Stellar flares from blended and neighbouring stars in *Kepler* short cadence observations

James A. G. Jackman,1,2,3*, Evgenya Shkolnik1, R. O. Parke Loyd1

1School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287
2Dept. of Physics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK
3Centre for Exoplanets and Habitability, University of Warwick, Coventry CV4 7AL, UK

ABSTRACT

We present the results of a search for stellar flares from stars neighbouring the target sources in the *Kepler* short cadence data. These flares have been discarded as contaminants in previous surveys and therefore provide an unexplored resource of flare events, in particular high energy events from faint stars. We have measured M dwarf flare energies up to $1.5 \times 10^{35}$ erg, pushing the limit for flare energies measured using *Kepler* data. We have used our sample to study the flaring activity of wide binaries, finding that the lower mass counterpart in a wide binary flares more often at a given energy. Of the 4430 flares detected in our original search, 298 came from a neighbouring star, a rate of $6.7 \pm 0.4$ per cent for the *Kepler* short cadence lightcurves. We have used our sample to estimate a $5.8 \pm 0.1$ per cent rate of false positive flare events in studies using *TESS* short cadence data.

Key words: stars: flare – stars: low mass

1 INTRODUCTION

Blending of multiple stars within apertures is a known problem in astronomy. Additional light from neighbouring stars will dilute astrophysical signals from a target star of interest, or even introduce signals where there were none before. A noted example of this is in exoplanet transit searches, where a deep signal from a faint eclipsing binary within a photometric aperture is diluted by the brighter target star, resulting in an apparent transit signal. False positive signals such as this, if not fully understood, can impact large scale statistical studies, such as the incidence of stars hosting a transiting planet. Vital tools in the analysis of astrophysical signals are the image centroid and the use of difference images. Measuring the position of the centroid, the centre of light, during the transit can identify a neighbouring star as the true source which can be flagged for removal (e.g. Bryson et al. 2013; Günther et al. 2017).

Another area of astronomy where blending can be an issue is the study of stellar flares. Stellar flares can be observed across a wide range of wavelengths, from the radio, through the infrared and optical up to X-rays (e.g. Jackson et al. 1989; Flaccomio et al. 2018). Believed to be analogous to Solar flares, stellar flares are explosive phenomena which are the result of magnetic reconnection events in the upper atmospheres of stars (e.g. Benz & Güdel 2010). Charged particles are accelerated downwards from the reconnection site in the corona, spiralling along the newly reconnected magnetic field lines until they reach the dense chromosphere. There they are suddenly decelerated, evaporating the surrounding plasma and heating local atmospheric layers, which in turn results in white-light emission (e.g. Antonucci & Dennis 1983; Hawley & Fisher 1992; Fletcher & Hudson 2008).

In recent years it is the white-light flare emission that has been the focus of ardent research. This is due to their detectability in wide-field optical exoplanet and transient surveys, which survey large patches of sky for long durations (e.g. Walkowicz et al. 2011; Davenport 2016; Lin et al. 2019; Schmidt et al. 2019). Such surveys have been used to study white-light flare energies, occurrence rates and their incidence rates. This has been both for individual stars and as a function of spectral type and age (Chang et al. 2015; Ilin et al. 2019; Jackman et al. 2020; Feinstein et al. 2020). The role of stellar flares in exoplanet habitability has also been called into question as well, with the high energy UV emission from flares potentially altering the chemistry and structure of exoplanet atmospheres (e.g. Venot et al. 2016; Chadney et al. 2017). At the same time, stellar flares have been suggested in the literature as a potential requirement for abiogenesis around M stars (e.g. Rimmer et al. 2018). In this scenario, flares provide the near-UV flux ($\approx 200$-300 nm; Ranjan et al. 2017) needed for the formation of amino acids which may lead to abiogenesis. Relative to solar-type stars, the near-UV flux from main sequence M stars is lacking and as such flares may be a requirement to provide needed UV flux. Large scale flare studies (such as those performed with *Kepler* and *TESS*, e.g. Günther et al. 2020) have aimed to understand the role of flares on the habitability of exoplanets around M stars by measuring the energies and flare occurrence rates for samples of M dwarfs and comparing these.
values with the predictions of atmospheric and biological models (e.g. Tilley et al. 2019). However, these properties can be affected by blended flare events (e.g. by atributing M star flares to a nearby brighter star), affecting the accuracy of the measured properties and any comparisons to theoretical models. Consequently, the rate of blended flare events in such surveys requires careful characterisation so that surveys can ensure more accurate measurements of flare energies and occurrence rates.

Previous large scale flare studies using Kepler observations have noted the issue of blending and taken various steps to get around it through a combination of vetting their input catalog and manual validation of flare events. For example, in their study of flares from G stars with Kepler long cadence observations, Shibayama et al. (2013) only studied stars in their sample which did not have another star (as specified from the Kepler input catalog) within 12”, to limit contamination from close neighbours. Flare candidates were also vetted to remove events where the spatial brightness distribution of the target between the flare and quiescence (i.e., the flare comes from a star elsewhere in the aperture). Similar approaches have been used in subsequent studies (e.g. Yang et al. 2017, 2018), although some have also added a step to remove eclipsing binaries from their search. Shibayama et al. (2013) found that approximately 10 per cent of flares in their long cadence sample were due to a neighbouring star in their aperture. Similar values have been reported by other surveys (Gao et al. 2016; Yang et al. 2017; Yang & Liu 2019).

However, by limiting their input samples before searching for flares, these studies did not measure the true rate of blended flare events, instead the rate from stars greater than some distance (typically 12”) from the host star. For studies which cannot spatially resolve stars down to this limit (i.e. those using TESS data) these previously reported values may underestimate the true false positive fractions of flares from neighbouring stars.

