MNRAS **000**, 1–13 (2018) Preprint 20 November 2018 Compiled using MNRAS LATEX style file v3.0

A Search for Red Giant Solar-like Oscillations in All *Kepler* Data

Marc Hon[1]★, Dennis Stello[1,2,3], Rafael A. García[4,5], Savita Mathur[6,7,8], Sanjib Sharma[2], Isabel L. Colman[2,3], and Lisa Bugnet[4,5]

[1]*School of Physics, The University of New South Wales, Sydney NSW 2052, Australia*
[2]*Sydney Institute for Astronomy (SIfA), School of Physics, University of Sydney, NSW 2006, Australia*
[3]*Stellar Astrophysics Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark*
[4]*IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*
[5]*Université Paris Diderot, AIM, Sorbonne Paris Cité, CEA, CNRS, F-91191 Gif-sur-Yvette, France*
[6]*Instituto de Astrofísica de Canarias, E-38200, La Laguna, Tenerife, Spain*
[7]*Universidad de La Laguna, Departamento de Astrofísica, E-38205, La Laguna, Tenerife, Spain*
[8]*Space Science Institute, 4750 Walnut Street Suite 205, Boulder, CO 80301, USA*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

The recently published *Kepler* Mission Data Release 25 (DR25) reported on ∼197,000 targets observed during the mission. Despite this, no wide search for red giants showing solar-like oscillations have been made across all stars observed in *Kepler*'s long-cadence mode. In this work, we perform this task using custom apertures on the *Kepler* pixel files and detect oscillations in 21,914 stars, representing the largest sample of solar-like oscillating stars to date. We measure their frequency at maximum power, $\nu_{\rm max}$, down to $\nu_{\rm max} \simeq 4\mu$Hz and obtain $\log(g)$ estimates with a typical uncertainty below 0.05 dex, which is superior to typical measurements from spectroscopy. Additionally, the $\nu_{\rm max}$ distribution of our detections show good agreement with results from a simulated model of the Milky Way, with a ratio of observed to predicted stars of 0.992 for stars with $10\mu{\rm Hz} < \nu_{\rm max} < 270\mu{\rm Hz}$. Among our red giant detections, we find 909 to be dwarf/subgiant stars whose flux signal is polluted by a neighbouring giant as a result of using larger photometric apertures than those used by the NASA *Kepler* Science Processing Pipeline. We further find that only 293 of the polluting giants are known *Kepler* targets. The remainder comprises over 600 newly identified oscillating red giants, with many expected to belong to the galactic halo, serendipitously falling within the *Kepler* pixel files of targeted stars.

Key words: asteroseismology – methods: data analysis – techniques: image processing – stars: oscillations – stars: statistics

1 INTRODUCTION

Red giants showing solar-like oscillations from NASA's *Kepler* Mission (Borucki et al. 2010) are critical to our understanding of stellar evolution and stellar populations. The asteroseismic study of such oscillations provides us with the ability to probe the stellar interior conditions such as core rotation rates (Beck et al. 2011; Mosser et al. 2012; Deheuvels et al. 2014), and potential magnetic fields in the core (Fuller et al. 2015; Stello et al. 2016a,b). Moreover, the abundance of oscillating red giants from *Kepler* has provided us with the opportunity to study and characterize large stellar populations through ensemble analyses (Chaplin & Miglio 2013; Hekker & Christensen-Dalsgaard 2017; García & Stello García & Stello) to inform galactic archaeology studies (e.g. Casagrande et al. 2015; Sharma et al. 2016; Aguirre et al. 2018).

From the 197,096 stars observed by *Kepler* in long-cadence ($\Delta T \simeq 30$min.) from the *Kepler* Data Release 25 (DR25), there is a current combined total of ∼19000 identified oscillating red giants reported in literature (Yu et al. 2018). This sample comes from ensemble studies of previous data releases (Hekker et al. 2011; Huber et al. 2011; Stello et al. 2013) and newly identified targets (Huber et al. 2014; Mathur et al. 2016; Yu et al. 2016), from which only ∼16000 have been analysed seismically based on the full end-of-mission 4-year light curves (Yu et al. 2018). This number represents a significant fraction of the total number of *Kepler* targets. However, recent population estimates from Gaia-derived radii on ∼178,000 *Kepler* DR25 stars showed that the number of *Kepler* red giants is approximately 21,000 (Berger et al. 2018, hereafter B18),

★ E-mail: mtyh555@uowmail.edu.au

proving that there is still a significant number of oscillating giants yet to be found from the end-of-mission data.

While red giants can be identified from their stellar parameters such as effective temperature ($T_{\rm eff}$) or surface gravity ($\log(g)$), the only way to guarantee that a giant shows solar-like oscillations is to directly detect oscillation modes within its power spectrum. Thus, we perform a fully-automated detection for such oscillations across all long-cadence *Kepler* targets from the DR25, a task that has not yet been done prior to this study. The conventional way of detecting oscillations is carried out using seismic pipelines, where model fitting and statistical tests are performed on the power spectrum of the star (see Stello et al. 2017 for a description of pipelines to date). While rigorous, such approaches can be time consuming and require careful consideration with respect to fitting mathematical models to approximate the power spectra. Moreover, in many cases the mathematical models do not capture the full complexity of real data.

Instead, we use a non-conventional way of detecting oscillations using machine learning. In particular, we use the method introduced by Hon et al. (2018, hereafter H18), which uses deep learning, an approach to artificial intelligence, to visually identify solar-like oscillations on plots of power density spectra. For brevity, from here onwards we use the term 'power spectrum' to simply refer the power density spectrum of the star. Besides identifying oscillating red giants, the frequency at maximum power of the star, $\nu_{\rm max}$, can also be measured by the deep learning visual expert based on the position of the oscillations within the power spectra. As shown by H18, this method achieves a high classification accuracy and is capable of predicting $\nu_{\rm max}$ with a human-level performance ($\sigma_{\nu_{\rm max}}/\nu_{\rm max} \sim 5\%$), while also being highly robust because no explicit model fitting of the power spectrum is required.

Our study in this paper is hence focused on the outcome of running our classification over all $\sim$ 197,000 *Kepler* targets ever observed during the mission, and characterizing the oscillating stars with the seismic parameter $\nu_{\rm max}$, which we then use to obtain $\log(g)$ estimates. A novel approach in our study includes the use of a larger photometric aperture, which produces an additional yield of serendipitous red giants. To date, this work represents the largest wide scale search for solar-like oscillations for giants with $\nu_{\rm max} \gtrsim 4\mu$Hz, which will be essential to the full *Kepler* legacy catalog for red giants (Garcia et al. in preparation).

2 METHODS

First, we describe the preparation of the data and our specific use of a different photometric aperture, followed by the detection of oscillating stars within the *Kepler* long-cadence dataset using deep learning methods.

2.1 *Kepler* Observations and Data Preparation

Our dataset comprises a total of 196,581 *Kepler* end-of-mission long-cadence light curves. The remaining 515 stars from the DR25 were only observed for very short durations, thus they are not included in our study. Compared to the simple aperture photometry used by the *Kepler* Science Processing Pipeline (Jenkins et al. 2010) that maximizes the signal-to-noise ratio for a specific *Kepler* target within a 6.5 hour window (Bryson et al. 2010), we use typically larger custom apertures that are designed to produce stable light curves on longer time scales for stellar rotation analysis (García et al. 2014b). We define the extent of the aperture by moving outwards pixel-by-pixel from the center of the point-spread function of the target star and calculating a reference flux value of each subsequent pixel. We compute the reference flux in a pixel as the 99.9th percentile flux value of the pixel's light curve to avoid taking outliers into account. If the reference flux value of a pixel is smaller than that of the previous pixel and is above a given threshold, it is included in the aperture. However, if the reference flux is below the threshold, it is considered to contain only background signal. We then do not include the pixel in the aperture and stop extending the aperture in the direction of this pixel. A pixel is also excluded if its reference flux value is greater than the previous pixel because this indicates that there is another object in the vicinity of the target star. Once we have defined the extent of the aperture, we calculate the total flux within it for each cadence to generate the light curve. We treat each quarter separately to allow the apertures to change between spacecraft roles. Finally, we correct the measured light curves by applying the methods described by García et al. (2011). In particular, the light curves are high-pass filtered with a triangular filter wit a cut-off providing a full transmission for signals shorter than 20 days and a smooth transition of the filter up to 40 days. No signals longer than this period remains in the data.

To increase the effectiveness of detecting solar-like oscillations, we need to prevent spectral leakage of high power from low frequency, which is an effect of the spectral window that can obscure the intrinsic spectral frequency profile of convective granulation, oscillations, and white noise (García et al. 2014a). Hence, we use the in-painting technique based on a multi-scale cosine transform described by Pires et al. (2015) to fill small gaps in the time series, which are mostly caused by regular missing points at a three day cadence from the angular momentum dump of the reaction wheels used for the attitude control of the spacecraft.

Once we obtain the power spectrum of a star, we convert it into a 2D grayscale image with a pixel dimension of 128x128 following the methods described by H18. The image is a white plot of the power spectrum in logarithmic axes over a black background within the frequency range of $3\mu{\rm Hz} \leq \nu \leq 283\mu{\rm Hz}$ and a power density (P) range of $3{\rm ppm}^2\mu{\rm Hz}^{-1} \leq P \leq 3\times10^7{\rm ppm}^2\mu{\rm Hz}^{-1}$.

2.2 Deep Learning Methods

We use the 2D deep learning classifier as described by H18 to automatically detect the presence of red giant solar-like oscillations in the power spectrum of each star in the *Kepler* long-cadence dataset. This method detects specific image features within the 2D grayscale image of the power spectrum in order to classify it as either a positive detection (showing solar-like oscillations) or a non-detection. In particular, the image of the power spectrum is required to show a broad hump of power excess on a granulation background to be classified as a positive detection by the classifier.

