Coefficients of variation for detecting solar-like oscillations

Keaton J. Bell,[1,2]⋆ Saskia Hekker[1,2] and James S. Kuszlewicz[1,2]

[1]*Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany*
[2]*Department of Physics and Astronomy, Stellar Astrophysics Centre, Aarhus University, Ny Munkegade 120, 8000 Aarhus C, Denmark*

ABSTRACT
Detecting the presence and characteristic scale of a signal is a common problem in data analysis. We develop a fast statistical test of the null hypothesis that a Fourier-like power spectrum is consistent with noise. The null hypothesis is rejected where the local "coefficient of variation" (CV)—the ratio of the standard deviation to the mean—in a power spectrum deviates significantly from expectations for pure noise (CV $\approx$ 1.0 for a χ^2 2-degrees-of-freedom distribution). This technique is of particular utility for detecting signals in power spectra with frequency-dependent noise backgrounds, as it is only sensitive to features that are sharp relative to the inspected frequency bin width. We develop a CV-based algorithm to quickly detect the presence of solar-like oscillations in photometric power spectra that are dominated by stellar granulation. This approach circumvents the need for background fitting to measure the frequency of maximum solar-like oscillation power, $\nu_{\max}$. In this paper, we derive the basic method and demonstrate its ability to detect the pulsational power excesses from the well-studied APOKASC-2 sample of oscillating red giants observed by *Kepler*. We recover the cataloged $\nu_{\max}$ values with an average precision of 2.7% for 99.5% of the stars with 4 years of *Kepler* photometry. Our method produces false positives for < 1% of dwarf stars with $\nu_{\max}$ well above the long-cadence Nyquist frequency. The algorithm also flags spectra that exhibit astrophysically interesting signals in addition to single, solar-like oscillation power excesses, which we catalog as part of our characterization of the *Kepler* light curves of APOKASC-2 targets.

Key words: methods: data analysis – methods: statistical – stars: oscillations

1 INTRODUCTION

Our ability to constrain the properties of stars has been recently revolutionized by space-based missions dedicated to obtaining extended, precise time series photometry such as CoRoT (Baglin et al. 2006) and *Kepler* (Borucki et al. 2010). The conditions of stellar interiors have been especially illuminated by detections of solar-like oscillations in thousands of red giant stars with convective outer envelopes (e.g., Hekker et al. 2009; Bedding et al. 2011; Mosser et al. 2011b; Mathur et al. 2011a; Stello et al. 2013; Yu et al. 2018, and references therein). Turbulent motions in these layers stochastically excite globally damped stellar eigenmodes with eigenfrequencies near the convective turnover timescale (Goldreich & Keeley 1977). These pulsations propagate through the stellar interior to manifest as incoherent photometric variations, which can be revealed through Fourier analysis of light curves. Detecting the presence of oscillations in a light curve gives the basic indication that a target is suited for asteroseismology (e.g., Chaplin et al. 2014a;

Miglio et al. 2016). The characteristic frequency of maximum power of these oscillations, $\nu_{\max}$, is a precise tracer of the surface gravity in red giants (e.g., Brown et al. 1991; Kjeldsen & Bedding 1995; Hekker et al. 2013). Asteroseismic determinations of this fundamental stellar parameter propagate to tight constraints on stellar and exoplanetary masses and radii (e.g., Ballard et al. 2014; Campante et al. 2015). Hekker & Christensen-Dalsgaard (2017) give a recent review of the field of giant star seismology and the science that is enabled by detecting solar-like oscillations.

The same turbulent convection that drives solar-like oscillations in red giants produces additional signatures from granulation that dominate the power spectra (Harvey 1985). These features can be described by multiple "super-Lorentzians" that decay to high frequency as ν^{-4} (Aigrain et al. 2004; Kallinger et al. 2014). The oscillation signals are comparatively sharp—the widths of their Lorentzian profiles ($\sim 0.1\,\mu$Hz; Baudin et al. 2011) are inversely proportional to the mode lifetimes (e.g., Christensen-Dalsgaard et al. 1989). To peakbagging efforts that aim to precisely measure pulsation mode characteristics, the granulation effectively contributes frequency-dependent noise to the power

⋆ E-mail: bell@mps.mpg.de

spectrum backgrounds. Obtaining and removing a global fit to these backgrounds is often the most computationally expensive step in asteroseismic analyses. However, this process could be greatly accelerated by independent determinations of ν_{max}, as these are tightly correlated with the background component parameters through empirical scaling relations (Kallinger et al. 2014).

Despite the challenges introduced by the background, spotting the power excesses of solar-like oscillations is not considered a difficult problem for the trained eye. We develop a new method for detecting solar-like oscillations that mimics the approach of an expert, who would prefer to visually inspect the logarithm of the power spectrum. In this representation, the power is distributed about a background that varies over orders of magnitude as a band with near-constant width. Near ν_{max}, the oscillations broaden the logarithmic power spectrum locally. We search for local excesses in power spectrum scatter by evaluating the "coefficient of variation" (CV; standard deviation divided by the mean) in limited frequency ranges. We reject the null hypothesis that a bin contains only noise where this CV value is sufficiently large. This method is successful in detecting sharp spectral features, while being relatively insensitive to slowly varying (in frequency) noise backgrounds. For solar-like oscillators, this enables us to quickly detect and measure the ν_{max} of pulsational power excesses without having to fit the granulation background.

Determining which targets of large time domain surveys exhibit oscillations is the crucial first step toward widespread asteroseismic characterization of stellar populations. While they have already yielded a multitude of important astero-seismic results, the CoRoT, *Kepler* and K2 (Howell et al. 2014) data sets remain treasure troves of unidentified solar-like oscillators. Over the next two years, the *Transiting Exo-planet Survey Satellite* (*TESS*; Ricker et al. 2015) promises to measure photometry for upwards of 20 million stars. By quickly evaluating which targets do or do not exhibit solar-like oscillations, the CV method stands to improve out efficiency in processing these data.

We develop the statistics for the CV spectrum in general terms in Section 2, including a significance criterion. In Section 3, we use the empirical seismic scaling relations to tune a CV-based algorithm for quickly detecting and measuring ν_{max} of solar-like oscillation power excesses. We demonstrate that our method yields the expected results for most dwarf stars (null detections; Section 4) and giants (power excess detections at accurate ν_{max} values; Section 5). Our method identifies stars that exhibit other interesting features in addition to solar-like oscillation power excesses, and these objects are flagged in our characterization of the APOKASC giants (Pinsonneault et al. 2014, 2018) in Appendix A. In Appendix B, we confirm from visual inspection 30 light curves that appear consistent with blends of multiple solar-like oscillators.

2   THE COEFFICIENT OF VARIATION (CV) METRIC

Fourier power spectra are commonly employed for detecting frequencies of intrinsic variability in time series data. Power from deterministic or stochastic signals is distributed dif-

ferently than power that represents background noise. We aim to exploit these differences to quickly distinguish the signals without having to precisely model the background. We develop an approach based on the approximate statistical properties of the background that we find to be a useful tool for identifying signals in the presence of frequency-dependent background noise. We will demonstrate this by detecting solar-like oscillations from time series photometry in later sections. In this section, we derive the coefficient of variation method in general terms, as it is widely applicable beyond the domain for which it has been crafted here.