Through their vetting, these studies removed thousands of flare events from neighbouring stars (e.g. Yang et al. 2017). These events are a currently unexplored resource and as such could provide an opportunity to measure some of the highest energy flares. Considering the case of a faint star nearby a brighter target star, only the highest energy events from the fainter star will contribute a detectable level of flux to the combined source. Studies using a similar technique of specifically targeting faint stars alone with wide-field surveys and waiting for them to flare into view (e.g. Schmidt et al. 2016; Jackman et al. 2019; Schmidt et al. 2019) have been able to detect some of the highest energy flares from M and L dwarfs, with Schmidt et al. (2019) measuring V band flare energies up to 10^15 erg from M dwarfs observed with ASAS-SN. Consequently, blended flares, specifically if they come from faint sources, can provide a way of studying higher energy flare events than those typically observed in standard large scale studies.

In this work we present the results of a search for blended flares from Kepler short cadence observations. These flares have been missed or discarded by previous flare surveys, making this a trove of new events. We will discuss our method for detrending the Kepler short cadence observations, detecting flare events and then identifying those which are from stars neighbouring the target source. We will detail our method for determining stellar properties and measuring flare amplitudes and energies. We will then discuss how these flares compare to those previously observed with Kepler and other surveys, along with measuring the fraction of flares which were blended in the Kepler short cadence observations.

2 DATA

The primary Kepler mission began in 2009 and ran for nearly four years, observing a 115 deg^2 patch of sky in the northern hemisphere (Borucki et al. 2010). During this time, Kepler observed ~ 115,000 stars (Batalha et al. 2010). The Kepler photometry was measured using two time cadences, a 30 minute long cadence mode and a 1 minute short cadence mode. The long cadence observations were performed for each target almost continuously throughout the duration of the mission. These long baseline lightcurves have been used for a variety of astrophysical studies, including but not limited to asteroseismology, stellar rotation and studies of eclipsing binaries and exoplanets (e.g. Bedding et al. 2010; McQuillan et al. 2014; Slawson et al. 2011). The short cadence observations were obtained for a small fraction of targets (0.3% in Q0; Gilliland et al. 2010) and for a fraction of the mission duration. The short cadence observations enabled studies of variability which were not possible with the long cadence mode, in particular the detection of short duration stellar flares.

Kepler did not obtain photometry for every single star within its field of view, instead targets were pre-selected (e.g. Batalha et al. 2010). A “postage stamp” of pixels was selected for each target per quarter, within which the data was collected and stored. Each postage stamp was selected to include the flux from the star and surrounding background. The size of each postage stamp varied with the magnitude of each target, with brighter stars requiring larger stamps. The average size of a postage stamp in the full set of short cadence observations is 12x6 pixels, approximately 48”x32”. The postage stamp pixel files were downloaded at the end of each Kepler quarter and made available for download as “Target Pixel Files”. These target pixel files include the individual pixel images, timestamps, quality flags, row and column image centroids and the optimal aperture. Lightcurves and ancillary products (e.g. image centroids) for each target were automatically generated from these target pixel files using the optimal apertures, which were designed to maximise the signal-to-noise ratio (Haas et al. 2010).

As each postage stamp includes some amount of sky around the target star, there is a chance nearby sources will have been included in the resultant target pixel files. If these stars are close enough to the target star and they flare with enough energy, then this flare will be visible in the Kepler lightcurve. By inspecting the individual images and the image centroids we can identify these events and use them to characterise previously unexplored stellar flares.

3 METHODS

3.1 Detrending and Flare Detection

In this work we used the Simple Aperture Photometry (SAP) data for all stars with short cadence observations. Before searching for flares we detrended the SAP lightcurves. This was done to remove instrumental trends and astrophysical variability (e.g. due to starspots) which may hinder accurate flare detections. Our detrending process is similar to those found in previous flare studies, in particular Yang & Liu (2019), and is as follows.

For each target lightcurve, we initially masked out all data points with quality bit flags equal to 1, 2, 3, 4, 5, 6, 7, 9, 13, 15, 16 and 17. These are the recommended quality flags masked by the Kepler data reduction pipeline and a flag for reaction wheel zero crossing (Thompson et al. 2016). Before smoothing a lightcurve using a median filter, we checked whether the target source had a known rotation period, using the catalogs of Reinhold et al. (2013) and...
McQuillan et al. (2014). This was to to ensure the selected window could effectively smooth out any rotational modulation present. If the source did not have a known rotation period, we searched for the strongest period with a generalised Lomb Scargle periodogram, using the astropy LombScargle routines (Astropy Collaboration et al. 2013). We sampled each lightcurve with periods between 100 minutes and half the duration of the full lightcurve. If the best period from the generalised periodogram had a power greater than 0.25 then this period was chosen. This power threshold was chosen empirically to avoid false positive periods in our dataset (e.g. Oelkers et al. 2018). After this routine we smoothed the lightcurve.

Before smoothing we initially split the lightcurve into separate continuous segments, using gaps in coverage greater than 30 minutes to split the data. This was done to avoid jumps in the measured flux between quarters, or segments with quarters, affecting our smoothing algorithm. Continuous segments less than one quarter of a day were masked. For each continuous segment we applied a median filter. The window size was selected to be equal to one tenth the measured period, with lower and upper limits of 30 minutes and 12 hours respectively. These limits were chosen to make sure lightcurves were not excessively smoothed at the lower level (in turn removing detected flares) or suffered from overly large window sizes. For sources without measured periods we used a standard window size of 100 minutes, which was used by Pugh et al. (2016) to remove lightcurve modulation without removing flares. Each segment was then iteratively smoothed. In each iteration, the original lightcurve was divided by a smoothed version. Outlying points greater than $3\sigma$ from the median of the resultant lightcurve were then masked. The equivalent points were masked in the original lightcurve and were interpolated over. This process was repeated up to 20 times, or until there were no more outlying points. Once the process was completed for a given continuous segment, the original segment (which contains the flares and rotational modulation) was divided by the final smoothed version (which should only contain the rotational modulation); to give what should be a flare-only segment. We then searched for flares in this detrended segment. To identify flares in the detrended lightcurve segment, we searched for consecutive outliers $3\sigma$ above the lightcurve median. Here $\sigma$ refers to the standard deviation of the detrended lightcurve. Regions with more at least three consecutive outliers were automatically flagged as flare candidates. The times at which the lightcurve first goes above and below the $3\sigma$ limit are recorded as the flare start and end time. Flare candidates starting and ending within 2.4 hours (0.1 days) of each other were combined to avoid multiple detections of the same event (e.g. due to a multi-peaked flare). Once all continuous segments had been run for a target, all flare candidates were combined and saved.