The dataset used to train the classifier is the same as that used by H18, comprising 31,123 *Kepler* targets, with 15,924 known oscillating red giants analysed by Yu et al. (2018) and the remaining 15,199 not showing red giant solar-like oscillations. Because not all long-cadence *Kepler* targets were observed across 4 years, we train two versions of the classifier. While the structure of both classifiers are identical, they differ only by the training data used. The first classifier uses the degraded form of the training data, where each target's power spectrum is made from only a 3-month segment of its light curve. The second classifier instead trains on a non-degraded (full-length) version of the data. By using two different classifiers that can classify both shorter and longer durations of data, we reduce the possibilities of missing the detection of poten-

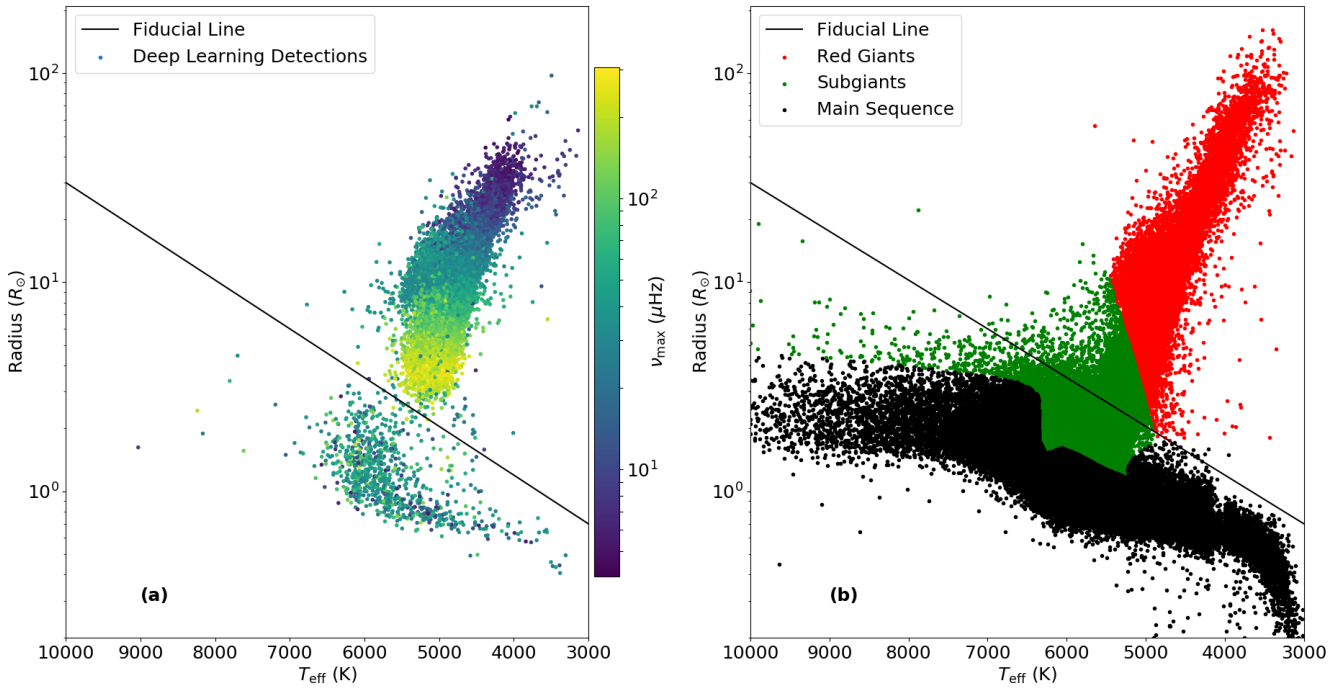

Figure 1. R-$T_{\rm eff}$ diagrams of *Kepler* long-cadence targets, with values of R and $T_{\rm eff}$ adopted from the catalog by B18. (a) Positive detections from our work, with points coloured by their measured $\nu_{\rm max}$ from the deep learning regressor. A fiducial line given by Equation 1 is drawn to delineate the boundary between the red giants and the non-red giants. 1260 detections do not have R and $T_{\rm eff}$ listed in the catalog and hence are not plotted. (b) 177,911 *Kepler* long-cadence targets from B18, coloured by their given evolutionary state from the same source.

tial red giants that were observed for different durations during the *Kepler* mission.

Once we determine the stars that show oscillations, we predict the frequency at maximum power, $\nu_{\rm max}$, using the 2D deep learning regressor described by H18. As the regressor uses the same 2D input as the deep learning classifier, it generally uses the same contextual clues from the power spectrum as the classifier to give a visual estimate of the oscillation power excess location. This visual estimate is the prediction of $\nu_{\rm max}$ for the star, with a typical measurement uncertainty of 5% as shown by H18.

Our deep learning classifier and regressor are constructed with the Keras library (Chollet 2015) using the Tensorflow back end (Abadi et al. 2015), with each utilizing an NVIDIA Titan Xp GPU and the NVIDIA cuDNN library (Chetlur et al. 2014) for training and prediction.

3 DETECTION OF OSCILLATING RED GIANTS

In this Section, we examine the results from the classifier and the validation of its results to produce a final list of detected solar-like oscillators.

3.1 Positive Detections from Deep Learning Classifiers

We first identify 21,468 stars for which both the 4-year and 3-month classifiers detect solar-like oscillations. Next, we identify ~ 8,000 stars for which the classifiers disagree and visually determine whether the star does indeed show a solar-like oscillation signal. This step is necessary because each classifier has a different ability to detect oscillation signals from one another as a result of training on different training sets. In this case, most of the ~ 8,000 stars are false positives from the 3-month classifier. This is because this classifier has been trained to look for oscillations in degraded data with lower signal-to-noise levels, hence in many instances it mistakenly considers white noise in full-length data to contain a pattern that is characteristic of oscillation power excess. Nonetheless, this step allows us to identify 1,130 false negatives from the 4-year classifier, from which many are stars observed for only a few quarters. By adding these false negatives to the detections from the mutual agreement of the classifiers, we have a list of 22,598 detections.

In general, while our method is automated, we find that it is useful to still visually inspect certain positive detections. This is because the non-oscillating stars that are used in our training set are randomly selected dwarf stars, hence our classifiers may not have seen many examples of stars of different types of stellar variability that are present in the *Kepler* long-cadence data, such as detached eclipsing binaries and δ Scutis, leading to potential false positives. In the following we describe how we identified such suspicious detections that were worth inspecting visually for verification.

3.2 Refining the List of Detections

By cross-matching our list of detections with stars from the *Kepler* Eclipsing Binary Catalog (Kirk et al. 2016)[1], we filter out 90 detached binaries that are false positives after visual inspection of their power spectra.

[1] http://keplerebs.villanova.edu/

We then plot a $R - T_{\rm eff}$ diagram for our detections using values from the Gaia(DR2)-*Kepler*(DR25) cross-matching catalog by B18, shown in Figure 1a, and find two well-separated groups of stars with a boundary that we delineate with a fiducial line, given by the following:

$$R = \left(\frac{T_{\rm eff} - 3000}{7000}\right)\log\left(\frac{300}{7}\right) \times 10^6 R_\odot. \quad (1)$$

As seen in Figure 1b, detections above the line evidently correspond to red giants with $R \gtrsim 2R_\odot$ and $T_{\rm eff} \lesssim 6250$ K, while detections below the line correspond to subgiants or main sequence stars that have $R \lesssim 2R_\odot$. We note that our fiducial line coincides with the minimum radius for a red giant following the B18 classification. While we initially detect 1459 stars below the fiducial line, we filter out 550 false positives (from which many are detached eclipsing binaries and δ Scutis) by visual inspection to retain 909 main sequence/subgiant stars showing red giant solar-like oscillations, which we will address in Section 4.

Finally, we cross-matched our results with a list of subgiant and red giant detections from the A2Z seismic pipeline (Mathur et al. 2010). A2Z performs a detection of excess power based on the measurement of the mean large separation of oscillation modes and the position of the maximum power in the power spectrum. A detection is confirmed when these two measurements follow the $\Delta\nu$-$\nu_{\rm max}$ relation (Stello et al. 2009) to within 10%. A2Z also uses the FliPer metric (Bugnet et al. 2018) to find outliers on which the pipeline is then re-run to search for modes in the frequency range predicted by FliPer.

We find 3847 stars that are predicted as detections by A2Z but not by the deep learning classifier. From these, we ignore the 2546 stars with $\nu_{\rm max} \lesssim 3\mu$Hz and 1210 stars with $\nu_{\rm max} > 283\mu$Hz because these are not detectable by our current version of the deep learning classifier by construction. By visual inspection of the power spectra of the remaining 91 stars, we confirm that the deep learning classifier missed 91 genuine oscillators (false negatives). These false negatives are typically stars with a power spectrum showing a superposition of solar-like oscillations with signatures from other forms of variability that show multiple sharp peaks, such as binarity or classical pulsators.

We also find 1195 stars predicted as detections by the deep learning classifier but not by A2Z. Again, we visually inspect their power spectra and find 135 false positives by the deep learning classifier. Thus, the remaining 1060 stars are genuine oscillating red giants detected by our deep learning classifier but not by A2Z.

After adding false negatives and removing false positives from our list of detections, we have a total of 21,914 oscillating red giants, each with a measured $\nu_{\rm max}$ from the deep learning regressor. 20,654 of these giants have $T_{\rm eff}$ and R from the B18 catalog and are plotted in Figure 1a. We tabulate these results in Table 1 (available online), and additionally indicate if the star is above the fiducial line (has red giant Gaia-derived radii), or below it (has dwarf/subgiant Gaia-derived radii).

3.3 Estimating the Surface Gravity of Detected Stars

From the measured $\nu_{\rm max}$ of each oscillating giant, we compute its surface gravity, $\log(g)$, using the following scaling relation (Brown et al. 1991; Kjeldsen & Bedding 1995):

$$\frac{g}{g_\odot} \simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max\odot}}\right)\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{1/2}, \quad (2)$$

where we use $\log(g_\odot) = 4.44$ dex, $\nu_{\rm max\odot} = 3050\ \mu$Hz and $T_{\rm eff\odot} =$ 5772K. We adopt $T_{\rm eff}$ and $\sigma_{T_{\rm eff}}$ values from the B18 catalog if available, otherwise we use values from the *Kepler* DR25 stellar properties catalog (Mathur et al. 2017, KSPC). We include our calculated $\log(g)$ and their uncertainties in Table 1.

To validate our estimated $\log(g)$ values, we compare them with those from the KPSC. We only select stars that have measured $\log(g)$ from asteroseismology from the KSPC because such measurements are the most precise. Additionally, we filter out stars that are also in our training set to prevent a biased validation result. As a result, our validation sample comprises 1430 stars with asteroseismic $\log(g)$ values that the deep learning regressor has not 'seen' during training.