In the Fourier transform of a time series of regularly sampled Gaussian white (independent and uncorrelated) noise of duration T, the real and imaginary components sampled at independent frequencies separated by $1/T$ are both Gaussian distributed. For time series that contain a large number of samples, the central limit theorem supports this result for arbitrary noise distributions. The power spectrum formed by summing the squares of these terms therefore consists of independent and identically distributed (i.i.d.) random variates from a χ^2 probability density function (p.d.f.) with 2 degrees of freedom[1] (d.o.f.; e.g., Robinson 2017, Section 9.4.2). This χ_2^2 distribution equals the exponential distribution with a rate parameter of $1/2$:

$$F(x) = \begin{cases} \frac{1}{2}e^{-x/2} & x \geq 0 \\ 0 & x < 0. \end{cases} \tag{1}$$

The power spectrum computed from the Lomb-Scargle periodogram is formulated to produce the same χ_2^2 noise statistics for unevenly sampled data[2] (Scargle 1982). We aim to exploit the behavior of this distribution to test the null hypothesis that a region of a power spectrum contains only noise—namely, that the underlying mean and standard deviation are equal[3]. The ratio of the standard deviation to the mean—called the "coefficient of variation" (CV)—for a χ_2^2 distribution is therefore equal to one. We expect to measure a CV close to this value in the power spectrum of pure white noise. The presence of any additional, localized signal will have the effect of increasing the CV, often to some statistical significance.

In reality, correlated stochastic noise is more common than white noise; however, these backgrounds often show only gradual frequency dependences, with the noise still distributed as χ_2^2 about a non-flat limit spectrum (as for solar-like oscillators; Duvall & Harvey 1986; Anderson et al. 1990). In the limit of increasingly narrow frequency ranges, the spectrum will appear *locally* white. Comparing the CV measured in a frequency range to expectations for χ_2^2 noise can reveal spectral features that are sharp relative to the

---

[1] Because we must estimate the variance from the time series, the power *measured* in a pure-noise spectrum is not exactly χ_2^2 distributed; however, this difference is negligible in the limit of a large number of observations (Schwarzenberg-Czerny 1998; Frescura et al. 2008).

[2] Structure in the spectral window for unevenly sampled data causes there to exist no frequencies at which the periodogram powers are truly independent (see, e.g., Frescura et al. 2008). For our purposes, we proceed to treat power sampled at $1/T$ as effectively independent for nearly evenly-spaced data.

[3] Generally, the mean and standard deviation of a χ^2 distribution with f d.o.f. are f and $\sqrt{2f}$, respectively.
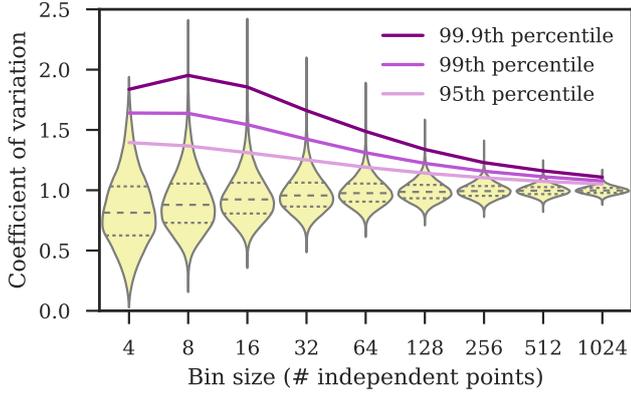
Figure 1. Simulated distributions of CV measurements for bins containing different numbers of χ_2^2 i.i.d.s. Solid lines mark the 95th, 99th, and 99.9th percentiles of each, corresponding to 5%, 1%, and 0.1% false alarm probabilities for pure noise in the power spectra, respectively. The dashed and dotted lines within each distribution indicate the medians and interquartile ranges.

Table 1. CV false alarm probabilities.

bin size (# pts)	5% FAP (CV)	1% FAP (CV)	0.1% FAP (CV)
4	1.397	1.640	1.830
8	1.368	1.632	1.942
16	1.314	1.540	1.865
32	1.253	1.421	1.664
64	1.191	1.311	1.486
128	1.141	1.225	1.341
256	1.102	1.157	1.227
512	1.073	1.110	1.157
1024	1.052	1.077	1.108

inspected bin width, while being insensitive to background features that vary over longer frequency scales. Because the CV calculation normalizes the scatter in the power spectrum by the local mean, it does not require a precise characterization of the overall noise background to identify sharp features of interest.

In practice, we rely on direct simulations to derive a significance criterion for rejecting the null hypothesis that a measured CV arose purely from χ_2^2 noise. While the χ_2^2 distribution does have a characteristic CV of unity, the measured standard deviation divided by the measured mean is a biased estimator of the CV for two reasons: (1) the standard deviation is biased (as a nonlinear transformation of an unbiased estimator: the variance); and (2) the variance and mean of i.i.d.s are not independent for non-Gaussian underlying p.d.f.s (Geary 1936). Closed forms for the p.d.f.s of variance measured for generic sample sizes drawn from exponential distributions are not known (e.g., Lam 1980).

The distributions of measured CVs simulated for nine different sample sizes containing between 4 and 1024 i.i.d. random χ_2^2 variates are displayed in Figure 1. Each distribution is determined from 100,000 sets of randomly generated values. The distributions converge with increasing sample size toward Gaussians centered on one, as expected from the central limit theorem. The solid lines mark the 95th,

99th, and 99.9 percentiles of each distribution. These correspond to 5%, 1%, and 0.1% false alarm probabilities (FAPs) that a detection above these thresholds was caused purely by χ_2^2 noise in the data. FAPs for bins of generic size can be interpolated from the values provided in Table 1. If multiple independent frequency bins are considered, the probability of a CV value in any bin exceeding these thresholds due solely to noise increases with the number of bins examined.

3 CV SPECTRA FOR DETECTING SOLAR-LIKE OSCILLATIONS IN GIANTS

Having presented a statistical metric for detecting sharp spectral features in the presence of frequency-dependent background noise, we now outline a CV-based strategy for detecting solar-like oscillations in red giants and measuring $\nu_{\max}$. The following method is specifically designed for long-cadence *Kepler* light curves. It is validated by null results for dwarf stars with super-Nyquist $\nu_{\max} > 500\,\mu$Hz, and by our recovery of cataloged $\nu_{\max}$ values for the APOKASC giants (Pinsonneault et al. 2014, 2018). This is just one of many possible CV-based algorithms, and it could easily be adjusted for other data sets or to be more/less susceptible to false positives/negatives.

The pulsation signatures in the power spectra of solar-like oscillators fall under an approximately Gaussian envelope centered on the frequency of maximum oscillation power, $\nu_{\max}$. Various other observational characteristics of red giant power spectra are tightly correlated with $\nu_{\max}$, as described by a set of empirical power laws (e.g., Mosser et al. 2010, 2012; Mathur et al. 2011b; Kallinger et al. 2014). These scaling relations set our expectations about the signatures of solar-like oscillations and inform our strategy for computing CV spectra that will best reveal their presence.