Some sources in the short cadence data are known eclipsing binaries. Deep eclipses in the Kepler lightcurves, if not masked, can increase the measured variance. This in turn will increase the $3\sigma$ limit used for detecting flares, meaning only the highest amplitude and energy events will be detected. Due to the effects of dilution, flares from neighbouring sources are unlikely to have large amplitudes in the original Kepler lightcurves. If a target source was a known eclipsing binary and it is included in the Kepler eclipsing binary catalog\footnote{http://keplerebs.villanova.edu/} (Prša et al. 2011; Slawson et al. 2011; Kirk et al. 2016), we masked its eclipse before smoothing the lightcurve and searching for flares. Eclipses were masked by iteratively applying a median filter to the lightcurve with a user-defined window of one fifth the orbital period, then masking outliers $3\sigma$ below the median of the lightcurve. These masked points were flagged and stored. The masked points were then interpolated over. This process was repeated either 20 times, or until no points were flagged as outliers. The stored masked points in the original lightcurve were then replaced using a cubic spline fit using anchor points two and one hour either side of the masked eclipses. This new lightcurve, with the eclipses replaced by a smoothed spline, was then put forward for the iterative smoothing and automated flare detection.

Flare candidates flagged by our automated method were then vetted to determine their authenticity. To vet flare candidates we first flagged candidates which had the same start and end times, as determined from the Kepler BJD time stamps. Stellar flares are known to be stochastic events, meaning any candidate events starting and ending at the same times on different stars are more likely to be instrumental than astrophysical in origin. Remaining flare candidates were visually inspected to remove those due to other sources of astrophysical variability (e.g. RR Lyrae or high frequency pulsators, e.g. Yang & Liu 2019), or those which may have been flagged as a result of our detrending algorithm.

### 3.1.1 Centroiding and Image Analysis

After vetting all automatically flagged candidates to obtain a clean sample, we inspected the centroids and images of all our flares. The purpose of this was to identify which flares arise from stars neighbouring the target stars. For all events we obtained a pre-flare image from two hours before the start of the flare and the image at the peak of the flare. We subtracted the pre-flare image from the image at the flare peak to obtain a residual, “flare only”, image. We compared the pre-flare, residual flare only and original flare images for every event. Flares from stars neighbouring the target source will have different source locations in the pre-flare and residual images. An example is shown in Fig. 1. We identified candidate flares from nearby stars through visual inspection of the residual images. This method of comparing images was used in Yang et al. (2017) to identify possible blends and flares from nearby sources. We also used the centroid position as an extra confirmation that flares came from a nearby object. Flares which come from stars other than the target star will shift the centre of light of the aperture towards the true flare source.

For all sources we inspected whether the pipeline aperture fully encompassed the PSF of the flaring source. In some cases the flare source lay in a pixel adjacent to the edge of the pipeline aperture. In these cases, light from the flare would enter the aperture of the target source, resulting in the appearance of a small flare event. Where required we created new apertures by eye, designed to cover the PSF of the flaring source. All the new apertures were made to include the light from both the target and neighbouring star, due to the separations not being large enough to fully disentangle the PSFs of the two stars. We used the new apertures to generate improved lightcurves directly from the target pixel files, upon which our flare detection and verification algorithms were run again to find previously undetected flares.

### 3.2 Determining the contribution of all stars

For apertures covering multiple stars it was important to determine the relative flux contribution from each source. Flux from nearby sources will dilute flare lightcurves and if not corrected for, this results in reduced measured flare amplitudes and energies. To determine the flux contribution from sources within an aperture we
calculated their *Kepler* magnitudes. As some sources in our apertures did not have entries in the *Kepler* input catalog, we decided to calculate the *Kepler* magnitudes of each source ourselves. Where possible, we calculated this by applying an offset to their catalogue *Gaia* G magnitude. The offset between *Gaia* G and the *Kepler* magnitude was determined by crossmatching all sources within the *Kepler* input catalog with those in *Gaia* DR2 which had passed the recommended photometric and astrometric quality checks (Arenou et al. 2018; Lindegren et al. 2018). When sources did not have a *Gaia* G band magnitude, but did have an entry in the *Kepler*-INT catalog, we used the available INT g, r and i magnitudes with Eq. (2)-(5) from Brown et al. (2011) to estimate the *Kepler* magnitude. The dilution is calculated as the ratio of the flux from the flaring stars to the flux from all stars within the aperture.

To obtain broadband photometry for all our identified flare sources, we performed crossmatching with various photometric catalogues. We crossmatched each source with *Gaia* DR2 (Gaia Collaboration et al. 2018), APASS (Henden & Munari 2014), 2MASS (Skrutskie et al. 2006), WISE (Cutri & et al. 2014), Pan-STARRS (Chambers et al. 2016) and the *Kepler*-INT survey. Sources which were unresolved in APASS, 2MASS or WISE, but are resolved in *Gaia*, were flagged for reference during our SED fitting.