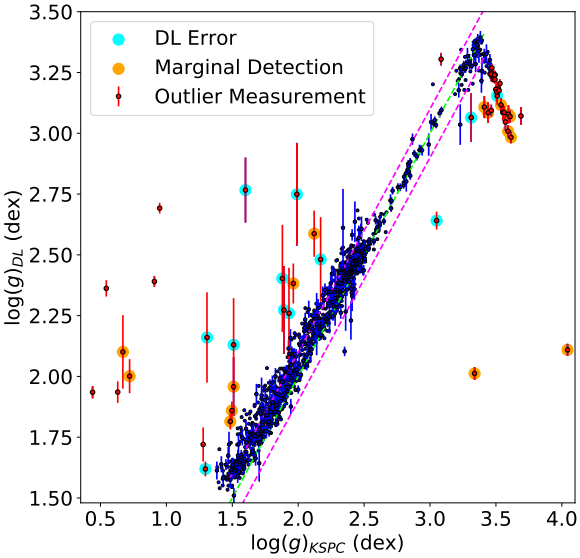

Figure 2. A comparison of estimated surface gravity values from the deep learning regressor, $\log(g)_{\rm DL}$, with asteroseismic surface gravity measurements from the *Kepler* DR25 Stellar Properties Catalog (Mathur et al. 2017), $\log(g)_{\rm KSPC}$, for the 1430 stars in our validation sample. The green dashed line indicates a one-to-one relation, while the magenta dashed lines indicate ±0.1 dex deviations from the green line. For clarity, the errorbars of only every 5th blue point is plotted. Outlier measurements are plotted in red with details in the text.

Table 1. *Kepler* long-cadence stars showing red giant solar-like oscillations. Values in brackets are 1σ uncertainties. The flag '0' indicates stars above the fiducial line in Figure 1, flag '1' indicates stars below the fiducial line, while flag '2' indicates stars that do not have $T_{\rm eff}$ and R from the B18 catalog. The full version of this table is available in a machine-readable form in the online journal, with a portion shown here for guidance regarding its form and content.

KIC	$\nu_{\rm max}(\mu$Hz)	$\log(g)$ (dex)	Flag
5611229	54.44 (1.53)	2.66 (0.01)	0
5611572	58.65 (5.25)	2.71 (0.03)	1
5616489	30.44 (1.30)	2.4 (0.02)	0
5616491	36.83 (7.06)	2.53 (0.09)	1
5616606	35.33 (1.20)	2.47 (0.02)	0
5621709	40.59 (4.14)	2.57 (0.04)	1
5629090	7.01 (0.36)	1.73 (0.02)	0
5613033	80.6 (2.83)	2.87 (0.02)	2
5648159	205.10 (6.34)	3.23 (0.01)	0
...	...	...	...

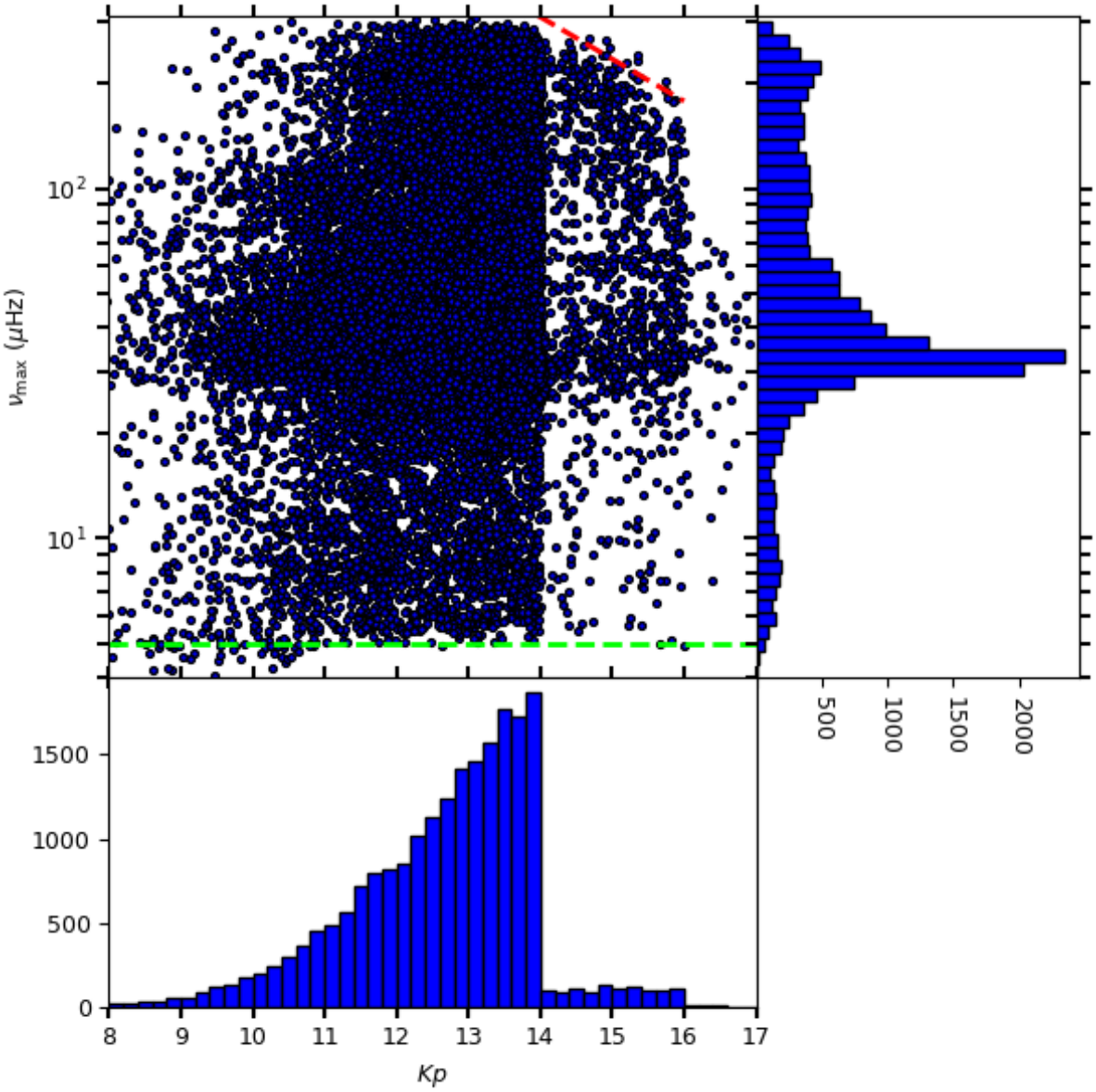

Figure 3. Measured ν_{max} vs. Kp for 21,005 positive detections above the fiducial line in Figure 1. The distributions of Kp and ν_{max} are shown as histograms in the bottom and right panels, respectively. The dashed red line indicates an empirical detection limit $\nu_{max-lim} < 1.5 \times 10^4 \cdot 2^{-0.4 \cdot Kp}$ for $Kp \leq 16$, above which the signal-to-noise is too low for the detection of solar-like oscillations. A large discontinuity in the detection limit occurs at $Kp = 16$, such that only stars with $\nu_{max} \lesssim 70\mu$Hz is detected at fainter magnitudes (see text). The dashed green line at $\nu_{max} = 5\mu$Hz indicates the minimum ν_{max} of the training set used to train the deep learning classifier and regressor, which imposes a limitation on the range of our measurements.

The comparison of our $\log(g)$ values, $\log(g)_{DL}$, with those from the validation sample, $\log(g)_{KSPC}$, is shown in Figure 2. In general, find that 88% and 96% of $\log(g)_{DL}$ lie within ± 0.05 dex and ± 0.1 dex of $\log(g)_{KSPC}$, respectively. For 2.0 dex $\leq \log(g)_{DL} \leq$ 2.6 dex, we find $\log(g)_{DL}$ to be overestimated by $\sim$ 0.04 dex, with $\sim$75% of stars within 2σ uncertainties of the differences between $\log(g)_{DL}$ and $\log(g)_{KSPC}$. For $\log(g)_{DL} < 2.0$ dex, this overestimation increases to $\sim$ 0.08 dex, with $\sim$50% of stars within 2σ uncertainties of the differences in $\log(g)$ values.

We highlight $\log(g)_{DL}$ values that deviate from $\log(g)_{KSPC}$ by more than 0.2 dex in red as shown in Figure 2. For $\log(g)_{DL} \lesssim 2.8$ dex, we find 25 stars that have a large scatter from the one-to-one relation (green line). We visually inspect their power spectra and find that 10 are caused by inaccurate ν_{max} measurements from the deep learning regressor (highlighted in cyan), while 9 stars have very low signal-to-noise levels and are considered marginal detections (highlighted in orange). For the remaining 6 stars, we find that our measured ν_{max} are consistent with the observed frequency range of the oscillation power excess within 2σ.

For $\log(g)_{DL} > 2.8$ dex, we observe a downward 'trail' of 30 stars with large $\log(g)$ deviations. This feature appears to approximate a reflection of the one-to-one relation around $\log(g)_{KSPC} \simeq 3.5$ dex, indicating that the detected oscillation power excess for these stars are potentially aliased counterparts of oscillations with frequencies beyond the Nyquist frequency of $\sim$283μHz (Yu et al. 2016). This results in a smaller $\log(g)_{DL}$ compared to $\log(g)_{KSPC}$.

