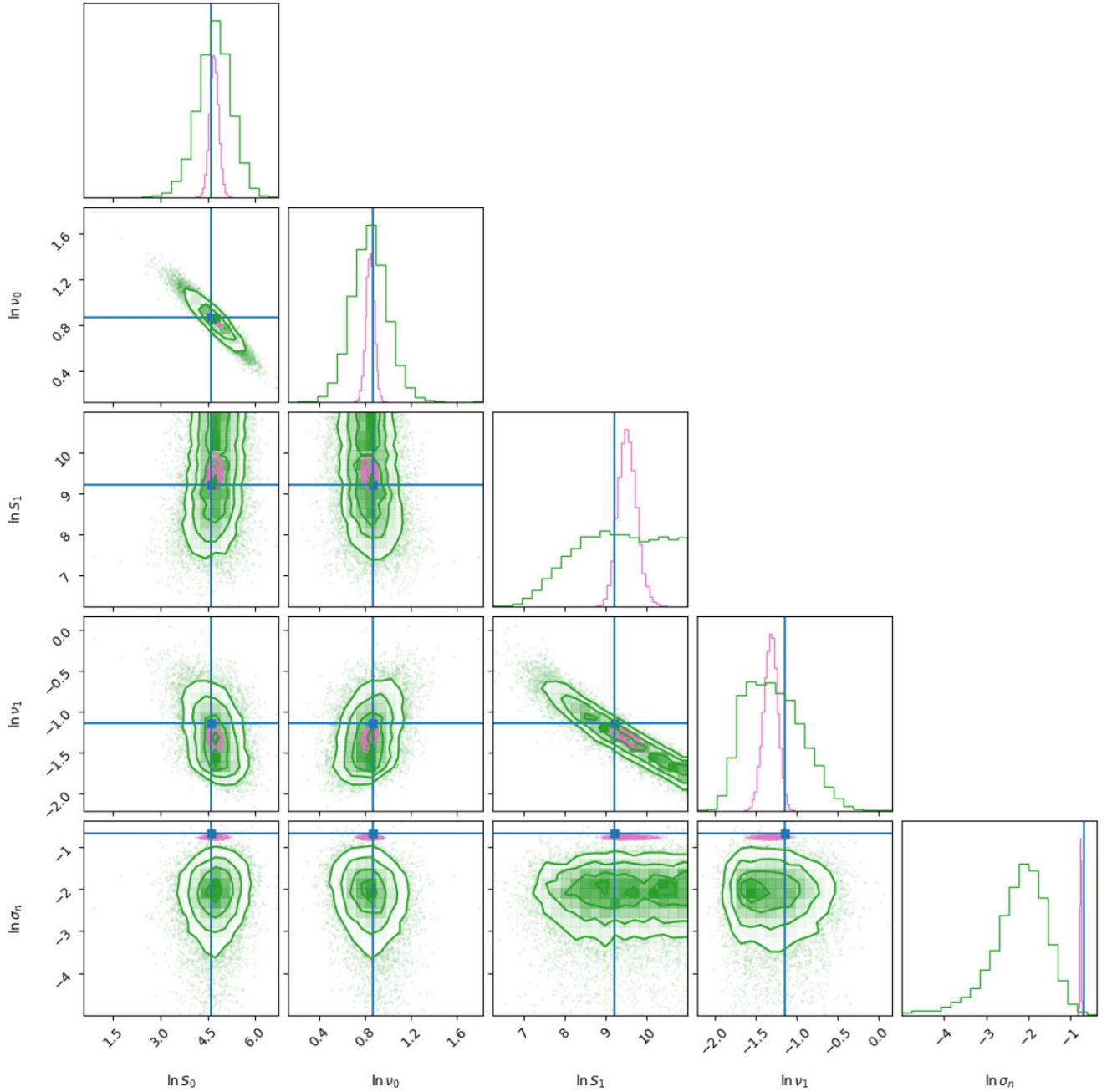


**Figure A5.** Corner plot for two-component models with one periodic and one aperiodic component with regular time sampling. The injected parameters are as follows:  $S_1 = 0.01$ ,  $v_1 = 20/2\pi$ ,  $Q_1 = 10$ ,  $S_2 = 1$ ,  $v_2 = 5/2\pi$ , and  $\sigma_n = 0.5$ . The GP MCMC posteriors are in pink, while the FFT/GLS posteriors are in green.