### 3.3 Stellar Properties

Where possible we fit the Spectral Energy Distribution (SED) of each of our flare stars using the PHOENIX v2 spectral library (Allard et al. 2012). In order to calculate flare energies, as we will discuss in Sect. 3.4, we required a star’s radius and effective temperature. For all stars we fitted for the effective temperature $T_{\text{eff}}$, the surface gravity $\log g$, interstellar extinction $A_V$ and an uncertainty scale factor $\sigma$. We used the uncertainty scale factor to account for underestimated variability in reported catalogue photometry. When fitting for interstellar extinction we assumed average Milky Way parameters for the reddening and used $R_V = 3.1$, where $R_V$ is the ratio of total to selective extinction (Cardelli et al. 1989). We used the 3D dust maps of Green et al. (2019) to apply Gaussian priors on our fitted extinction values. These dust maps are based on *Gaia* DR2, Pan-STARRS and 2MASS and can be used to estimate the expected level of extinction along a given line-of-sight, both in total and as a function of distance. Each template SED was multiplied by a scale factor $S$, equal to $(R_* / d)^2$, where $R_*$ is the stellar radius and $d$ is the distance. For stars which had a distance measurement from Bailer-Jones et al. (2018), we fit directly for the stellar radius $R_*$ and the distance $d$. We applied a Gaussian prior to the distance, using the values from Bailer-Jones et al. (2018). We included the distance in our fitting to propagate its uncertainty to our fitted ra-

![Figure 1](image.png)

*Figure 1.* Example of the flare image and centroiding analysis, for a flare from the M3 dwarf KIS J192818.57+383757.2. The top panel shows the pre-flare, residual and peak flare images. The green squares indicate the position of the original *Kepler* aperture. The green circle is the position of the target source in the *Kepler* input catalog. The purple crosses indicate the positions of stars in the Kepler-INT survey. The bottom three panels show the detrended lightcurve and the positions of the row and the column centroid. Note the clear shift in centroid position during the flare, showing it did not come from the target source.
dius. For stars without a distance measurement, we fit the scale factor only. To determine the radii of these objects, we assumed the star was on the main sequence and calculated the radius using the temperature-radius relations of Bailer-Jones et al. (2012, 2017).

Some sources were too close to a nearby, brighter, neighbour to have non-blended photometry in all of our crossmatched catalogues. In these cases, where possible, we fit the SED of both the flaring star and the nearby neighbour simultaneously, following a method similar to that used in Jackman et al. (2019). In this scenario we fit for the effective temperature, surface gravity, interstellar extinction and scale factor of both stars simultaneously, along with the uncertainty scale factor σ. We fit the combined SED of both stars to blended catalogue magnitudes and the individual SEDs to Gaia resolved catalogue photometry. This is primarily the Gaia G band photometry, although depending on the separation can include Pan-STARRS.

To fully explore the posterior parameter space we used the emcee Python package (Foreman-Mackey et al. 2013) to generate a Markov Chain Monte-Carlo (MCMC) process. For all SED fits we ran the MCMC process for 10,000 steps using 32 walkers and discarded the first 3000 as a burn-in. Some stars did not have enough broadband photometry for SED fitting. For these stars we used the $T_{\text{eff}}$ and radii colour-colour relations from Stassun et al. (2019) with the Gaia G, RP and BP photometry and the distance measurements from Baker-Jones et al. (2018). These relations were created for the TESS Input Catalog (TIC).

3.4 Calculating Flare Energies

Before calculating the energy we normalised each flare lightcurve by dividing it by the quiescent flux value. We assumed that any stellar variation (e.g. due to rotational modulation) during the flare changed on timescales longer than the flare duration and as such could be fitted with a linear baseline, which was obtained by linearly interpolating the lightcurve just before and after each flare event. The flare lightcurve was then divided by this baseline to obtain the normalised flare lightcurve. The quiescent baseline of this normalised lightcurve was set to zero. To account for the possible effects of dilution from other stars in our apertures, we divided each normalised flare lightcurve by the dilution values calculated in Sect. 3.2. This renormalised flare lightcurve was then used to measure the flux amplitude and calculate the flare energy. We express the flare amplitude using $\Delta F/F_0$ where $\Delta F$ is the flux due to the flare and $F_0$ is the quiescent flux of the star (e.g. Hawley et al. 2014). Using this formulation, a value of 0 indicates no flare is present, while a value of 1 indicates the flare emits as much flux (in the chosen filter) as the quiescent star, or a doubling of the total observed flux.

To calculate the energy of each flare we followed the method outlined by Shibayama et al. (2013). We assumed the flare spectrum can be modelled by a single blackbody with a temperature of 9000 ± 500 K. Multi-colour photometric and spectroscopic flare observations (Hawley & Fisher 1992; Kowalski et al. 2013) have shown that flares can be approximated as a 9000 K blackbody. However, these and more recent studies have also shown that flares exhibit changing temperatures between individual events, with continuum blackbody temperatures up to 15,000 and in some cases even up to 40,000 K being measured from optical and far ultraviolet observations of the impulsive phase of flares (Loyd et al. 2018; Froning et al. 2019; Howard et al. 2020). In addition to this, the flare temperature has been observed to change within events, cooling during the flare decay. We calculated how the measured flare energy, using the Shibayama et al. (2013) method with the Kepler bandpass, would change if we used temperatures of 6000 K and 16000 K, values representative of the minimum and maximum flare temperatures measured spectroscopically by Kowalski et al. (2013). We found that the measured energies would range from 0.8 to 2.5 times that the value calculated when a 9000 K blackbody is assumed. Consequently, the assumption of a 9000 K blackbody is only an approximation to the true behaviour of white-light flares and may underestimate the energy, particularly during the flare rise when the flare temperature reaches its maximum. Studies apply a 500 K error to partially account for the changing flare temperature, as the changing flare temperature behaviour cannot be accurately predicted. We note that this uncertainty may need to be revised in future, as more measurements of flare temperatures become available. However, as this method has been used several times in previous flare studies using Kepler and TESS (e.g. Shibayama et al. 2013; Yang et al. 2018; Günther et al. 2020; Tu et al. 2020) it allows for a comparison between our results and previous studies and as such we use it in this work.

4 RESULTS

From our survey of the Kepler short cadence data we identified 4430 flares from 403 stars. We identified that 515 flares in 26 Kepler short cadence lightcurves were either due to a nearby source, or due to there being flares from both the target star and a close companion (likely excluding it from other surveys). From the 26 Kepler lightcurves there were 34 individual flaring stars, 26 of which were neighbouring the target source. The remaining eight were cases when both stars in the postage stamp flared. 30 of the 34 stars had enough catalogue photometry for SED fitting or the Stassun et al. (2019) colour-colour relations. Further inspection with revised apertures, as discussed in Sect. 3.1.1, resulted in an extra 68 flare detections from nearby stars, increasing the total number of flares to 4498, with 583 flares from 34 stars. A full table of our results is provided in the online version of this work.