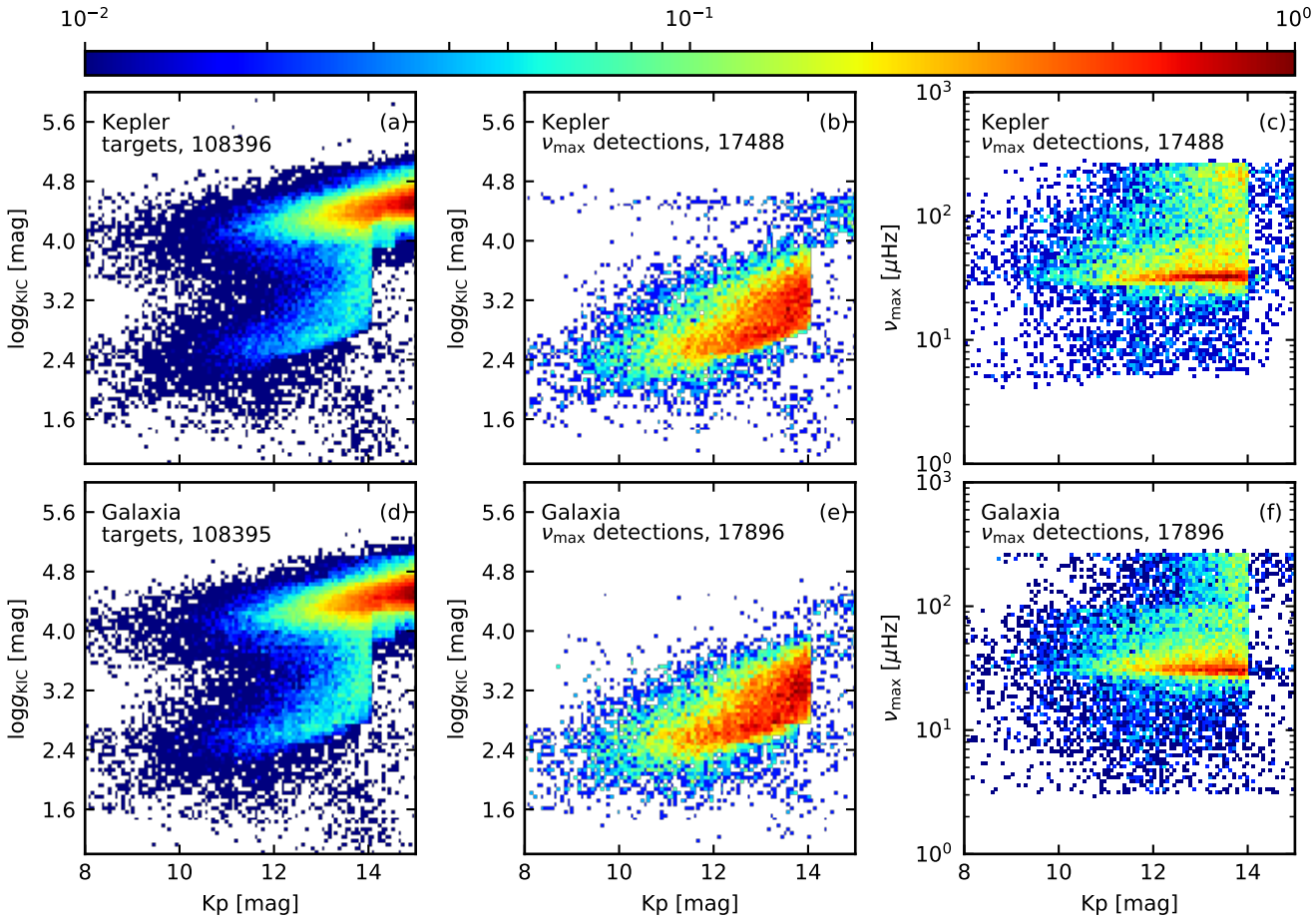

Figure 4. Comparison of the number of observed oscillating *Kepler* giants from this study with predictions from Galaxia. Panels in the top row show results from *Kepler* while panels in the bottom row show results for stars simulated with Galaxia that are sampled to match the distribution of *Kepler* targets in $(\log g_{\rm KIC}, Kp)$ space. Within each panel, the colour represents target density normalized by its maximum density value, while the listed number indicates the number of stars. (a), (d): Distribution of all *Kepler* targets in $(\log g_{\rm KIC}, Kp)$ space. (b), (e): Distribution of targets with detected solar-like oscillations with $3\mu{\rm Hz} < \nu_{\rm max} < 270\mu{\rm Hz}$ in $(\log g_{\rm KIC}, Kp)$ space. (c), (f): Same as panels (b) and (e), but in $(\nu_{\rm max}, Kp)$ space.

Upon visual inspection of their power spectra, we find that 7 stars have low signal-to-noise levels, which we flag as marginal detections. We also find 2 stars with inaccurate deep learning regressor $\nu_{\rm max}$ measurements. The remaining 21 stars have high signal-to-noise levels and $\nu_{\rm max}$ values consistent with the frequency range of the oscillation power excess in their power spectra.

We find that $\log(g)_{\rm DL}$ values based on inaccurately measured $\nu_{\rm max}$ generally have $\sigma_{\log(g),\rm DL} > 0.05$ dex, while ~96% of our non-outlier measurements (blue) have $\sigma_{\log(g),\rm DL} \leq 0.05$ dex. Hence, selecting stars with $\sigma_{\log(g),\rm DL}$ below this threshold will generally provide reliable surface gravity estimates.

3.4 Analysis of Oscillating Giants with Red Giant Gaia-Derived Radii

3.4.1 Distributions of $\nu_{\rm max}$ and Kepler magnitude

A comparison between the stars above the line in Figures 1a and 1b shows that our detections lack stars with $R \gtrsim 60 R_\odot$. This is due to the inherent limitation of our method that is caused by our training set having no stars with $\nu_{\rm max} < 5\mu$Hz. Despite this, we note that our method can still detect stars down to $\nu_{\rm max} \simeq 4\mu$Hz, which is seen in the $\nu_{\rm max}$ vs. *Kepler* magnitude (Kp) diagram in Figure 3.

In Figure 3, we also see a decrease of the number of stars with $\nu_{\rm max} \gtrsim 230\mu$Hz with fainter magnitudes for $Kp \geq 14$. This is due to the lower signal-to-noise levels with fainter magnitudes and higher $\nu_{\rm max}$, which limits the detectability of solar-like oscillations. This can be described by an empirical detection limit as shown in Figure 3. The limit appears to have a discontinuity at $Kp \simeq 16$, which is likely caused by stars at $Kp \gtrsim 16$ originating from a different stellar sample that prioritizes the observation of cool dwarfs (e.g. Batalha et al. 2010). Hence, we infer that this discontinuity is from the selection effect of the sample rather than a detection bias/limit of our method.

To make sure that the slope of the detection limit is not influenced by potentially spurious $\nu_{\rm max}$ measurements for low-luminosity stars, we investigate ~ 20 stars that lie just above the fiducial line in Figure 1a but have a measured $\nu_{\rm max} \lesssim 200\mu$Hz. We find that such stars are not predominantly faint and thus we rule out inaccuracies with measured $\nu_{\rm max}$ values as a significant factor that influences the slope of the observed detection limit.

Additionally, we note that this limit is consistently above the detection limit for the K2 Mission (Howell et al. 2014), which

varies as $2.6 \times 10^6 \cdot 2^{-Kp}$ (Stello et al. 2017). This is expected because K2 data generally have lower signal-to-noise levels compared to that from *Kepler*.

3.4.2 *Detection completeness*

We estimate the detection completeness of the detected oscillating giants in our study by comparing with predictions from theoretical models of the Milky Way. For this purpose, we use Galaxia (Sharma et al. 2011) to simulate a mock *Kepler* survey, from which we resample stars to match the selection function of the *Kepler* targets. Here, we use a Galactic model different from the default model in Galaxia described by Sharma et al. (2011). In particular, our model has a thick disc with mean $\log Z/Z_\odot$ of -0.18 to match the mean $\log Z/Z_\odot$ of stars between 1 kpc $< |z| <$ 2 kpc and 5 kpc $< R <$ 7 kpc in the GALAH survey (Buder et al. 2018).

Because stars in the *Kepler* mission were mainly selected based on *Kepler* magnitude *Kp* and photometric surface gravity $\log g_{\rm KIC}$ (Batalha et al. 2010; Sharma et al. 2016), we assume the selection to be a function of Kp and $\log g_{\rm KIC}$. To simulate this distribution of stars using Galaxia, we bin the full list of *Kepler* targets into 28 x 48 bins in $(Kp, \log g_{\rm kic})$ space as shown in Figure 4a. Next, we resample stars from Galaxia to match the number of *Kepler* targets as shown in Figure 4d. Details of estimating $\log g_{\rm KIC}$ for Galaxia stars are described in Sharma et al. (2016). In short, the synthetic photometry in the g band is re-calibrated to match the *Kepler* pass bands, where the *Kepler* Input Catalog Bayesian scheme (Brown et al. 2011) is then used to estimate the stellar parameters.

Next, we use the Chaplin et al. (2011) method to estimate the probability of detecting $\nu_{\rm max}$ for a light curve spanning 4 years, $p_{\rm detect}$. We then compare the number of detected oscillating giants from *Kepler* (Figure 4b) with the number of Galaxia stars having $p_{\rm detct} > 0.95$ (Figure 4e). The $\nu_{\rm max} - Kp$ distribution for both observed and synthetic oscillating giants is shown in Figures 4c and 4f, respectively. From these results, we find that the distribution of *Kepler* stars in $(Kp, \log g_{\rm kic})$ and $(\nu_{\rm max}, Kp)$ space matches well with predictions from Galaxia. Notably, a majority of *Kepler* targets that show red giant oscillations in Figure 4b have $\log g_{\rm KIC} < 4$, which is in agreement with the distribution shown by Galaxia in Figure 4e. However, in Figure 4f we find that Galaxia predicts a sharp difference in the number of stars just below and above $\nu_{\rm max} \simeq 100\mu$Hz, which is not observed for the *Kepler* targets in Figure 4c. This is because the predicted population of secondary clump stars in Galaxia appear as an overdensity with a sharp maximum $\nu_{\rm max}$ cutoff. As the comparison shows, this cutoff is potentially too sharp to describe the population of observed secondary clump stars.

We find the ratio of observed to predicted stars to be 0.977 and 0.992 for stars with 3μHz $< \nu_{\rm max} < 270\mu$Hz and 10μHz $< \nu_{\rm max} < 270\mu$Hz, respectively. Hence, under the assumption that the Galaxia model provides a good representation of the true population, our sample of detected oscillating giants seems to be nearly complete except for stars with 3μHz $< \nu_{\rm max} < 10\mu$Hz, which is expected due to the current limitation of our deep learning methods as discussed in Section 3.4.1.

4 DWARF/SUBGIANT STARS WITH DETECTED GIANT OSCILLATIONS

In this Section, we address the 909 main sequence/subgiant stars in Section 3.2 that show red giant oscillations. We identify these stars as blended targets, with the following Sections detailing the process.

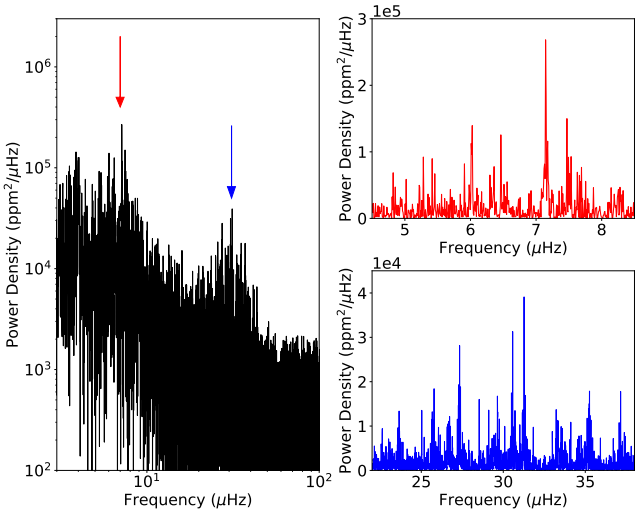

Figure 5. Power spectrum of a blended target, KIC 5824221. The two power excesses showing solar-like oscillations denoted by the arrows (left panel) are determined to be caused light pollution from nearby red giants KIC 5824237 (red) and KIC 5824232 (blue), respectively, with their corresponding oscillation modes in the spectrum shown in the right. The identification of the polluting stars is shown in Figure A1.