The calculation of CV spectra requires binning of the power spectra, and we want to adopt a binning scheme that well-captures the oscillatory signal near any $\nu_{\max}$. The pressure-modes of solar-like oscillations are distributed according to a "universal pattern" (Mosser et al. 2011a), with subsequent radial overtones of the same spherical degree exhibiting a nearly even frequency spacing (Tassoul 1980). Many studies have empirically constrained the asteroseismic scaling relation for this "large frequency separation", $\Delta\nu$, all finding roughly $\Delta\nu \propto \nu_{\max}^{0.75-0.78}$ (e.g., Hekker et al. 2009; Stello et al. 2009; Huber et al. 2010; Mosser et al. 2012; Yu et al. 2018). Since we expect each acoustic radial order to exhibit pulsation signals within the power excess, we compute CV values in bins with widths that scale with the expected $\Delta\nu$ at each searched frequency. The power-law form of the scaling relation makes it convenient to consider our spectra with log-frequency scales. The "large logarithmic frequency separation", $\Delta\log\nu$ (in units of dex; Allen 1951), decreases linearly with $\log\nu_{\max}$ from $\Delta\log\nu \approx 0.067\,$dex at $\nu_{\max} = 10\,\mu$Hz to $\Delta\log\nu \approx 0.033\,$dex at $\nu_{\max} = 200\,\mu$Hz.

The FAP significance thresholds given in Section 2 are only independently applicable to non-overlapping bins. We numerically calculate edges of 43 contiguous bins that span a range between $1\,\mu$Hz and the Nyquist frequency of the long-cadence *Kepler* data, $283.2\,\mu$Hz (following the most recent relation from Yu et al. 2018, $\Delta\nu \approx 0.267 \times \nu_{\max}^{0.764}$). These bin edges are indicated in Figure 2, which displays a rep-
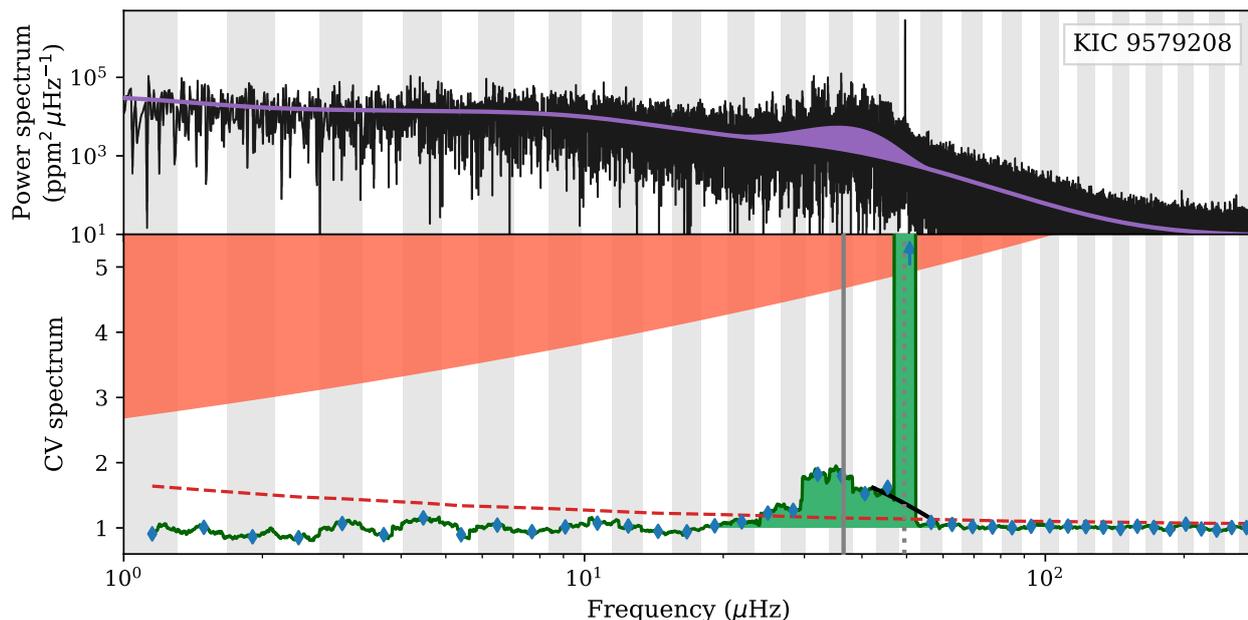
Figure 2. Top: Power density spectrum of the APOKASC giant KIC 5793628 on a log-log scale. A spike exceeding 2.7×10^6 ppm$^2 \mu$Hz^{-1} falls near the pulsation signals. A background model fit to the power spectrum following Kuszlewicz et al. (in prep.; fit with spike excluded) consisting of three super-Lorentzians, white noise, and a Gaussian power excess (filled) is overplotted. Gray and white bands indicate the 43 independent bins that scale with $\Delta\nu$ in both panels (see text). Bottom: Corresponding CV spectrum. The dashed red curve is the 0.1% FAP significance threshold. The blue diamonds mark the CV measurements from independent bins, with the arrow indicating one value above the plotted range. The oversampled CV spectrum is dark green. The spike causes the oversampled CV spectrum to exceed an upper expectation limit for solar-like oscillations (shaded red), and the black line segment displays the interpolation over this feature. The thick gray line indicates the ν_{max} value that we measure from the CV spectrum, and the dotted gray line is the value from APOKASC-2 (Pinsonneault et al. 2018; coincident with the large spike in this case; see Section 5). The oversampled spectrum is filled where it is continuously above a level of CV = 1 surrounding the reported detection of solar-like oscillations.

resentative CV spectrum in comparison with the original power spectrum for the red giant KIC 5793628. The width of each bin is 1.029 times the average expected $\Delta \log \nu$. We compare the CV values calculated in these bins to the 0.1% FAP threshold corresponding to the number of independent power spectrum frequencies they contain. As indicated by the dashed line in the bottom panel of Figure 2, lower-frequency bins generally have higher significance thresholds because their narrower linear frequency ranges contain fewer independent frequencies of the power spectrum.

While the CV spectrum calculated with non-overlapping bins is convenient for significance testing, it sparsely samples pulsational power excesses. Therefore, we compute a second CV spectrum with 2 000 overlapping bins, evenly spaced in log-frequency over the same range and following the same scaling relation for $\Delta\nu$. We refer to this as the "oversampled" CV spectrum, in contrast to the former "independent" spectrum. The oversampled spectrum is displayed in green in the bottom panel of Figure 2. This resembles a smoothed, background-removed power spectrum that enables precise determinations of ν_{max}.

Before further analysis, we search our oversampled CV spectra for contaminating signals that may impact our measurements. Signatures of, e.g., binarity, rotation, instrumental noise, or blends with classical pulsators, often introduce large spikes to our CV spectra. By inspection of the maximum CV values near the cataloged ν_{max} of APOKASC stars (Section 5), we determine a rough upper limit for CV values expected from solar-like oscillations: CV$_{max} \approx$

$2.69 \times (\nu/\mu$Hz$)^{0.154}$. We linearly interpolate across any regions of the oversampled spectrum that share a dependence with CV values that exceed this threshold. The black line segment in the bottom panel of Figure 2 shows the interpolation across such a spike. Cleaning the spectrum in this way enables robust ν_{max} measurements, even when a large contaminating signal falls near the power excess. Our measurements do not lose precision in the few cases where CVs from solar-like oscillations themselves exceed this threshold.

While we expect solar-like oscillations to span multiple acoustic radial orders within the pulsational power excesses, the lower-amplitude signals near the edges of the Gaussian power envelope may not be sufficient to increase the CV above the FAP threshold. Nevertheless, they are likely to exceed the limiting expectation value of 1.0. For $\nu_{max} < 100\,\mu$Hz, Mosser et al. (2012) find that the full-widths at half-maximum (FWHM) of the oscillation envelopes follow $\delta\nu_{env} \approx 0.66\nu_{max}^{0.88}$. This trend is observed to be shallower above $100\,\mu$Hz. We define our algorithm to expect $\delta\nu_{env}$ to scale with $\Delta\nu$, fixing $\delta\nu_{env} \approx 4.2\Delta\nu$ for $\nu_{max} > 100\,\mu$Hz. We adopt a candidate solar-like oscillation power excess if its independent CV spectrum contains values between the 0.1% FAP threshold and the empirical CV$_{max}$ limit in a region where the oversampled CV spectrum remains sustained above 1.0 for the span of at least $\delta\nu_{env}$.