4.1 Flare Amplitudes

We measured the flare amplitudes using the method outlined in Sect. 3.4, by accounting for the effects of dilution from nearby stars. 515 of the flares in the final sample of 583 flares were from M stars and 16 of these had corrected flare amplitudes greater than 1. Note that these 515 flares are not the full 515 originally detected using the standard Kepler apertures. The full sample of M dwarf flares have amplitudes between 0.002 and 35.7. The highest amplitude flare, from the M3 dwarf KIS J192818.57+383757.2 (PS1 1543529207740999705), is equivalent to $\Delta m_{\text{Kp}}=3.9$. Assuming a 9000 K blackbody, we estimate that this would be equivalent to a 4.8 magnitude change in the V filter and 7.5 magnitudes in the U filter for that star. The original Kepler lightcurve for this flare is shown in Fig. 1. This flare is also greater in amplitude than any of those studied from K2 observations of M dwarfs by Lin et al. (2019) and those in the sample of M dwarf hyperflares studied using Kepler and LAMOST by Chang et al. (2018). However, while having higher amplitudes than many flares studied previously using Kepler or K2 data, the larger flares in our sample are comparable to those from other surveys which studied much greater numbers of M dwarfs, in particular EvryScope (Howard et al. 2019), ASASSN (Schmidt et al. 2019; Rodriguez Martínez et al. 2020) and TESS (Günther et al. 2020).
4.2 Flare Energies

The flare energies were calculated following the method outlined in Sect. 3.4. The flare energies from all stars in our sample range from $3.7 \times 10^{29}$ to $1.5 \times 10^{35}$ erg. The flare energies from the target sources sit within this range. The maximum flare energy as a function of spectral type, compared with the sample of flares detected from Kepler and K2 by Yang et al. (2018) and the sample of flares detected with TESS short cadence observations by Günther et al. (2020) is shown in Fig. 2. We can see that the flares from neighbouring M stars in our sample have maximum energies comparable to those previously detected using TESS short cadence data, while some appear to push the limit for flares detected with the Kepler long cadence data. We believe this is due to the larger number of M dwarfs observed with TESS relative to Kepler, aiding the detection of high energy events.

We believe that the reason why our M star energies are preferentially higher than previous studies with Kepler is due to a selection effect related to our detection and confirmation methods, which we will explain here. Due to their low luminosities, an M star neighbouring a brighter star (e.g. a G star) will not contribute much flux in the Kepler aperture. Therefore, only the highest energy flares, with the exact energy depending on the flux ratio between the two stars, will contribute enough flux to trigger our detection method and also cause an observable centroid shift. Small flares are less likely to trigger our detection method. As we go to higher luminosity stars, which contribute more quiescent flux, we can detect lower energy flares. This can be seen in Fig. 2, where our K stars broadly reside compared to those in previous studies. Some stars in our sample flared multiple times, allowing us to trigger our detection method.

The maximum energies of the F, G and K spectral type stars are comparable to those measured in previous studies. We can also see some flares in Fig. 2 that are at the lower limit of the Kepler sensitivity. These sources are the hotter, brighter, components of pairs of stars in which both components flare. Some of these sources are in wide binaries, which we will discuss in Sect. 3.4, and are less active than their lower mass companions over the Kepler lifetime. Consequently, these stars flare less often and in a given time span appear to flare with a lower maximum energy.

4.3 Flare Occurrence Rates

Some stars in our sample flared multiple times, allowing us to measure the flare occurrence rate. Previous studies have shown that flares occur with a power-law-like distribution in energy, written as

$$dN(E) \propto E^{-\alpha} dE$$

(1)

where $N(E)$ is the number of flares in a given duration with energy $E$ and $\alpha$ is the power law index. The number of flares which occur with an energy greater than $E$ can then be written as

$$\log N(E_{f1} > E) = C + \beta \log E$$

(2)

where $C$ is a normalisation constant and $\beta = 1 - \alpha$. To measure $C$ and $\alpha$ for stars in our sample, we fit the flare occurrence rates using the “powerlaw” Python package (Alstott et al. 2014). powerlaw is designed to fit to heavy-tailed distributions and has previously been used to fit occurrence rates for flares observed with Kepler (e.g. Lin et al. 2019). For each flaring star in our sample we used powerlaw to identify the energy above which flares were consistent with a power-law distribution. Flares with energies lower than this value were those affected by selection effects due to the sensitivity of our detection method. This effect is well documented in stellar flare literature (e.g. Gershberg 1972; Kunkel 1973; Hawley et al. 2014) and results in a turnover in the observed occurrence rate at low energies. If a star had five or more flares above the limiting energy, then we fitted Eq. 2 to its measured occurrence rate using a least squares fitting. We have presented the best fitting values of $C$, $\beta$ and $\alpha$, along with the information about each star and the number of flares detected and fit to, in Tab. 1.

We can see in Tab. 1 that many of the stars in our sample have $\alpha > 2$. A value of $\alpha$ greater than 2 indicates that the total flare energy distribution of a star is dominated by low energy flare events (e.g. Güdel et al. 2003). If they dominate the total energy distribution, then low energy events such as nanoflares have been suggested as being able to heat the observable X-ray corona (Parker 1988; Hudson 1991). We note then that if our fitted power law distributions hold to nanoflare energies ($\sim 10^{24}$ erg; Parker 1988), then flares may be a major contributor to the heating of the quiescent corona for some of the stars in our sample.