4.1 Blended Targets

As a consequence of using a larger photometric aperture on the *Kepler* targets, there is a greater likelihood of light pollution, where flux from a nearby star is captured by the aperture as well. Although this is mitigated when defining the extent of our aperture as described in Section 2.1, flux from a nearby star may still be captured. As a result, the oscillation signal from this polluting star can be superimposed in the power spectrum of the original target as shown in Figure 5.

To confirm that the 909 dwarf/subgiant stars showing giant oscillations are indeed blended targets, we have to identify the polluting star for each target by searching for other nearby stars within its vicinity in the sky. We do this by examining their *Kepler* target pixel files (TPFs). TPFs show a 'postage stamp' image of pixels around *Kepler* targets, with one image at every cadence ($\Delta T \simeq$ 30min.) and each pixel spanning 4 arcseconds across the sky. These files are available to download from the Mikulski Archive for Space Telescopes (MAST)[2], which we retrieve and access using the `lightkurve`[3] Python package (Vinícius et al. 2018).

A single pixel image from the TPF provides a 'flux map' (Figure 6, right), which describes light intensities in regions surrounding a *Kepler* target for a given observation timestamp. Since the full set of pixel images from the TPF provides a light curve for each individual pixel, we create power spectra for each pixel. First, we prepare the light curves by passing them through a high-pass boxcar filter of 20-day length and perform 3σ clipping to remove outliers before taking its Fourier transform, similar to the approach

[2] http://archive.stsci.edu/kepler/
[3] https://lightkurve.keplerscience.org/

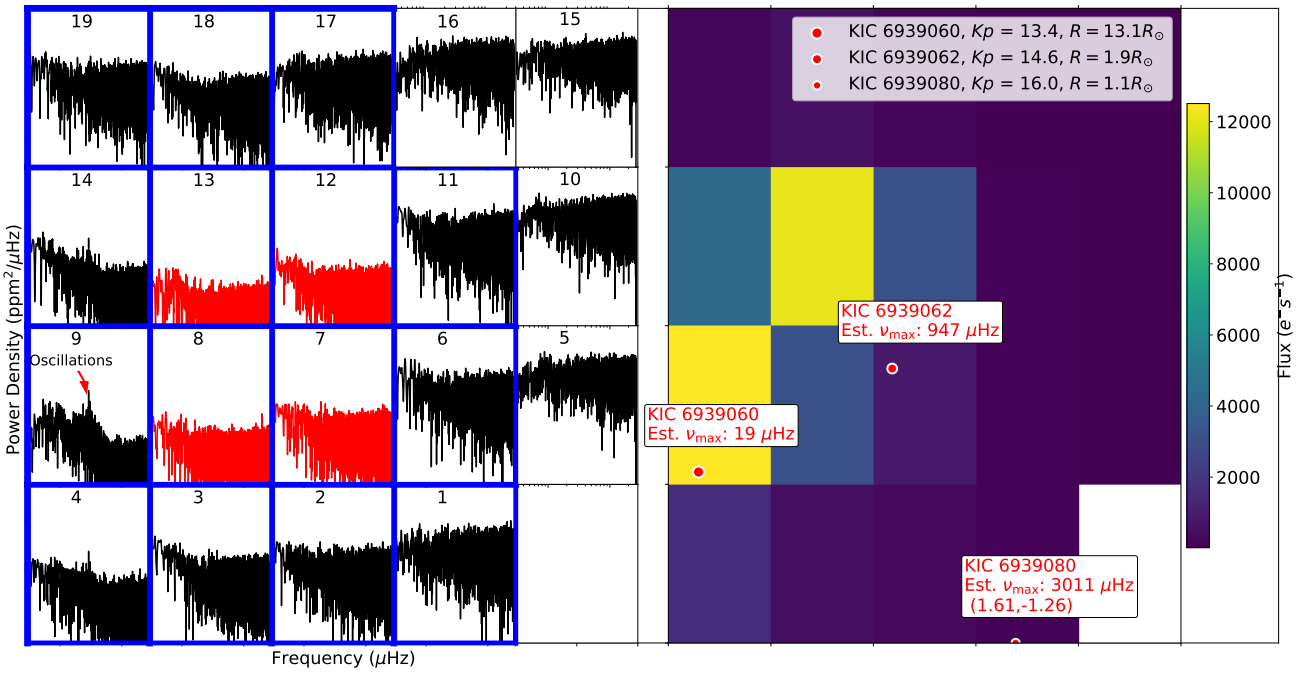

Figure 6. Target Pixel File (TPF) for Quarter 4 of the blended target KIC 6939062, with its power spectrum map (left) and flux map (right). A star map with the positions of the *Kepler* target and nearby stars are overlaid on the flux map, with plot point sizes proportional to the negative logarithm of the *Kepler* magnitude, Kp. The pixel aperture mask used in this study is highlighted with blue borders, while the mask used by the *Kepler* Science Processing Pipeline is shown by red plots in the power spectrum map. KIC 6939060 is identified as the polluting star because it has a $\nu_{\rm max,sc} \simeq 19\mu$Hz using Equation 3, which matches the frequency of the oscillation power excess that can be seen in pixels 8, 9, and 14 in the power spectrum map, indicated by the arrow in the left panel. These pixels also correspond to a high light intensity in the flux map, which are indicated by bright colours (right panel). Another star, KIC 6939080, is located beyond the bounds of the pixel image but is still within our search boundaries for *Kepler* targets (5 pixels beyond TPF) and hence is still shown at the border of the star map.

by Colman et al. (2017). This effectively produces a 'power spectrum map' (Figure 6, left) to complement the flux map. To further aid the identification of the polluting star, we overlay a sky map on the flux map, where we plot and compare the celestial coordinates of the main *Kepler* target with that of other nearby known stars from various catalogs.

Using these three maps, we can identify the pixels that have high incident fluxes and show oscillations, which we then supplement with spatial information from the star map. Such an approach will greatly help to narrow down potential candidates for the polluting star. Because the orientation of the telescope during the mission changes by 90° every quarter, we examine each quarter of pixel data separately. In the following, we describe how we search and identify the polluting stars from different survey catalogs, in decreasing order of priority.

4.2 Identification of Polluting Stars

4.2.1 Kepler targets

We first attempt to identify the polluting stars with targets in the KSPC. In the power spectrum map of each suspected blended target, we initially identify which pixels contain the polluting red giant oscillation signal and determine if these pixels correspond to high flux values in the flux map. This allows us to approximately locate the position of the polluting star in the sky with respect to the blended target.

Next, we plot all other *Kepler* targets that can be seen within the TPF on the sky map. In other words, we find all other targets with celestial coordinates within an area of sky centered around the blended target, with the area boundary determined by the dimensions of the TPF pixel image. Because light from a star can fall onto several different pixels due to the point-spread function of the *Kepler* telescope (Bryson et al. 2010), in some scenarios the polluting star may be located outside the TPF bounds. Thus, we extend our search boundary to also include stars that are 5 pixels (20 arcseconds) beyond each boundary of the TPF.

For every star we locate in the vicinity of a blended target, we calculate $\nu_{\rm max,sc}$, which is an estimate the star's frequency at maximum power using the following scaling relation (Brown et al. 1991; Kjeldsen & Bedding 1995):

$$\frac{\nu_{\rm max,sc}}{\nu_{\rm max\odot}} \simeq \left(\frac{M}{M_\odot}\right)\left(\frac{R}{R_\odot}\right)^{-2}\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{-1/2}, \quad (3)$$

where M, R, and $T_{\rm eff}$ are the stellar mass, radius, and effective temperature, respectively, with $\nu_{\rm max\odot} = 3050\ \mu$Hz and $T_{\rm eff\odot} = 5772$K. For our estimates, we adopt $T_{\rm eff}$ and R values from B18. For all our estimates, we use a constant value of $M = 1.2M_\odot$, which is approximately the median stellar mass of *Kepler* red giants and is a good estimate to within a factor of two for the vast majority of the giants (see Yu et al. 2018, their Figure 5a). We select the polluting star by identifying a neighbouring star that satisfies the following requirements:

(i) The potential candidate is located at a position on the sky map that

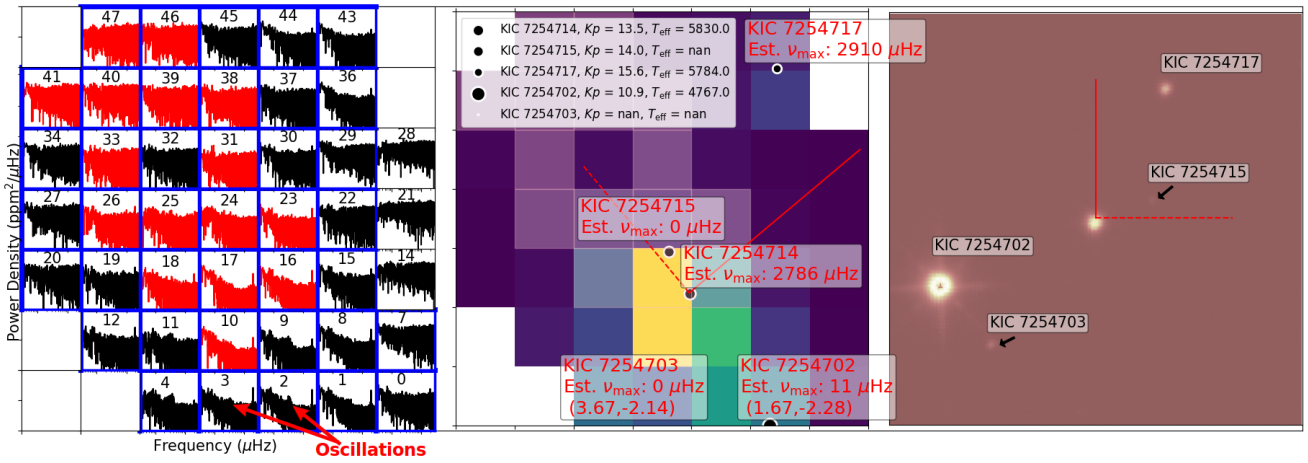

Figure 7. Target Pixel File for Quarter 4 of the blended target KIC 7254714, with its power spectrum map (left) and flux/star maps(center). Additionally, a 45 x 45 arcsecond *Kepler* field image cutout from the UKIRT WFCAM survey centered on the blended target is displayed on the right. To indicate scale, compass arms are drawn on both the star map and the field image and both are 15 arcseconds across. The solid line points north, while the dashed line points east. $\nu_{\rm max,sc} = 0$ is assigned to stars without $T_{\rm eff}$ or R values. The polluting star is identified as KIC 7254702, which is a very bright target such that its oscillation signal (as indicated in pixels 2 and 3 in the power spectrum map) can be detected across many pixels. The use of the UKIRT image allows us to visually identify bright polluting stars beyond our search range for non-*Kepler* targets (2 pixels beyond the TPF).

coincides with the location of the polluting star as inferred from the flux and power spectrum maps.