After accepting that a frequency range likely contains ν_{max}, we refine the measurement by computing local moments of the oversampled CV spectrum. Searching over the full CV > 1 power excess, we identify the $\delta\nu_{env}$-wide fre-

Table 2. Bit field flag definitions.

bit	value	meaning
1	1	additional significant, independent CVs
2	2	CV spike exceeds solar-like expectation
3	4	more than one $\nu_{\max}$ candidate detected
4	8	FliPer $\nu_{\max}$ disagreement
5	16	$\nu_{\max}$ candidate near Nyquist frequency

quency window containing the largest average CV (effectively a second smoothing step). We then record the $\nu_{\max}$ value to be the center-of-mass of the oversampled CV spectrum (first moment divided by the integrated CV) in that window.

Our algorithm can naturally return multiple $\nu_{\max}$ values when the time series captures blended sources. We require that these are sufficiently resolved (separated by at least $\delta\nu_{\mathrm{env}}$) to all be reported. We also record whether the independent CV spectra contain additional significant peaks outside the expected Gaussian 3σ ranges surrounding accepted $\nu_{\max}$ values, which may be scientifically interesting.

Signals from intrinsic oscillations with frequencies that exceed the Nyquist ($283.2\,\mu$Hz for *Kepler* long cadence) will be reflected back into the sub-Nyquist regime (e.g., Chaplin et al. 2014b). This folding of the power spectrum introduces systematics in our measurements for $\nu_{\max} \gtrsim 245.6\,\mu$Hz. The span of CV > 1 required at this limit is also restricted, so false positives are somewhat more likely. For stars with $\nu_{\max}$ above the Nyquist, our algorithm is only sensitive to the sub-Nyquist aliases.

As a final consistency check, we compare our $\nu_{\max}$ values to independent estimates based on the stellar granulation. Because the granulation signatures that dominate the power in the light curves are correlated with surface gravity (e.g., Mathur et al. 2011b), the variance can be used as a rough proxy for $\nu_{\max}$ (Hekker et al. 2012). The most recent tool for validating $\nu_{\max}$ in this way is the "FliPer" metric (Bugnet et al. 2017). This empirical relationship scales from the mean of the power density spectrum, corrected for the photon noise. To be less affected by high-frequency signal, we use Bugnet et al.'s recommended prescription based on the parameterized instrumental noise level from Jenkins et al. (2010). We flag $\nu_{\max}$ results that disagree with the FliPer estimate by more than 0.5 dex. Potential causes of such disagreements include: false positive detections of $\nu_{\max}$; detection of aliases of super-Nyquist power excesses; or, the $\nu_{\max}$ and variance not agreeing with scaling relations, e.g., for blended sources. We note that the ~1-dex-wide acceptable $\nu_{\max}$ ranges estimated from FliPer (Bugnet et al. 2017, Figure 3) span about 40% of the log-frequency range that our CV algorithm searches.

In the end, our CV analysis of a power spectrum outputs any (possibly multiple) $\nu_{\max}$ measurements. In addition, we flag a number of conditions to indicate that the measurements may be suspect or that the spectrum may have additional features of scientific interest. Each bit in the binary representation of a single returned value represents the condition of each of five flags, as defined in Table 2.

4   FALSE POSITIVE TEST ON *Kepler* DWARF SAMPLE

We have designed an algorithm to detect solar-like oscillation signatures under the hypothesis that the frequency-dependent granulation backgrounds do not mimic pulsational power excesses in CV spectra. We test this by applying it to long-cadence light curves of 391 dwarfs and subgiants with $\nu_{\max} > 500\,\mu$Hz measured from short-cadence *Kepler* data by Chaplin et al. (2014a). The pulsations of these stars are well above the Nyquist frequency, and their sub-Nyquist aliases will be suppressed below observational limits by the 30-minute *Kepler* exposures. The power spectra of these data still have coloured granulation backgrounds and potentially other contaminating signals.

We utilize the preprocessed light curves available on the *Kepler* Asteroseismic Science Operations Center (KASOC) website[4]. The *Kepler* light curves used in this work have been corrected following Handberg & Lund (2014).

All power spectra in this work were calculated from unweighted time series with the fast implementation of the Lomb-Scargle periodogram from the GATSPY package (VanderPlas 2016). We rescale the spectra by the variance of the light curve to obtain units of ppm$^2\mu$Hz^{-1} (ppm is parts per million) that agree with Parseval's theorem. These units are useful for computing $\nu_{\max}$ estimates from FliPer for validation, yet constant multiplicative offsets to the power spectra do not affect the CV spectra.

The KASOC light curves omit data that were acquired during *Kepler* reaction wheel desaturation events that occurred every three days. The regularity of these missing points introduces a significant comb of aliases into the spectral window, causing power to be broadly redistributed through frequency space. The resulting backgrounds in the power spectra of such time series generally do not follow χ_2^2 distributions, and the correlated features every $3.86\,\mu$Hz can increase the CV above our significance threshold. The effects of the window function on stellar oscillation spectra were analyzed by, e.g., García et al. (2014). Linear interpolation across single missing observations improves the spectral window sufficiently to avoid most spurious CV detections[5].

No features in the CV spectra of 369 out of 391 dwarfs examined match the height and width requirements to be considered candidate solar-like oscillation power excesses. The top panels of Figure 3 display the power and CV spectra of a representative dwarf star for which our algorithm does not detect any pulsations: KIC 11414712. Of the 22 candidate $\nu_{\max}$ values returned, the majority are at frequencies below $10\,\mu$Hz and correspond to localized rotation signatures. However, the FliPer (Bugnet et al. 2017) validation test rejects all but one of these candidates. We display the power and CV spectra of this target, KIC 9579208, in the bottom panel of Figure 3. While FliPer is not a very precise proxy for $\nu_{\max}$, it is very useful for vetting candidate power excess detections. The $\nu_{\max}$ returned for this target is also flagged for falling in the "near-Nyquist" ($\nu_{\max} \gtrsim 245.6\,\mu$Hz) range where a relaxed width requirement makes false posi-