4.4 Wide Binaries

After stars arrive on the zero-age main sequence and start to undergo magnetic braking, their flaring activity is expected to decrease. This behaviour has been observed in observations of open clusters (Ilin et al. 2019), which provide a snapshot of a specific time of stellar evolution. The change in flare activity is expected to have both a dependence on age and stellar mass, as solar-type stars for example spin down faster than their lower mass counterparts. Another way of comparing how flaring activity changes with mass and age is to use wide binaries. Wide binary systems are pairs of stars with similar proper motions and distances. Their similar kinematics means wide binaries are thought to have formed together and be coeval, making them excellent laboratories for testing age-activity relations (e.g. Gunning et al. 2014).

We have investigated whether any of our flaring sources are components of wide binaries or common proper motion pairs. We crossmatched the position of each of the 34 individual flaring stars with Gaia DR2 to obtain the parallax and proper motion information. We crossmatched sources with total proper motions greater...
than 40 mas yr\(^{-1}\) with the SUPERWIDE catalog (Hartman & Lépine 2020) of wide binaries. 7 flaring stars in our sample were listed as members of wide binaries in the SUPERWIDE catalog. For sources with total proper motions less than 40 mas yr\(^{-1}\), we checked to see if there were nearby sources with similar parallaxes and proper motions. For each candidate pair we calculated the difference in distance, their projected physical separation and the difference in their tangential velocities. For a pair to be considered a wide binary, we required their projected separation be less than 10000 AU, the difference in tangential velocity be less than 10 km s\(^{-1}\) and that at a bolometric energy of 10\(^{33}\) erg. We have used this energy as an upper limit when fitting the occurrence rates for GJ 1245A and B. We find that, on average, GJ 1245A flares more than twice as often as GJ 1245B and that at a bolometric energy of 10\(^{33}\) erg (assuming a 9000K blackbody) GJ 1245A flares once every eight days and GJ 1245B once every 17 days.

We find, for all binary pairs in our sample that, as expected, the lower temperature component flares more often than the higher temperature component. We have either inferred this from the absence of flares detected from the hotter component, or from measured flare occurrence rates. This result is in agreement with Clarke et al. (2018). For pairs with flares from both stars, we extrapolated the occurrence rate measured from the lower temperature component to the maximum energy detected from the higher temperature component to estimate a ratio of the flare occurrences. The flare energies, effective temperatures and the estimated ratios are given in Tab. 2.

### 5 DISCUSSION

#### 5.1 Incidence of originally detected flares

As stated in Sect 4, 515 flares in our original search came from stars nearby the original Kepler target source, or were from multiple flaring stars contributing to the same lightcurve. This value comprises 11.7 ± 0.5 per cent of the flares in our original search. This value is similar to the 10-12 per cent false positive rate reported by previous Kepler flare studies (which excluded stars with separations less than 12″, e.g. Shibayama et al. 2013; Gao et al. 2016). We investigated our flare sample further to identify only those flares which came from a neighbouring star. This excludes the flares in wide binaries which come from the target Kepler source. We found that 298 flares

---

**Table 1.** Best fitting power law parameters for the flare occurrence rates of stars in our sample. \(E_{\text{lim}}\) is the minimum energy limit used when fitting the power law distribution. \(N_{\text{flares,fit}}\) is the number of flares used in the fitting. \(E_{\text{max}}\) is the maximum flare energy. The last column gives the predicted waiting time to observe a flare with an energy of 10\(^{33}\) erg or greater. The * symbol indicates this is an extrapolation of our fitted power law. The † symbol indicates that these stars, in GJ 1245, had an upper energy limit of 2 × 10\(^{33}\) erg imposed when fitting.

| Name | Teff (K) | C | α | \(E_{\text{lim}}\) (erg) | \(E_{\text{max}}\) (erg) | \(N_{\text{flares,fit}}\) | Time for 10\(^{33}\) erg (days) |
|------|---------|---|---|-------------------|-------------------|----------------|---------------------|
| Gaia DR2 2099133865420292480 | 4353 ± 134 | 41.4 ± 9.0 | 2.26 ± 0.27 | 2.5 × 10\(^{33}\) | 1.3 × 10\(^{34}\) | 13 | 1.4 |
| Gaia DR2 2076733526527061767 | 3790 ± 263 | 41.3 ± 12.1 | 2.27 ± 0.37 | 7.1 × 10\(^{32}\) | 7.2 × 10\(^{33}\) | 11 | 5.5 |
| Gaia DR2 2101855569017361280 | 4142 ± 60 | 15.5 ± 9.6 | 1.50 ± 0.29 | 5.6 × 10\(^{32}\) | 1.2 × 10\(^{34}\) | 5 | 9.2 |
| Gaia DR2 2104786042380468736 | 3686 ± 125 | 29.2 ± 3.3 | 1.91 ± 0.10 | 1.8 × 10\(^{32}\) | 1.9 × 10\(^{33}\) | 21 | 5.9 |
| Gaia DR2 2105165163343324000 | 5544 ± 175 | 26.6 ± 11.3 | 1.81 ± 0.33 | 6.4 × 10\(^{33}\) | 5.2 × 10\(^{34}\) | 7 | 1.0 |
| Gaia DR2 2105392991397882624 | 2964 ± 29 | 14.9 ± 4.4 | 1.48 ± 0.14 | 8.7 × 10\(^{31}\) | 2.3 × 10\(^{33}\) | 9 | 10.1 |
| Gaia DR2 2105930840143687040 | 3275 ± 291 | 29.0 ± 6.6 | 1.94 ± 0.20 | 6.1 × 10\(^{32}\) | 4.4 × 10\(^{33}\) | 11 | 106 |
| Gaia DR2 2126410068846203776 | 5099 ± 269 | 47.6 ± 21.1 | 2.46 ± 0.62 | 6.0 × 10\(^{33}\) | 1.7 × 10\(^{34}\) | 6 | 2.7 |
| Gaia DR2 2130462537972407044 | 4424 ± 28 | 45.6 ± 11.2 | 2.39 ± 0.34 | 1.0 × 10\(^{33}\) | 6.5 × 10\(^{33}\) | 12 | 2.5 |
| Gaia DR2 2131725447307160321 | 3602 ± 260 | 42.6 ± 22.9 | 2.39 ± 0.72 | 6.7 × 10\(^{31}\) | 2.5 × 10\(^{32}\) | 6 | 2000 |
| Gaia DR2 2131725456195069441 | 3190 ± 276 | 54.7 ± 4.8 | 2.74 ± 0.15 | 7.3 × 10\(^{31}\) | 3.7 × 10\(^{32}\) | 31 | 735 |
| Gaia DR2 2132768956904826624 | 3431 ± 156 | 50.1 ± 4.8 | 2.58 ± 0.15 | 2.3 × 10\(^{32}\) | 1.3 × 10\(^{34}\) | 36 | 95 |
| Gaia DR2 2079073928612891860 | 2776 ± 131 | 51.0 ± 4.9 | 2.68 ± 0.16 | 8.4 × 10\(^{30}\) | 2.0 × 10\(^{31}\) (†) | 35 | 17400 |
| Gaia DR2 2079073928612817602 | 2763 ± 121 | 36.8 ± 1.6 | 2.23 ± 0.05 | 2.5 × 10\(^{30}\) | 1.6 × 10\(^{31}\) (†) | 66 | 4830 |