(ii) $\nu_{\rm max,sc}$ of the potential candidate (Equation 3) matches the $\nu_{\rm max}$ as measured by the deep learning regressor within 2σ. This can also be confirmed by visually identifying $\nu_{\rm max}$ from the power spectrum maps.

An example of this selection process is shown in Figure 6.

4.2.2 *Kepler Input Catalog (KIC) stars*

If there are no *Kepler* targets that are suitable candidates for the polluting star, we next search for candidates within the KIC (Brown et al. 2011). Such stars typically have stellar parameters in the KIC but do not have light curves as they were not targeted by *Kepler*. We perform a similar analysis as described in Section 4.2.1, except that here we restrict our search boundary to 2 pixels (8 arcseconds) beyond each TPF boundary. This is because searching for a polluting candidate star beyond the TPF becomes increasingly difficult with larger search boundaries due to the high density of stars now being included from the KIC.

This time, we adopt R and $T_{\rm eff}$ values directly from the KIC. Because not all entries in the KIC have these values measured, three scenarios are possible:

(i) The polluting candidate has both $T_{\rm eff}$ and R listed, allowing a determination of $\nu_{\rm max,sc}$ using Equation 3, hence allowing us to match $\nu_{\rm max,sc}$ with $\nu_{\rm max}$ values from the deep learning regressor and $\nu_{\rm max}$ observed from the power spectrum maps.

(ii) The polluting candidate only has $T_{\rm eff}$ measured. We identify stars with $3000{\rm K} \leq T_{\rm eff} \leq 6000{\rm K}$ as more likely candidates because they have the expected $T_{\rm eff}$ for a red giant. The availability of a $T_{\rm eff}$ value also indicates the presence of multi-band photometry for this star, hence it is typically not too faint.

(iii) The polluting candidate has neither $T_{\rm eff}$ nor R listed. In this scenario, we are only able to rely on the spatial requirement (i) in Section 4.2.1.

For scenarios where identifying the polluting star is difficult, which is however uncommon, we also compare the flux, power spectra, and sky maps from each TPF with higher resolution images of the same area of sky using *Kepler* field image cutouts from the UKIRT WFCAM (UK Infrared Telescope Wide Field Camera) survey (Lawrence et al. 2007). An example of this is shown in Figure 7.

4.2.3 *Gaia targets within the Kepler field*

For polluting stars that are not identified as *Kepler* targets or as entries in the KIC, we next identify them using data from the Gaia mission (Prusti et al. 2016), where we search across 10.3 million Gaia targets from the Gaia DR2 catalog (Brown et al. 2018) that lie within 8° from the center of the *Kepler* field. We use a similar approach to that in Section 4.2.2, except that we adopt $T_{\rm eff}$ and R values from the Gaia DR2 catalog.

4.3 Analysis of Polluting Stars

From the 909 dwarf/subgiant stars with giant oscillations, we identify that 293 have polluting stars that are *Kepler* targets (hence with previously known light curves on MAST), 587 are non-*Kepler* targets in the KIC, 14 are only identified as Gaia targets, with the remaining 15 unidentified.

First, we verify that our identified polluting stars are indeed red giants. In Figure 8a, we plot the stellar parameters of the 293 polluting stars that are known *Kepler* targets. As expected, they all appear above the fiducial line and hence are red giants. In Figure 8b, we do the same for the 587 polluting non-*Kepler* targets that are in the KIC, except that we plot the $\log g$ and $T_{\rm eff}$ values from the KIC. We find that most have $5\mu{\rm Hz} \lesssim \nu_{\rm max,sc} \lesssim 283\mu{\rm Hz}$. However, there is a significant number of stars with $\nu_{\rm max,sc} > 283\mu{\rm Hz}$. This can be explained by the fact that $\log g$ values from the KIC are typically overestimated by about 0.5 dex (Brown et al. 2011). Alternatively,

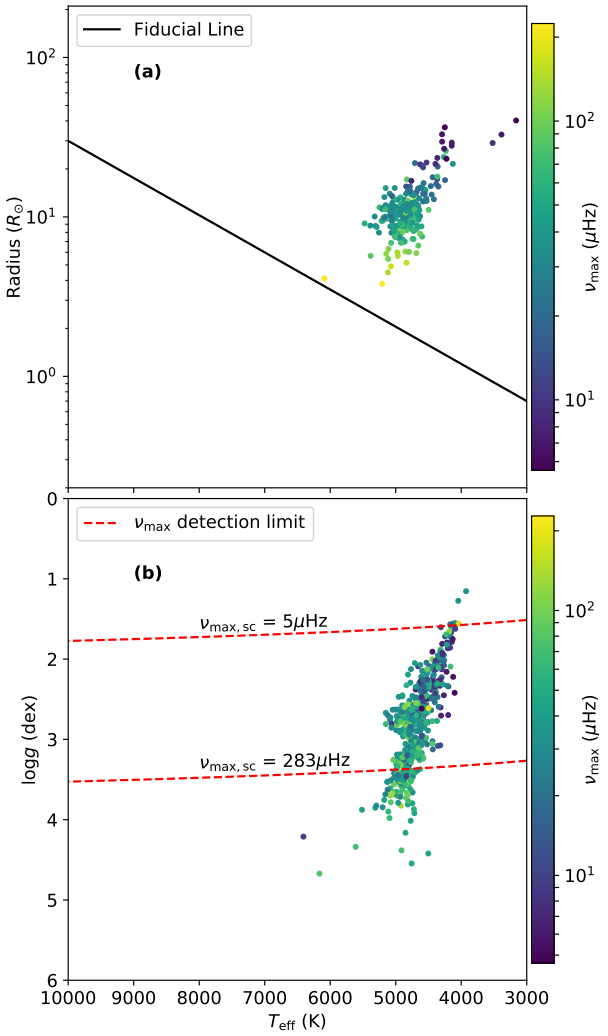

Figure 8. (a) 293 polluting stars that are known *Kepler* targets, with R and $T_{\rm eff}$ values adopted from the B18 catalog. The drawn fiducial line is the same as in Figure 1. (b) 562 polluting stars that are non-*Kepler* targets in the KIC, with $\log g$ and $T_{\rm eff}$ values adopted directly from the KIC. 25 stars do not have $\log g$ values from the KIC and hence are not plotted. The dashed red lines are lines of constant $\nu_{\rm max,sc}$ and is a rough guide to indicate the detection limit of our method.

the detected oscillations from these stars may be aliased counterparts of super-Nyquist modes as discussed in Section 3.3. We also find a few stars with $\log g > 4$ dex in Figure 8b, whereupon visual inspection we find that these stars have marginal detections, such that their identification as oscillating red giants are uncertain.

To investigate the properties of these new red giants, we compare their Kp and $\nu_{\rm max}$ distributions with that of the positive detections above the fiducial line. The result of this is shown in Figure 9. For $\nu_{\rm max} \gtrsim 100\mu$Hz, the number of new red giants (red) decreases more rapidly with $\nu_{\rm max}$ compared to the detections above the fiducial line (blue), which is expected because each photometric aperture that measures the flux from these new red giants is not centered around them. As a result, these new giants are measured 'indirectly', causing them to have lower signal-to-noise levels as compared to *Kepler* targets with apertures centered around them.

As a whole, the Kp distribution of these new giants differ significantly from the distribution for targeted giants, which shows a

Table 2. A list of blended targets, each assigned with the identity of its polluting star. Flag '0' indicates that the polluting star is a *Kepler* target, flag '1' indicates that the polluting star is a non-*Kepler* target in the KIC, flag '2' indicates that the polluting star is only identified in the Gaia DR2 catalog (for which the 19-digit Gaia source ID is listed instead), while flag '3' indicates an unidentified polluting star. The full version of this table is available in a machine-readable format in the online journal, with a portion shown here for guidance regarding its form and content.

Blended Target ID	Polluting Star ID	Flag
7102071	7102068	0
7207089	7124606	1
7211526*	7211529	1
7222444	7222454**	1
7797928*	-	3
7811846	7811847	0
7812628	7812622	1
7881258	7881261	0
7984047	-	3
8016196	2105691730722492544**	2
...	...	...

* Marginal detection
** No $T_{\rm eff}$ or R available for this star

cut-off at $Kp \simeq 14$ from their selection function (Batalha et al. 2010; Sharma et al. 2016). Specifically, it can be seen that a significant fraction of the new giants are fainter than the majority of the targeted giants. They therefore potentially represent a population of distant stars similar to the giants residing in the halo of the Galaxy as identified by Mathur et al. (2016). We find that our sample of new red giants has no overlap with the sample by Mathur et al. (2016), hence they will prove to be a very valuable sample for galactic archaeology.

We also investigate the difference in Kp distributions between the new giants (red) and the *Kepler* targets that they pollute (green) as shown in Figure 10a. We see that these two distributions are significantly different from one another, which we can better quantify by examining the distribution of their Kp differences as shown in Figure 10b. On average, the polluting star is of comparable or brighter magnitude compared to the blended target, which is an expected result. Interestingly, there are also polluting stars that are 2~4 magnitudes fainter than the blended target that are still detected. This can be informative about the Kp detection limit for neighbouring stars using our custom aperture. While further investigations into these distributions are beyond the scope of this paper, the sample of blended targets and their polluting stars will nonetheless be of significant interest towards modeling populations of giants in the *Kepler* field.

Besides polluting stars with entries in the KIC, we find 14 polluting stars that are only identified as Gaia targets. In most of these scenarios, the flux and power spectrum maps indicate that the polluting star should be located within an adjacent pixel to the blended target, where we indeed find a nearby Gaia target within such an area. However, all Gaia targets that we identify have no measured $T_{\rm eff}$ from the Gaia DR2 catalog, such that we can only determine polluting star candidates from their location on the star map relative to the blended target. This requirement alone is not strong evidence towards the identity of the actual polluting star. Hence, it is possible that the polluting stars for these targets may instead be from unresolved binary members or chance alignments from other background stars.