[4] http://kasoc.phys.au.dk/
[5] No gap interpolation was performed on the pre-computed APOKASC power spectra available on the KASOC website[4], so they are not suited for this analysis.
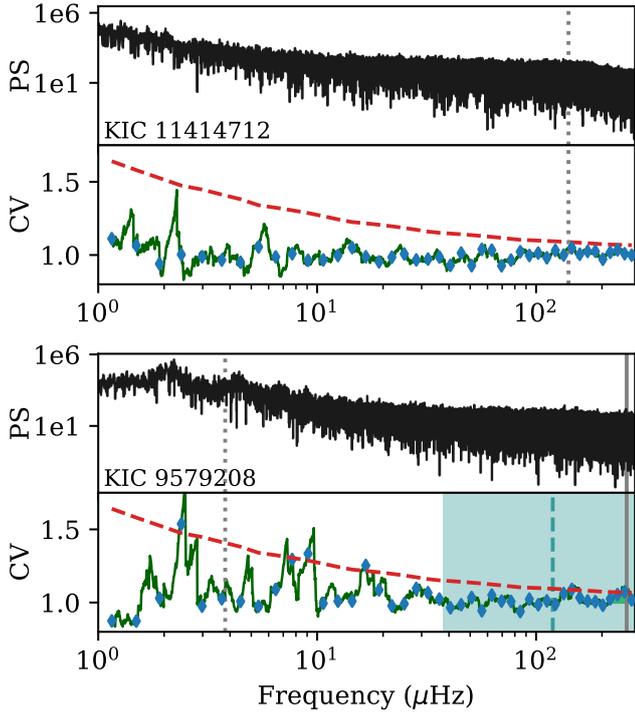
Figure 3. Power and CV spectra of two dwarf stars: KIC 11414712 (top) and KIC 9579208 (bottom). The power spectra are in units of ppm^2 μHz^{-1}. The vertical dotted lines indicate the sub-Nyquist aliases of their intrinsic values of $\nu_{max} \approx 707 \, \mu$Hz and $\nu_{max} \approx 1129 \, \mu$Hz, respectively (Chaplin et al. 2014a). The CV spectra from independent bins are plotted as blue diamonds, and the oversampled CV spectra are displayed in dark green (see text). The red dashed curves give the 0.1% FAP significance thresholds. The power spectrum of KIC 11414712 exhibits a frequency-dependent background, but no significant deviations from expectations for noise are present in the CV spectrum. KIC 9579208 is the only dwarf target that yields a false positive detection (filled green) that agrees with the acceptable range of ν_{max} values given the FliPer metric (shaded blue).

tives more likely. We conclude that fewer than 1% of ν_{max} detections returned by our algorithm that agree with the estimate from FliPer are false positives.

5   APPLICATION TO THE APOKASC GIANTS

The second APOKASC catalog (APOKASC-2; Pinsonneault et al. 2018) characterizes 6676 evolved stars with both *Kepler* light curves and APOGEE spectra (Pinsonneault et al. 2014). It includes ν_{max} values combined from an ensemble of pipelines. We use these as a benchmark for evaluating the sensitivity, accuracy, and precision of our method of detecting ν_{max} from CV spectra.

We apply the CV method to the 6656 APOKASC-2 stars with preprocessed light curves available from KASOC. APOKASC-2 provides seismic measurements (ν_{max}) for 6556 of these. Where available on KASOC, we inspect the short-cadence (1-minute) light curves to identify stars with intrinsic ν_{max} values above the long-cadence Nyquist frequency of 283.2 μHz (e.g., Chaplin et al. 2014b). To test

our method in this regime, we reflect the cataloged sub-Nyquist ν_{max} aliases across the Nyquist frequency to recover the intrinsic ν_{max} values of five super-Nyquist pulsators: KIC 10394814, KIC 6430804, KIC 4351319, KIC 7341231, and KIC 3329196.

Our algorithm returns ν_{max} measurements for 99.7% of the 4080 longest light curves—those spanning > 3.5 years with > 85% duty cycle. We compare our detections to the APOKASC-2 ν_{max} values for these stars in Figure 4. The results are separated into panels based on the flags returned (defined in Table 2). Stars can appear in multiple panels if more than one flag is set. Stars with flag 3 set appear multiple times per panel (marked with crosses), once per candidate ν_{max} value. Detections in the near-Nyquist region are also flagged (flag 5), and we consider these separately from the breakdown by flags 1–4. Seismic measurements for most near-Nyquist stars were omitted in APOKASC-2.

The 3644 (89.3%) ν_{max} measurements that are not accompanied by flags 1–4 are in close agreement with the cataloged values, accounting for reflection at the Nyquist frequency (top panel of Figure 4). Figure 5 displays the histogram of the percentage difference between measured and cataloged ν_{max} for sub-Nyquist stars that do not return any flags. The median difference is 0.12%, and 68.3% of our results agree within 3.0% (σ) of the ν_{max} values from APOKASC. The uncertainties quoted by APOKASC-2 imply that their ν_{max} measurements are distributed about the intrinsic values with $\sigma = 1.3\%$. Subtracting this in quadrature from the σ of the differences between measured and cataloged ν_{max} values for these stars, we conclude that our ν_{max} measurements have an average precision of 2.7%. This includes both our intrinsic measurement errors and any systematic differences, which are typically of-order 1% for ν_{max} values output by different pipelines (e.g., Pinsonneault et al. 2018). We detect a ν_{max} within 20% of the cataloged value (i.e., the same power excess) for 99.5% of stars, regardless of flags.

The most conspicuous disagreement for a ν_{max} value that carries no flags—marked with a square in the top panel of Figure 4—is KIC 5610200. We display the power and CV spectra of this star as the top example in Figure 6. The ν_{max} value from the CV method (thick vertical line at 158.9 μHz) coincides with the oscillation signatures visible in the power spectrum, while the value from APOKASC-2 does not appear to match (dotted line at 95.4 μHz). The low mode heights observed suggest that the intrinsic ν_{max} of this star may be above the long-cadence Nyquist frequency.