---

**MNARAS 000, 1–10 (2021)**
in our original search came from a neighbouring star (26 stars), whether that was the fainter component of a wide binary or an unrelated nearby source. This new filter reduces the false positive rate to $6.7 \pm 0.4$ per cent.

Filtering our sample even further to keep only those sources where the nearby star flared (i.e. removing wide binaries in which both components flared) results in a subset of 61 flares from 19 stars. This is $1.4 \pm 0.2$ per cent of the total number of flares in our original search. In this analysis we have kept flares from non-target stars which have their own entry in the Kepler input catalog but lack a distinct short cadence lightcurve.

### 5.2 Application to TESS flare studies

TESS has since visited the Kepler field and studies are now emerging which are combining data from both missions (e.g. Davenport et al. 2020). Two of the main differences between the Kepler and TESS telescopes are the field of view and the pixel scale. The Kepler telescope had a total field of view of 115 deg$^2$ and a pixel scale of approximately 4", while each camera on the TESS telescope has a field of view of 24x24 deg$^2$ and a pixel scale of 21" (Ricker et al. 2014). The larger pixel scale of TESS will not only cause a greater number of stars to be blended within each aperture, but pairs of stars which were previously resolved with Kepler may be placed on the same TESS pixel. To confirm whether this was the case for the flare stars in our sample, we calculated both their TESS pixel position and the pixel positions of nearby sources in the TESS input catalog. This was done using the Lightkurve Python package (Lightkurve Collaboration et al. 2018) with the TESS sector 14 observations. We found that 21 of the 26 neighbouring flare stars in our sample fell on the same TESS pixel as the original Kepler target source. The remaining five fell on an adjacent pixel, but all were within 21".

In Sect. 5.1, we calculated that $6.7 \pm 0.4$ per cent of flares in our sample were due to a nearby star either in or neighbouring the pipeline Kepler aperture. The smaller pixel size of Kepler relative to TESS means we can use our survey to estimate the false positive rate of flares in TESS flare studies. Specifically, due to the smaller aperture and postage stamp sizes of Kepler compared to the TESS stamps and full frame images, we can estimate the false positive rate of flares from stars within 21", i.e. one TESS pixel distance of the target star.

Due to the different noise properties between Kepler and TESS we cannot directly apply our calculated estimate of false positive flare events to large scale flare studies using TESS data. In addition, the TESS bandpass is not the same as the one used by Kepler and it probes a different portion of the flare spectrum. To estimate the number of flares and stars which might cause false positive detections in large scale flare studies using TESS short cadence observations, we initially calculated the amplitude of all of our detected flares if they had been observed with the TESS bandpass. This combines the flares detected in our original search and those detected after we had created new apertures. To do this for each star we used the best fitting effective temperatures from Sect. 3.3 to normalise a 9000 K flare blackbody spectrum to the observed amplitude in the Kepler bandpass. Both the normalised flare blackbody spectrum and the PHOENIX v2 spectrum were then convolved with the TESS filter to estimate the new amplitude. We then recalculated the dilution from nearby stars in our apertures, using their TESS magnitudes obtained from the TESS input catalog. The calculated flare amplitude in the TESS bandpass was multiplied by the TESS dilution coefficient to estimate the flare amplitude in initial TESS lightcurves. When calculating the dilution we consider the contributions from stars within a 40 arcsecond radius of the target star (e.g. Tu et al. 2020), to better estimate the dilution in TESS apertures.

We then used the sample of flares detected with the two minute cadence TESS observations from Günther et al. (2020) to determine the distribution of detectable flares in amplitude-TESS magnitude space. We used this distribution to estimate a lower flare amplitude limit as a function of TESS magnitude, below which no flares could be detected. We compared this lower limit to our measured flare amplitudes, from our detrended Kepler lightcurves, and the TESS magnitude of each target. We have used the total sample of flares here, i.e after we revised the apertures to fully include neighbouring stars, to better estimate the larger apertures used for TESS relative to Kepler (e.g. Tu et al. 2020). We found that of the 4498 flares detected in our adjusted sample, 674 would have an amplitude large enough to be detectable with TESS short cadence observations. Of these 674 detectable flares, 39 were from stars other than the target source. From this we estimate a false positive rate of $5.8 \pm 1.0$ per cent for flares from stars within 21". This value can be used as an estimate of the rate of false positive events from stars within one pixel as the target in TESS flare studies. This value shows that for large scale flare studies using TESS short cadence data, a non-negligible fraction of flares will come from stars close to the target source. This effect can be avoided by carefully vetting input catalogs to select only stars with no nearby neighbours.