Finally, we are unable to identify the polluting stars for 15 dwarfs/subgiants that show giant oscillations. Eleven of these are

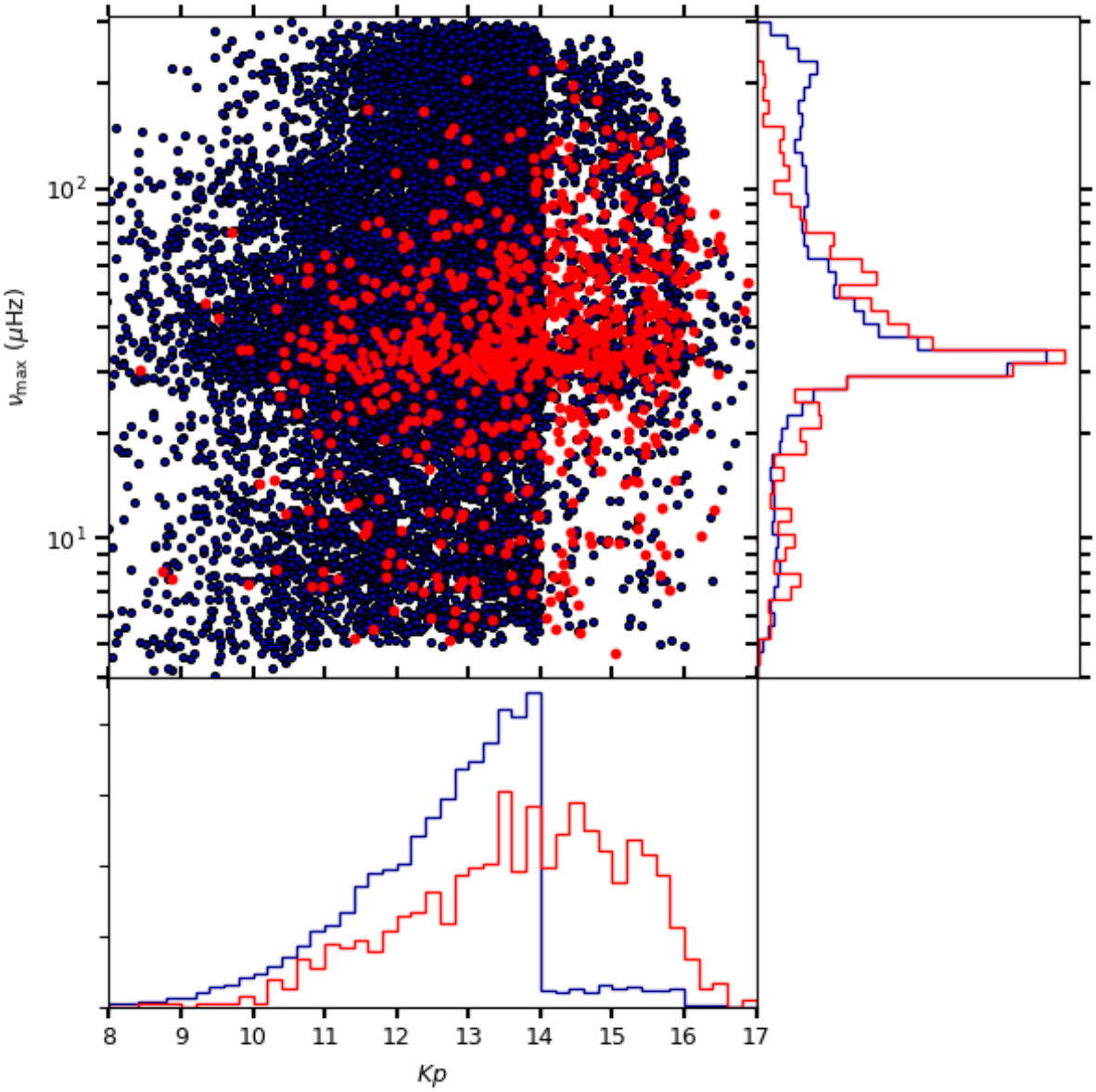

Figure 9. Comparison of ν_{max} vs. Kp for 21,005 giants above the fiducial line in Figure 1a (blue) and for the 880 polluting giants (red). The distributions of the red and blue dots are shown as histograms in Kp and ν_{max}. 29 polluting giants do not have measured *Kepler* magnitudes, and hence are not shown.

marginal detections where most have been observed for only a few quarters. We infer that the remaining 4 may have unresolved polluting stars, because their TPFs indicate polluting stars that are close to the blended target, but no such star can be identified from either the KIC or the Gaia DR2. In Table 2, we tabulate the 909 blended targets along with their corresponding identified polluting star. The ν_{max} for each polluting star can be found by matching the blended target ID in Table 2 with the KIC ID in Table 1.

5 DISCUSSION AND CONCLUSION

From all ~197,000 *Kepler* stars observed in long-cadence, we detected a total of 21,914 stars showing solar-like oscillations using a deep learning classifier, yielding the current largest list of stars showing solar-like oscillations providing many more targets of interest for various asteroseismic and Galactic archaeology analyses. We also predict ν_{max} for each detection, which will provide useful prior values for more precise measurements using asteroseismic model-fitting pipelines. In addition, our ν_{max} values provide $\log(g)$ estimates, for which ~88% are good to within 0.05 dex, which is still superior to typical spectroscopic $\log(g)$ determinations. Because our method is limited to stars with $\nu_{max} \gtrsim 4 - 5\mu$Hz, we can currently only identify oscillating red giants with $R \lesssim 40 R_{\odot}$. We will explore extending our methods to lower frequencies with the greater availability of highly luminous giants that are analysed seismically. From a machine learning perspective, the increased number of identified oscillating *Kepler* red giants will be of great importance to increasing the robustness of our deep learning methods by bolstering the quantity of training examples in preparation for data from the Transiting Exoplanet Survey Satellite (TESS) (Ricker et al. 2014). The ability of our current method is limited by the

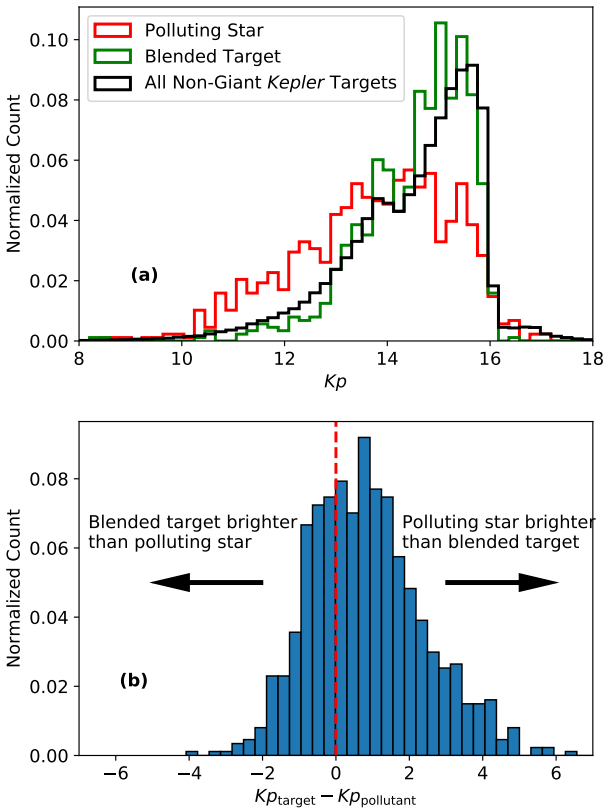

Figure 10. (a) Comparison of Kp distributions between blended targets (green), which are the stars below the fiducial line in Figure 1a, and their polluting stars (red), which are red giants neighbouring the blended targets that cause red giant solar-like oscillations to be seen in the power spectrum of the blended targets. The Kp distribution for non-giant *Kepler* targets (black) is also shown for comparison. (b) Kp differences between a blended target ($Kp_{\rm target}$) and its identified polluting star ($Kp_{\rm pollutant}$). The dashed red line indicates the zero point, where both stars' magnitudes are equal.

types of stars that are 'seen' during training, hence the inclusion of many more stars during training will greatly help the classifier distinguish solar-like oscillations across a wider range of stellar variability.

Although our deep learning methods can efficiently make predictions for a wide range of power-frequency profiles of an oscillating red giant, it is possible that the classifier may label an oscillating red giant as a non-detection. This can happen if the star shows many other sharp peaks in its power spectrum. These are typically signals from a detached binary or from a classical pulsator, which make the power spectrum look more complicated. While we have corrected many such cases by visual inspection of selected samples of suspicious cases, there can still be a small fraction of such stars that were missed by our analysis. However, such stars that do get classified correctly as positive detections is expected to have large $\nu_{\rm max}$ uncertainties because the regressor has not 'seen' such complicated power spectra before. Thus, it will be useful and interesting to validate our results with measurements from more classical seismic pipelines, as well as with the FliPer metric (Bugnet et al. 2018), or statistical tests on the power spectrum (Bell et al. 2018).

From our list of 21,914 detections, we identified 21,005 *Kepler* targets with Gaia-derived radii representative of red giants. This number is within estimates of the total number of red giants (~21,000) as reported by B18 even though we do not account for giants with $R \gtrsim 40R_{\odot}$. This is because we also detect giant oscillations in stars with $R \lesssim 10R_{\odot}$ that were classified as subgiants by B18 (see Figure 1). Additionally, we find 1702 stars that are classified as red giants by B18 but are not in our list of detections, indicating that these stars potentially have suppressed oscillations caused by physical interactions such as rapid rotation in close binary systems (Gaulme et al. 2014; Tayar et al. 2015).

By observing the $\nu_{\rm max}$-Kp distribution of the 21,005 detected giants, we derived an empirical estimate for the detection limit for solar-like oscillations. We note that the power spectra in this study have higher white noise levels compared to the photometry by the *Kepler* Science Processing Pipeline because we use larger photometric apertures, making it possible that we miss out a small fraction of high $\nu_{\rm max}$ stars at $Kp > 14$. Hence, our derived detection limit should be taken as a lower bound to the detectability of solar-like oscillations.

We also compare our detections with predictions from a synthetic model of the Milky Way by Galaxia, where we find a good agreement of the distributions of the giants in $(Kp, \log g_{\rm kic})$ and $(\nu_{\rm max}, Kp)$ space. Furthermore, a comparison of the number of observed and simulated oscillating giants indicate that our detections are ~99% complete for stars with $10\mu{\rm Hz} < \nu_{\rm max} < 270\mu{\rm Hz}$. This level of completeness decreases to ~97% for stars with $3\mu{\rm Hz} < \nu_{\rm max} < 270\mu{\rm Hz}$ as a result of the limitation of our deep learning classifier at low $\nu_{\rm max}$.

The 909 serendipitous detections from our results show solar-like oscillations but have Gaia-derived radii representative of dwarf stars. Using the TPFs for such stars, we determined that most were blended targets as a result of using a larger photometric aperture. We identified the polluting stars for most of the 909 blended targets, where 293 are *Kepler* targets, while 587 are non-*Kepler* targets in the KIC. For polluting stars identified in the KIC, we found that their Kp distribution differs significantly from the distribution of *Kepler* red giant targets, with a significant fraction of them having $Kp > 14$. Additionally, by comparing the magnitudes for a blended target with its corresponding polluting star, we found that polluting stars have magnitudes that are on average brighter than the blended targets, yet there are also cases where the blended targets can be brighter than the polluting stars.