The most commonly returned flag is flag 1 (8.5% of stars). This indicates that there are additional significant peaks in the independent CV spectrum besides any detected solar-like oscillations or spikes large enough to raise flag 2. Most ν_{max} results that are accompanied by this flag still agree with the cataloged values to the same precision. All results that are in disagreement with the APOKASC values have additional flags set (marked with crosses in the second panel of Figure 4), with the exception of KIC 10735519 (marked with a square). We display the power and CV spectra of this target as the second example in Figure 6. Zhou (2010) characterized KIC 10735519 as an Algol-type eclipsing binary with a 0.9070-day orbital period. The series of spikes caused by the binary signal increases the CV values
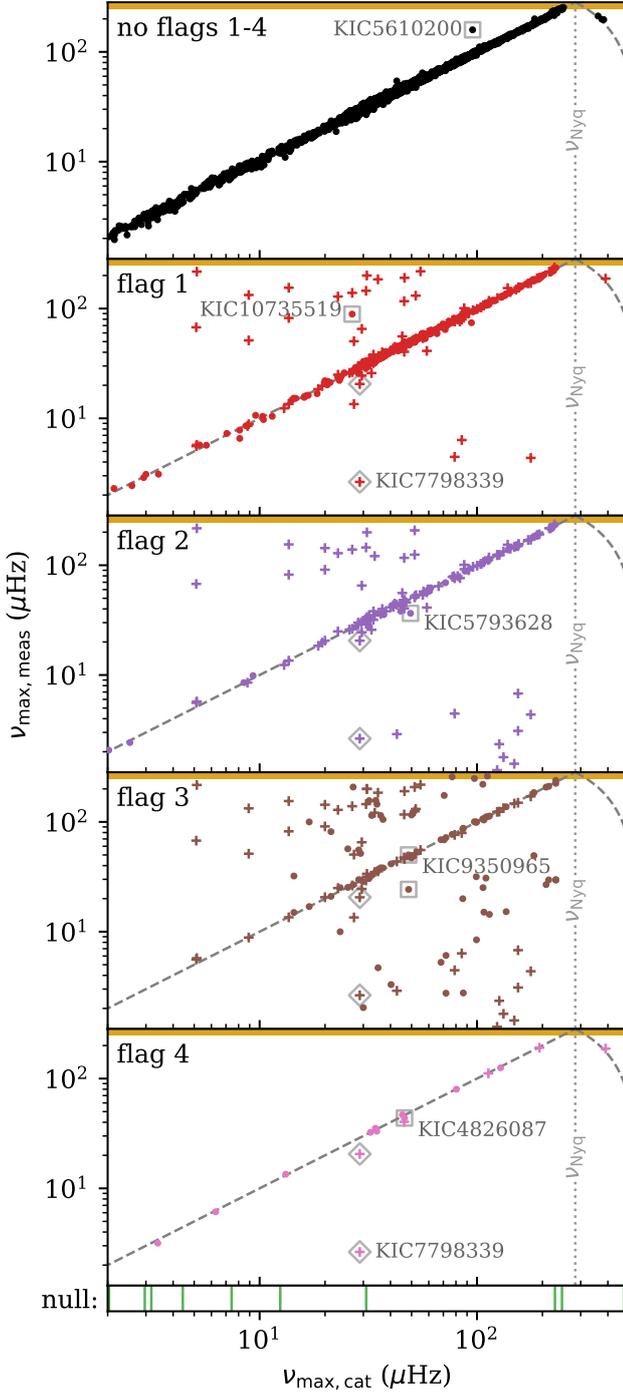
Figure 4. Comparison between the ν_{max} values measured by our algorithm and the values from APOKASC-2 for the 4080 longest KASOC-processed *Kepler* light curves. The results are separated by the flags returned by our algorithm, which are defined in Table 2. Stars marked with crosses are present in multiple panels. Stars with multiple ν_{max} values returned are represented for each detection (flag 3). Flag 5 is set for stars with ν_{max} measured in the orange region near the Nyquist limit. The dashed line marks exact agreement, accounting for reflection above the Nyquist frequency, which is indicated with a dotted line. Vertical bars in the bottom panel mark the cataloged ν_{max} values of stars for which our algorithm does not return a ν_{max} detection. Objects with spectra displayed as examples in Figures 2 and 6 are outlined with boxes and labeled (see text for discussion).
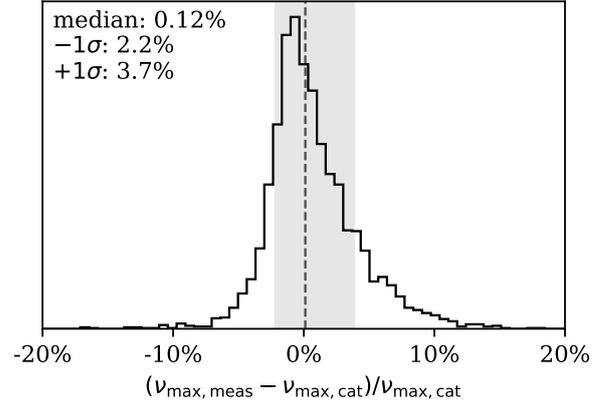
Figure 5. Histogram of the percentage difference between the APOKASC-2 cataloged ν_{max} values and the results from the CV method for sub-Nyquist stars that do not return any flags.

above the expectation for χ_2^2 noise through most of the spectrum, corrupting our automated ν_{max} detection.

We displayed an example of a star with an anomalously high spike in its CV and power spectra in Figure 2 (KIC 5793628; square in third panel of Figure 4). Situated near the pulsational power excess, this spike threatens to skew any ν_{max} measurement if not accounted for. Figure 2 demonstrates how our algorithm avoids giving these spikes too much influence by interpolating across them in the oversampled CV spectra. Flag 2 is set for 2.7% of stars for which we perform this interpolation step. This correction explains why our ν_{max} measurement for KIC 5793628 is 26% lower than the APOKASC value, which coincides with the spike frequency. All ν_{max} measurements from spike-interpolated spectra that are more discrepant have other flags set.

Of the 69 targets (1.7%) that our method returns multiple ν_{max} candidates for (flag 3), 63 have two ν_{max} values recorded, and six have three values. KIC 9350965, a star with two reported power excesses (two squares in the fourth panel of Figure 4), is the third example in Figure 6. Both detections correspond to solar-like oscillations in a blended light curve. The natural ability to find multiple power excesses is a strength of our approach, as it can enable the efficient search for asteroseismic binary systems, of which only three are known (Appourchaux et al. 2015; White et al. 2017; Beck et al. 2018). Since APOKASC only lists one ν_{max} for each target, additional ν_{max} detections from our algorithm will necessarily disagree with the catalog, even when these correspond to real solar-like oscillation power excesses.

Flag 4 indicates that none of our returned ν_{max} values are within 0.5 dex of the estimate from FliPer, implying a disagreement between the power excess and the variance of the light curve expected from scaling relations. Only the 15 targets (0.4%) in the fifth panel of Figure 4 carry this flag. Our ν_{max} values may not be good initial guesses for fitting the backgrounds of these power spectra. An interesting example is KIC 4826087, whose variance is influenced by granulation corresponding to a pulsational power excess below our searched frequency range, near 0.7 μHz. We extend the displayed frequency range for this target in Figure 6. There are few known examples of solar-like oscillations at such low frequencies (e.g., Stello et al. 2014).

Our last example in Figure 6 is the only star with the
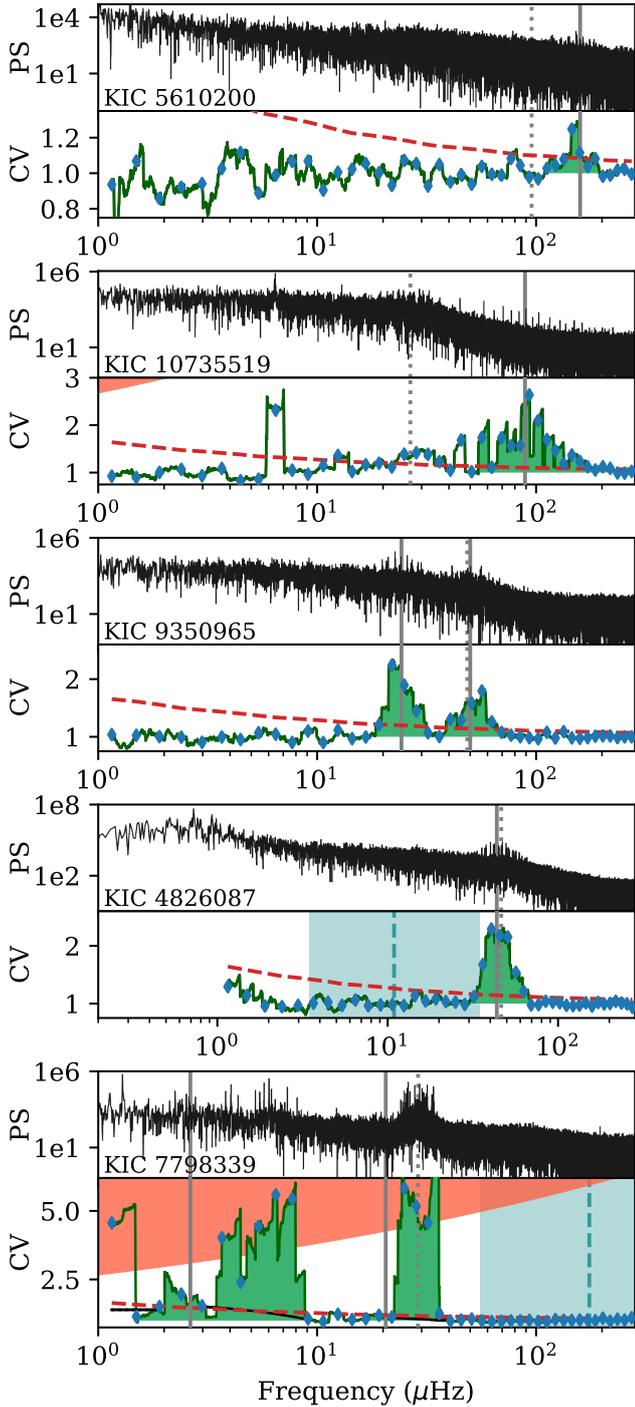
Figure 6. Power and CV spectra for targets, from top to bottom, KIC 5610200, KIC 10735519, KIC 9350965, KIC 4826087, and KIC 7798339. The CV spectra are displayed as in Figures 2 and 3 (blue diamonds: independent CV; dark green curves: over-sampled CV; black segments: interpolated CV; dashed red lines: 0.1% FAP threshold; filled green; CV > 1 regions around detections; shaded red: upper CV limit; dotted vertical lines: $\nu_{\max}$ from APOKASC-2; solid vertical lines: measurements from our algorithm). The acceptable range of $\nu_{\max}$ given the estimate from FliPer (vertical dashed line) is shaded in blue for stars that are flagged for disagreement. We extend the displayed frequency range for KIC 4826087 to reveal the power excess at a $\nu_{\max}$ below the range searched by our algorithm. See text for discussion.