In this analysis we estimated the dilution in TESS by using all stars from within a 40 arcsecond radius. However, many of the Kepler image stamps do not extend to 40 arcseconds from the target star, meaning flaring stars at greater separations will have been missed. Consequently, for real TESS apertures which extend further than 21 arcseconds from the target source, our estimated false positive rate will be a lower limit. We also note that the false positive fraction will increase in more crowded regions of sky, such as open clusters and those nearer to the galactic plane. As such, caution should be applied when using the quoted values in these regions for studies which have not applied careful vetting of their input catalogues.

| Source ID  | $T_{\text{eff},1}$ (K) | Source ID  | $T_{\text{eff},2}$ (K) | Flare Energy (erg) | Ratio of 1 to 2 |
|-----------|----------------------|-----------|----------------------|-------------------|----------------|
| Gaia DR2 2076733562507061766 | 3790$^{+233}_{-253}$ | Gaia DR2 207673356259065856 | 3807$^{+513}_{-338}$ | $1.2 \times 10^{31}$ | 5.0 |
| Gaia DR2 2131725451649506944 | 3190$^{+203}_{-222}$ | Gaia DR2 213172547350716032 | 3602$^{+462}_{-59}$ | $2.5 \times 10^{32}$ | 3.8 |
| Gaia DR2 213276895604826624 | 3437$^{+166}_{-176}$ | Gaia DR2 2132768952604988672 | 4759$^{+320}_{-310}$ | $2.9 \times 10^{32}$ | 26.9 |
| Gaia DR2 207907928612819840 | 2776$^{+356}_{-34}$ | Gaia DR2 207907928612821760 | 2763$^{+33}_{-33}$ | $1.7 \times 10^{31}$ | 1.9 |

Table 2. Flare occurrence rate ratios for stars in wide binaries where at least one component had a measured flare occurrence rate. The flare energy is the maximum measured energy for the less active component. For non-equal mass binaries, this is the hotter component.
6 CONCLUSIONS

In this work we have presented the results of a search for stellar flares in the Kepler short cadence data. We specifically searched for flares which came from stars neighbouring the target sources, which have been excluded from previous surveys. From our search of the short cadence lightcurves using the original Kepler apertures we initially detected 4430 flares from 403 stars. Of these, 515 flares came from either a neighbouring source or were from multiple flaring stars contributing to the same lightcurve. We calculated that these flares comprised $11.7 \pm 0.5$ per cent of the Kepler short cadence sample. We found that this value is similar to that reported by previous surveys for the fraction of false positives and blends. Of these 515 flares, 298 were from a neighbouring star. This resulted in a measured false positive rate of $6.7 \pm 0.4$ per cent in our original sample. We have used our sample and the measured false positive rates to calculate the fraction of flares which would come from sources other than the target star in TESS two minute cadence observations. We calculated that $5.8 \pm 1.0$ per cent rate of flares in these surveys would come from stars within $21^\circ$ of the target source and recommend careful vetting of input catalogs to minimise the effect of flares from blended sources.

Our samples and reported false positive rates includes neighbouring stars within $12^\circ$. This is a sample which has been ignored by many previous flare studies using the Kepler data. By including these close neighbours in our sample we have been able to study the flaring behaviour of wide binaries, the components of which reside close together on the sky. From our sample of wide binaries, we found that the lower mass components are more active than their higher mass partners during our observations, in agreement with previous studies for the activity of coeval pairs of stars. Our sample highlights how the careful treatment of flares from neighbouring stars can aid in the study of these coeval systems and how flare rates and energies vary between spectral type.

Where possible we calculated the flare amplitudes and energies for events in our sample, accounting for the flux from both the target source and the flaring neighbour. We found that the flares from neighbouring sources in our sample are some of the highest energy white-light flares observed with Kepler, with energies up to $1.5 \times 10^{45}$ erg. We found that the majority of these flares come from M dwarfs which neighbour the Kepler target sources. Our study shows how the flares from stars neighbouring the target sources in Kepler, previously discarded as contaminants, can be used to study high energy flare events and flare occurrence rates. We have focused on the Kepler short cadence data in this study, however this data only comprises a small fraction of the total available Kepler data. We anticipate that a similar study for the long cadence data will be able to expand upon on the results presented in this study, however will likely be biased to higher energy events due to the 30 minute cadence missing short duration flares.

ACKNOWLEDGEMENTS

JAGJ acknowledges support from grant HST-GO-15955.004-A and R.O.P.L. acknowledges support from grant HST-GO-14784.001-A from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–26555. This paper includes data collected by the Kepler mission. Funding for the Kepler mission is provided by the NASA Science Mission directorate. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. The Pan-STARRS1 Surveys (PS1) and the PS1 public science archive have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, the Queen’s University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation Grant No. AST-1238877, the University of Maryland, Eotvos Lorand University (ELTE), the Los Alamos National Laboratory, and the Gordon and Betty Moore Foundation. We acknowledge with thanks the variable star observations from the AAVSO International Database contributed by observers worldwide and used in this research.

DATA AVAILABILITY

All lightcurves and catalogue broadband photometry used in this work are publicly available. Derived stellar and flare properties are available upon request to JAGJ.

REFERENCES

Allard F., Homeier D., Freytag B., 2012, Philosophical Transactions of the Royal Society of London Series A, 370, 2765
Ailott J., Bullmore E., Plenz D., 2014, PLoS ONE, 9, e85777
Antonucci E., Dennis B. R., 1983, Sol. Phys., 86, 67
Arenou F., et al., 2018, A&A, 616, A17
Astropy Collaboration et al., 2013, A&A, 558, A33
Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, AJ, 156, 58
Batalha N. M., et al., 2010, ApJ, 713, L109
Bedding T. R., et al., 2010, ApJ, 713, L176
Benz A. O., Güdel M., 2010, ARA&A, 48, 241
Borucki W. J., et al., 2010, Science, 327, 977
Boyajian T. S., et al., 2012, ApJ, 757, 112
Boyajian T. S., et al., 2017, ApJ, 845, 178
Brown T. M., Latham D. W., Everett M. E., Esquerdo G. A., 2011, AJ, 142, 112
Bryson S. T., et al., 2013, PASP, 125, 889
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