Finally, we identified potential polluting stars identified only in the Gaia DR2 catalog for 14 blended targets, while the identification for the 15 remaining blended targets remains uncertain. We suggest that stars that show clear detections among these 15 stars may have unresolved nearby polluting stars. Nevertheless, because a large fraction of the polluting stars for the 909 blended targets were not targeted by *Kepler*, this opens up opportunities for performing asteroseismology on oscillating red giants without previous light curves of their own, which will be valuable additions to both stellar population studies and Galactic archaeology.

ACKNOWLEDGEMENTS

Funding for this Discovery mission is provided by NASA's Science Mission Directorate. We thank the entire *Kepler* team without whom this investigation would not be possible. D.S. is the recipient of an Australian Research Council Future Fellowship (project number FT1400147). R.A.G. acknowledges the support from CNES. S.M acknowledges support from NASA grant NNX15AF13G, NSF grant AST-1411685, and the Ramon y Cajal fellowship number

RYC-2015-17697. I.L.C. acknowledges scholarship support from the University of Sydney. We would like to thank Nicholas Barbara and Timothy Bedding for providing us with a list of variable stars that helped to validate a number of detections in this study. We also thank the group at the University of Sydney for fruitful discussions. Finally, we gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

REFERENCES

Abadi M., et al., 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, https://www.tensorflow.org/
Aguirre V. S., et al., 2018, Monthly Notices of the Royal Astronomical Society
Batalha N. M., et al., 2010, The Astrophysical Journal, 713, L109
Beck P. G., et al., 2011, Nature, 481, 55
Bell K. J., Hekker S., Kuszlewicz J. S., 2018, preprint (arXiv:1809.09135)
Berger T. A., Huber D., Gaidos E., van Saders J. L., 2018, preprint (arXiv:1805.00231)
Borucki W. J., et al., 2010, Science, 327, 977
Brown T. M., Gilliland R. L., Noyes R. W., Ramsey L. W., 1991, The Astrophysical Journal, 368, 599
Brown T. M., Latham D. W., Everett M. E., Esquerdo G. A., 2011, The Astronomical Journal, 142, 112
Brown A. G. A., et al., 2018, Astronomy & Astrophysics, 616, A1
Bryson S. T., et al., 2010, The Astrophysical Journal, 713, L97
Buder S., et al., 2018, Monthly Notices of the Royal Astronomical Society, 478, 4513
Bugnet L., García R. A., Davies G. R., Mathur S., Corsaro E., Hall O. J., Rendle B. M., 2018, preprint (arXiv:1809.05105)
Casagrande L., et al., 2015, Monthly Notices of the Royal Astronomical Society, 455, 987
Chaplin W. J., Miglio A., 2013, Annual Review of Astronomy and Astrophysics, 51, 353
Chaplin W. J., et al., 2011, The Astrophysical Journal, 732, 54
Chetlur S., Woolley C., Vandermersch P., Cohen J., Tran J., Catanzaro B., Shelhamer E., 2014, CoRR, abs/1410.0759
Chollet F., 2015, Keras, https://github.com/fchollet/keras
Colman I. L., et al., 2017, Monthly Notices of the Royal Astronomical Society, 469, 3802
Deheuvels S., et al., 2014, Astronomy & Astrophysics, 564, A27
Fuller J., Cantiello M., Stello D., Garcia R. A., Bildsten L., 2015, Science, 350, 423
García R. A., Stello D., , in Tong V. C. H., Garcia R. A., eds, , Extraterrestrial Seismology. Cambridge University Press, Chapt. Asteroseismology of red giant stars, pp 159–169, doi:10.1017/cbo9781107300668.014
García R. A., et al., 2011, Monthly Notices of the Royal Astronomical Society: Letters, 414, L6
García R. A., et al., 2014a, Astronomy & Astrophysics, 568, A10
García R. A., et al., 2014b, Astronomy & Astrophysics, 572, A34
Gaulme P., Jackiewicz J., Appourchaux T., Mosser B., 2014, The Astrophysical Journal, 785, 5
Hekker S., Christensen-Dalsgaard J., 2017, The Astronomy and Astrophysics Review, 25
Hekker S., et al., 2011, Monthly Notices of the Royal Astronomical Society, 414, 2594
Hon M., Stello D., Zinn J. C., 2018, The Astrophysical Journal, 859, 64
Howell S. B., et al., 2014, Publications of the Astronomical Society of the Pacific, 126, 398
Huber D., et al., 2011, The Astrophysical Journal, 743, 143
Huber D., et al., 2014, The Astrophysical Journal Supplement Series, 211, 2
Jenkins J. M., et al., 2010, The Astrophysical Journal, 713, L87
Kirk B., et al., 2016, The Astronomical Journal, 151, 68
Kjeldsen H., Bedding T. R., 1995, A&A, 293, 87
Lawrence A., et al., 2007, Monthly Notices of the Royal Astronomical Society, 379, 1599
Mathur S., et al., 2010, Astronomy and Astrophysics, 511, A46
Mathur S., García R. A., Huber D., Regulo C., Stello D., Beck P. G., Houmani K., Salabert D., 2016, The Astrophysical Journal, 827, 50
Mathur S., et al., 2017, The Astrophysical Journal Supplement Series, 229, 30
Mosser B., et al., 2012, A&A, 548, A10
Pires S., Mathur S., García R. A., Ballot J., Stello D., Sato K., 2015, Astronomy & Astrophysics, 574, A18
Prusti T., et al., 2016, Astronomy & Astrophysics, 595, A1
Ricker G. R., et al., 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
Sharma S., Bland-Hawthorn J., Johnston K. V., Binney J., 2011, The Astrophysical Journal, 730, 3
Sharma S., Stello D., Bland-Hawthorn J., Huber D., Bedding T. R., 2016, The Astrophysical Journal, 822, 15
Stello D., Chaplin W. J., Basu S., Elsworth Y., Bedding T. R., 2009, Monthly Notices of the Royal Astronomical Society: Letters, 400, L80
Stello D., et al., 2013, ApJ, 765, L41
Stello D., Cantiello M., Fuller J., Garcia R. A., Huber D., 2016a, Publications of the Astronomical Society of Australia, 33
Stello D., Cantiello M., Fuller J., Huber D., García R. A., Bedding T. R., Bildsten L., Aguirre V. S., 2016b, Nature, 529, 364
Stello D., et al., 2017, The Astrophysical Journal, 835, 83
Tayar J., et al., 2015, The Astrophysical Journal, 807, 82
Vinícius Z., Barentsen G., Hedges C., Gully-Santiago M., Cody A. M., 2018, KeplerGO/lightkurve, doi:10.5281/zenodo.1181928
Yu J., Huber D., Bedding T. R., Stello D., Murphy S. J., Xiang M., Bi S., Li T., 2016, Monthly Notices of the Royal Astronomical Society, 463, 1297
Yu J., Huber D., Bedding T. R., Stello D., Hon M., Murphy S. J., Khanna S., 2018, The Astrophysical Journal Supplement Series, 236, 42

APPENDIX A: DOUBLE BLENDED TARGET KIC 5824221

In Figure A1, we demonstrate an example of the identification of the polluting star for the blended target KIC 5824221. The power spectrum for this blended target is shown in Figure 5. The two power excesses arises as a result of pollution from two nearby red giants, which we identify to be KIC 5824237 and KIC 5824228.

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

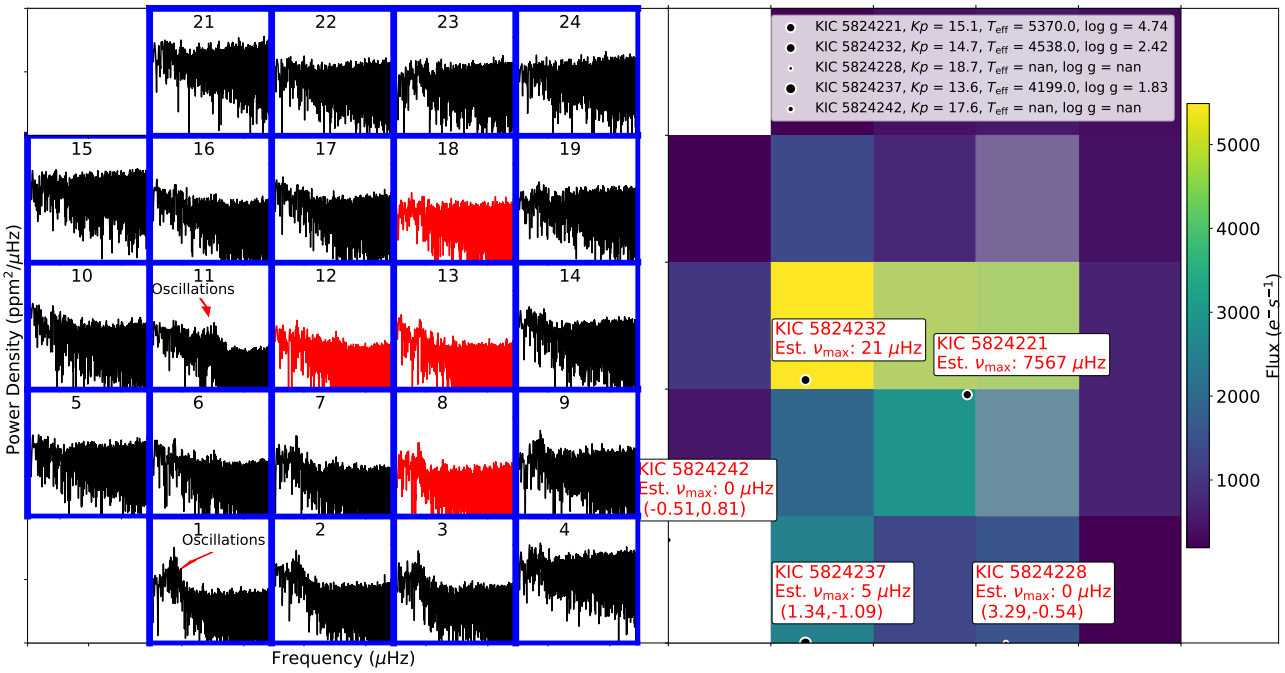

Figure A1. Target Pixel File for Quarter 3 of the double blended target KIC 5824221 with its power spectrum map (left) and its flux/star map (right) . $\nu_{\rm max,sc} = 0$ is assigned to stars without $T_{\rm eff}$ or R values from the B18 catalog. The source of the polluting signals can be seen mainly in pixels 1 and 11, which agrees with the position and $\nu_{\rm max}$ for the stars KIC 5824237 and KIC 5824232, respectively.