first four flags all set: KIC 7798339 (marked with diamonds throughout Figure 4). This is not a solar-like oscillator at all, but rather a γ Doradus variable (Uytterhoeven et al. 2011).

The cataloged $\nu_{\max}$ values for the 11 stars that our algorithm does not return a detection for are marked in the bottom panel of Figure 4. Five of these stars have $\nu_{\max} \leq 5\,\mu$Hz, where our significance threshold is highest and the mode lifetimes are not as well sampled. We also do not detect the Nyquist-aliased pulsations in another with an intrinsic $\nu_{\max} \approx 475\,\mu$Hz implied by APOKASC—its power excess is not obvious in the long-cadence observations. We are, however, able to recover the Nyquist-reflected $\nu_{\max}$ for a target with intrinsic $\nu_{\max} \approx 390\,\mu$Hz.

Our $\nu_{\max}$ measurements from CV spectra of the full sample of 6656 APOKASC-2 stars with KASOC data are tabulated in Appendix A. We also catalog the associated flags and emphasize that the vast majority of our $\nu_{\max}$ values are robust, even when accompanied by flags. We find that the 94 light curves that are flagged for having multiple candidate power excesses warrant visual inspection. We note those that appear to the authors to genuinely display multiple solar-like oscillation power excesses in Appendix B.

6   DISCUSSION AND CONCLUSIONS

Detecting the presence of solar-like oscillations is a prerequisite to red giant asteroseismology. The frequency of maximum pulsation power, $\nu_{\max}$, is one of the fundamental seismic parameters of a solar-like oscillator. We have developed a rapid and robust method of detecting and measuring $\nu_{\max}$ of these oscillations in the presence of granulation backgrounds from coefficient of variation (CV) spectra. The $\nu_{\max}$ values obtained in our analysis of the APOKASC-2 giants typically agree with the cataloged values (Pinsonneault et al. 2018) to an average precision of 2.7% from four years of *Kepler* data. Our algorithm detects > 99% of solar-like oscillations in giants with < 1% false positives determined from a test on dwarf stars that do not exhibit visible pulsations in long-cadence data. Computing the CV spectra and searching for power excesses in 6656 APOKASC stars takes 70 seconds on a single CPU and is easily parallelized (most additional overhead is from reading in the time series and computing the power spectra).

Our algorithm was designed to produce robust measurements in the presence of source blending, and it can even detect multiple power excesses in a single time series. Additional signals that may be scientifically interesting or obstructive can also be identified from CV spectra. For example, Colman et al. (2017) flagged *by eye* 168 red giant power spectra that featured anomalous peaks, of which 110 overlapped with the APOKASC-2 sample that we analyzed. Our algorithm flags 103 of these as containing high spikes (66 targets) and/or additional significant CV values. Our method can quickly produce larger catalogs of anomalous features in solar-like oscillator power spectra, some of which Colman et al. (2017) argue are signatures of compact binaries orbiting the oscillating giants. The flags returned for all 6656 APOKASC-2 stars that we analysed are included with the $\nu_{\max}$ measurements in Appendix A.

This data-driven approach to detecting $\nu_{\max}$ pairs naturally with methods such as the power spectrum of the power

spectrum (e.g., Hekker et al. 2010; or equivalently, the auto-correlation of the time series; Mosser & Appourchaux 2009) to measure the large frequency separation, $\Delta\nu$. This second asteroseismic parameter is usually easier to detect than ν_{max} (e.g., Chaplin et al. 2014a), and can be used as additional validation of a candidate power excess detection or to resolve whether a measured ν_{max} is intrinsically super-Nyquist (Chaplin et al. 2014b).

By providing a shortcut to ν_{max}, the CV method will accelerate the process of fitting the granulation background to power spectra (esp. where the detected ν_{max} agrees with the prediction from FliPer; Bugnet et al. 2017). These fits could validate and refine our ν_{max} values, and will benefit from knowledge of spike locations that the CV algorithm flags. This will shorten the time to bagging individual peaks for detailed asteroseismic analysis, and we plan to incorporate the CV method into the code of García Saravia Ortiz de Montellano et al. (2018). The CV spectrum may also be a valuable data representation for training machine learning algorithms for detecting ν_{max}, as has been attempted from the regular power spectrum by Hon et al. (2018).

Ingredients of our approach have been applied to red giants and other astrophysical variables in the past. Ransom et al. (2002) provide a detailed description of Fourier techniques that are relevant to the analysis of long time series. In particular, Israel & Stella (1996) took a similar approach to detecting coherent signals in coloured power spectra, motivated for the study of pulsars. The "Moving-Windowed-Power-Search" of Lund et al. (2012) also considers statistical excesses in the binned spectra of solar-like oscillators; however, it relies on first dividing by an optimized background model fit. By combining elements of these approaches, we avoid the need to fit a background to reliably identify the presence and ν_{max} of stochastic, solar-like oscillations.

We have analyzed here only a subset of the red giants observed by *Kepler*. Yu et al. (2018) recently identified 16,000 red giant stars in the *Kepler* data. We plan to extend and optimize this algorithm for application to a larger *Kepler* sample, as well as to the CoRoT data and observations along the ecliptic from K2. As currently defined, the algorithm encounters challenges in searching for power excesses in short time series; however, widening the bins used to calculate the CV spectra can lower the FAP significance thresholds at the cost of admitting more false positives. We will explore these limitations and ways to mitigate them in future work, with the eventual goal of applying this detection technique to the upcoming data from TESS.