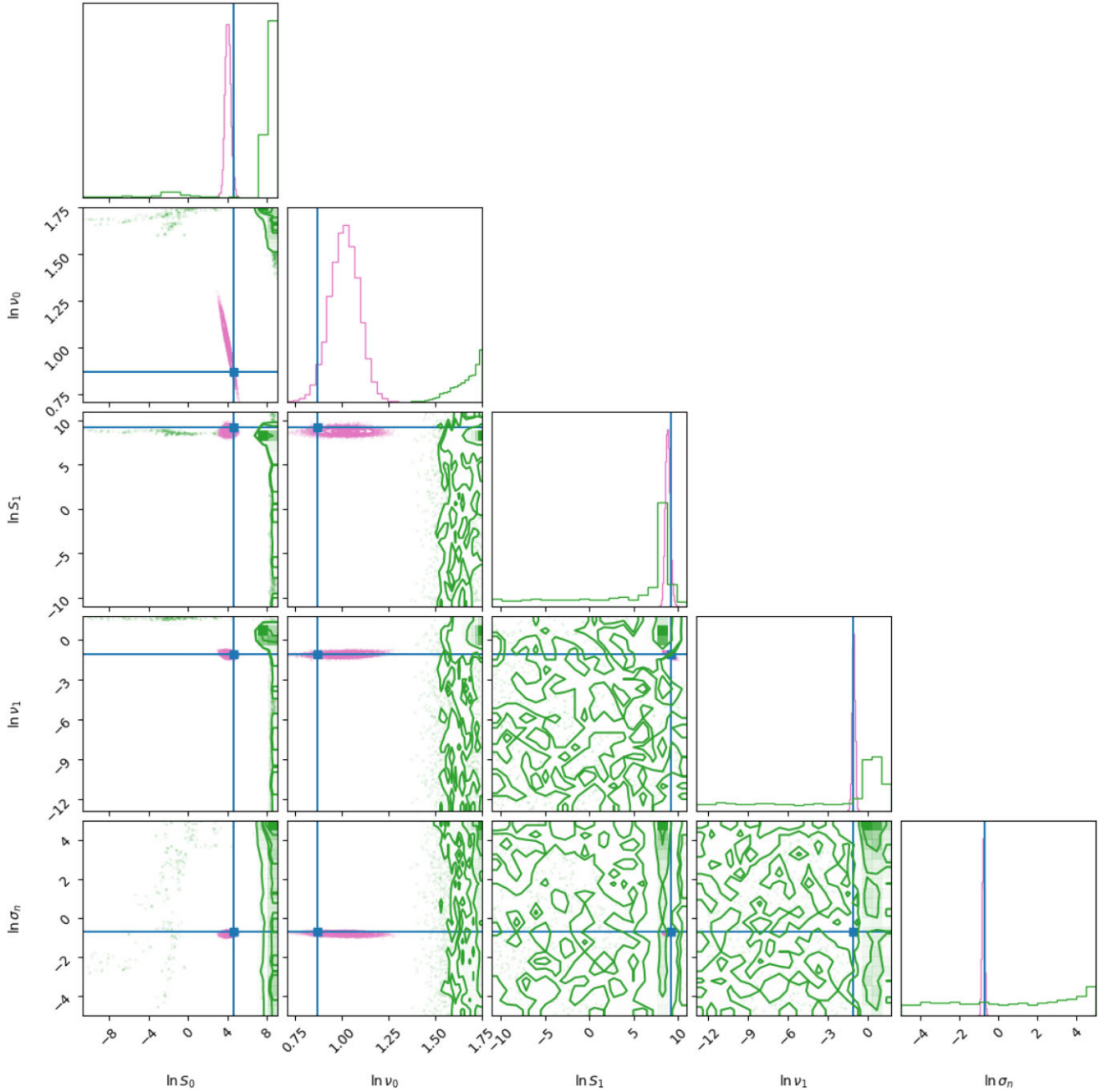


**Figure A6.** Corner plots for two-component models with one periodic and one aperiodic component with irregular time-sampling. The injected parameters are as follows:  $S_1 = 0.01$ ,  $v_1 = 20/2\pi$ ,  $Q_1 = 10$ ,  $S_2 = 1$ ,  $v_2 = 5/2\pi$ , and  $\sigma_n = 0.5$ . The GP MCMC posteriors are in pink, while the FFT/GLS posteriors are in green.





**Figure A7.** Corner plots for two-component models with two aperiodic components with regular time-sampling. The injected parameters are as follows:  $S_1 = 0.01$ ,  $v_1 = 20/2\pi$ ,  $Q_1 = 10$ ,  $S_2 = 1$ ,  $v_2 = 5/2\pi$ , and  $\sigma_n = 0.5$ . The GP MCMC posteriors are in pink, while the FFT/GLS posteriors are in green.



**Figure A8.** Corner plots for two-component models with two aperiodic components with irregular time-sampling. The injected parameters are as follows:  $S_1 = 0.01$ ,  $v_1 = 20/2\pi$ ,  $Q_1 = 10$ ,  $S_2 = 1$ ,  $v_2 = 5/2\pi$ , and  $\sigma_n = 0.5$ . The GP MCMC posteriors are in pink, while the FFT/GLS posteriors are in green.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.