ACKNOWLEDGEMENTS

REFERENCES

Aigrain S., Favata F., Gilmore G., 2004, A&A, 414, 1139
Allen C. W., 1951, The Observatory, 71, 157
Anderson E. R., Duvall Jr. T. L., Jefferies S. M., 1990, ApJ, 364, 699
Appourchaux T., et al., 2015, A&A, 582, A25
Baglin A., Auvergne M., Barge P., Deleuil M., Catala C., Michel E., Weiss W., COROT Team 2006, in Fridlund M., Baglin A., Lochard J., Conroy L., eds, ESA Special Publication Vol. 1306, The CoRoT Mission Pre-Launch Status - Stellar Seismology and Planet Finding. p. 33
Ballard S., et al., 2014, ApJ, 790, 12
Baudin F., et al., 2011, A&A, 529, A84
Beck P. G., et al., 2018, A&A, 612, A22
Bedding T. R., et al., 2011, Nature, 471, 608
Borucki W. J., et al., 2010, Science, 327, 977
Brown T. M., Gilliland R. L., Noyes R. W., Ramsey L. W., 1991, ApJ, 368, 599
Bugnet L., García R. A., Davies G. R., Mathur S., Corsaro E., 2017, in Reylé C., Di Matteo P., Herpin F., Lagadec E., Lançon A., Meliani Z., Royer F., eds, SF2A-2017: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. pp 85–88 (arXiv:1711.02890)
Campante T. L., et al., 2015, ApJ, 799, 170
Chaplin W. J., et al., 2014a, ApJS, 210, 1
Chaplin W. J., Elsworth Y., Davies G. R., Campante T. L., Handberg R., Miglio A., Basu S., 2014b, MNRAS, 445, 946
Christensen-Dalsgaard J., Gough D. O., Libbrecht K. G., 1989, ApJ, 341, L103
Colman I. L., et al., 2017, MNRAS, 469, 3802
Duvall Jr. T. L., Harvey J. W., 1986, in Gough D. O., ed., NATO Advanced Science Institutes (ASI) Series C Vol. 169, NATO Advanced Science Institutes (ASI) Series C. pp 105–116
Frescura F. A. M., Engelbrecht C. A., Frank B. S., 2008, MNRAS, 388, 1693
García Saravia Ortiz de Montellano A., Hekker S., Themeßl N., 2018, MNRAS, 476, 1470
García R. A., et al., 2014, A&A, 568, A10
Geary R. C., 1936, Supplement to the Journal of the Royal Statistical Society, 3, 178
Goldreich P., Keeley D. A., 1977, ApJ, 212, 243
Handberg R., Lund M. N., 2014, MNRAS, 445, 2698
Harvey J., 1985, in Rolfe E., Battrick B., eds, ESA Special Publication Vol. 235, Future Missions in Solar, Heliospheric & Space Plasma Physics.
Hekker S., Christensen-Dalsgaard J., 2017, A&ARv, 25, 1
Hekker S., et al., 2009, A&A, 506, 465
Hekker S., et al., 2010, MNRAS, 402, 2049
Hekker S., et al., 2012, A&A, 544, A90
Hekker S., Elsworth Y., Mosser B., Kallinger T., Basu S., Chaplin W. J., Stello D., 2013, A&A, 556, A59
Hon M., Stello D., Zinn J. C., 2018, ApJ, 859, 64
Howell S. B., et al., 2014, PASP, 126, 398
Huber D., et al., 2010, ApJ, 723, 1607
Israel G. L., Stella L., 1996, ApJ, 468, 369
Jenkins J. M., et al., 2010, ApJ, 713, L120
Kallinger T., et al., 2014, A&A, 570, A41

Kjeldsen H., Bedding T. R., 1995, A&A, 293, 87

Lam H.-k., 1980, Communications in Statistics - Simulation and Computation, 9, 639

Lund M. N., Chaplin W. J., Kjeldsen H., 2012, MNRAS, 427, 1784

Lund M. N., Handberg R., Davies G. R., Chaplin W. J., Jones C. D., 2015, ApJ, 806, 30

Mathur S., et al., 2011a, ApJ, 733, 95

Mathur S., et al., 2011b, ApJ, 741, 119

Miglio A., et al., 2016, MNRAS, 461, 760

Mosser B., Appourchaux T., 2009, A&A, 508, 877

Mosser B., et al., 2010, A&A, 517, A22

Mosser B., et al., 2011a, A&A, 525, L9

Mosser B., et al., 2011b, A&A, 532, A86

Mosser B., et al., 2012, A&A, 537, A30

Pinsonneault M. H., et al., 2014, ApJS, 215, 19

Pinsonneault M. H., et al., 2018, preprint, (arXiv:1804.09983)

Ransom S. M., Eikenberry S. S., Middleditch J., 2002, AJ, 124, 1788

Ricker G. R., et al., 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Robinson E., 2017, Data Analysis for Scientists and Engineers. Princeton University Press, https://press.princeton.edu/titles/10911.html

Scargle J. D., 1982, ApJ, 263, 835

Schwarzenberg-Czerny A., 1998, MNRAS, 301, 831

Stello D., Chaplin W. J., Basu S., Elsworth Y., Bedding T. R., 2009, MNRAS, 400, L80

Stello D., et al., 2013, ApJ, 765, L41

Stello D., et al., 2014, ApJ, 788, L10

Tassoul M., 1980, ApJS, 43, 469

Uytterhoeven K., et al., 2011, A&A, 534, A125

VanderPlas J., 2016, gatspy: General tools for Astronomical Time Series in Python, Astrophysics Source Code Library (ascl:1610.007)

White T. R., et al., 2017, A&A, 601, A82

Yu J., Huber D., Bedding T. R., Stello D., Hon M., Murphy S. J., Khanna S., 2018, ApJS, 236, 42

Zhou A.-Y., 2010, preprint, (arXiv:1002.2729)

APPENDIX A: MEASUREMENTS OF ν_{max} FOR THE APOKASC SAMPLE OF GIANTS

We applied the coefficient-of-variation algorithm for detecting solar-like oscillations (Section 3) to the KASOC-processed light curves available for 6656 stars from APOKASC-2 (see Section 5). We provide the full output of our algorithm for every star in Table A1 (full machine readable table available online). Stars with multiple candidate solar-like oscillation power excess detections are included once per ν_{max} value. The binary bit flags returned by our algorithm (defined in Table 2) are summed and presented as base-10 numbers in the "flags" column. The cataloged values from APOKASC-2 (Pinsonneault et al. 2018) and the estimates from FliPer (Bugnet et al. 2017) are included for comparison.

APPENDIX B: VALIDATION OF MULTIPLE ν_{max} DETECTIONS

Out of 6656 light curves analysed, our algorithm returns multiple candidate ν_{max} detections for 94. The authors inspected these light curves and their power spectra by eye to confirm which appear to exhibit two well-separated solar-like

Table A1. CV-based ν_{max} measurements for 6656 giants in APOKASC. Full machine readable table available online.

KIC	$\nu_{\mathrm{max,meas}}$ (μHz)	$\nu_{\mathrm{max,cat}}$ (μHz)	$\nu_{\mathrm{max,FliPer}}$ (μHz)	flags
8037095	2.07	2.02	1.29	2
7672453		2.02	1.91	0
7351928	2.07	2.03	1.56	0
9697618		2.03	1.29	3
2011145	2.13	2.09	1.61	0
6615133	2.25	2.12	2.04	0
8085217	2.30	2.14	1.66	1

Table B1. APOKASC-2 targets that exhibit multiple solar-like oscillation power excesses.

KIC	notes
2161409	
2422558	a
2449518	
2568888	
4173334	
4260884	
6206407	
6501237	
6689517	a
6888756	
7510604	
8004637	
8479383	
8636389	
9350965	
9392650	
9412408	
9725292	
9893437	b
9893440	b
10083224	
10592924	
10602015	
10937954	
10973854	
11090673	b
11090674	b
11299484	
12117790	a
12165692	

[a] Second power excess tentatively supported by visual inspection.
[b] These same blends appear twice in the APOKASC sample.

oscillation power excesses in reality. We list these 30 targets in Table B1, with footnotes indicating those for which the second detection is not completely convincing or unique. Besides these stars with multiple sets of solar-like oscillations, some additional detections arise from, e.g., rotation signatures, remaining signal from photometric binaries, classical oscillations, and spurious detections from noise.

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