



# Stellar flares detected with the Next Generation Transit Survey

James A. G. Jackman<sup>1,2,3\*</sup>, Peter J. Wheatley<sup>1,2</sup>, Jack S. Acton<sup>4</sup>, David R. Anderson<sup>1,2</sup>,  
Daniel Bayliss<sup>1,2</sup>, Joshua T. Briegal<sup>5</sup>, Matthew R. Burleigh<sup>4</sup>, Sarah L. Casewell<sup>4</sup>, Boris T. Gänsicke<sup>1,2</sup>,  
Samuel Gill<sup>1</sup>, Edward Gillen<sup>1,5,6†</sup>, Michael R. Goad<sup>4</sup>, Maximilian N. Günther<sup>7,‡</sup>, Beth A. Henderson<sup>4</sup>,  
Simon T. Hodgkin<sup>8</sup>, James S. Jenkins<sup>9,10</sup>, Chloe Pugh<sup>1</sup>, Didier Queloz<sup>5</sup>, Liam Raynard<sup>4</sup>,  
Rosanna H. Tilbrook<sup>4</sup>, Christopher A. Watson<sup>11</sup> and Richard G. West<sup>1</sup>

<sup>1</sup>Department of Physics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK

<sup>2</sup>Centre for Exoplanets and Habitability, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK

<sup>3</sup>School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287, USA

<sup>4</sup>School of Physics and Astronomy, University of Leicester, Leicester LE1 7RH, UK

<sup>5</sup>Astrophysics Group, Cavendish Laboratory, J.J. Thomson Avenue, Cambridge CB3 0HE, UK

<sup>6</sup>Astronomy Unit, Queen Mary University of London, Mile End Road, London E1 4NS, UK

<sup>7</sup>Department of Physics, and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>8</sup>Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

<sup>9</sup>Departamento de Astronomía, Universidad de Chile, Camino El Observatorio 1515, Las Condes, Casilla 36-D, Santiago, Chile

<sup>10</sup>Centro de Astrofísica y Tecnologías Afines (CATA), Casilla 36-D, Santiago, Chile

<sup>11</sup>Astrophysics Research Centre, Queen's University of Belfast, 1 University Road, Belfast BT7 1NN, UK

Accepted 2021 April 5. Received 2021 February 24; in original form 2020 December 23

## ABSTRACT

We present the results of a search for stellar flares in the first data release from the Next Generation Transit Survey (NGTS). We have found 610 flares from 339 stars, with spectral types between F8 and M6, the majority of which belong to the Galactic thin disc. We have used the 13-s cadence NGTS light curves to measure flare properties such as the flare amplitude, duration, and bolometric energy. We have measured the average flare occurrence rates of K and early to mid-M stars and present a generalized method to measure these rates while accounting for changing detection sensitivities. We find that field age K and early M stars show similar flare behaviour, while fully convective M stars exhibit increased white-light flaring activity, which we attribute to their increased spin-down time. We have also studied the average flare rates of pre-main-sequence K and M stars, showing they exhibit increased flare activity relative to their main-sequence counterparts.

**Key words:** stars: flare – stars: rotation – starspots.

## 1 INTRODUCTION

The study of stellar flares in recent years has benefited greatly from the introduction of wide-field long-duration surveys. The first observations of these stellar phenomena were either from targeted observations (e.g. Lacy, Moffett & Evans 1976; Alekseev et al. 1994; Andrews & Panagi 1996), or were by chance (e.g. F and G star superflares Schaefer, King & Deliyannis 2000). Targeted and wide-field studies have revealed that stars down to and including the L spectral type (e.g. Schmidt et al. 2007, 2016; Gizis et al. 2017) could produce flares with radiated energies hundreds to thousands of times that produced by the Sun ( $10^{32}$  erg for the 1859 Carrington event; Carrington 1859; Tsurutani et al. 2003). These flares occur on nearly all main-sequence (MS) stars with an outer convective envelope, with the trigger believed to be due to magnetic reconnection (Shibata 1999; Benz & Güdel 2010). These reconnection events cause the

acceleration of charged particles into the chromosphere, resulting in plasma evaporation and the release of energy from radio to X-ray wavelengths (Benz & Güdel 2010; Benz 2017). When observed in visible wavelengths, these flares are termed ‘white-light flares’ (e.g. Paudel et al. 2018). The observable behaviour of flares is known to change with spectral type, with K and M stars showing more frequent detectable high-energy flares than their higher mass counterparts (Balona et al. 2015). This is due in part to the increased contrast between the flare and quiescent spectrum for lower mass stars and in part due to the decreased photospheric luminosity of these stars relative to their solar-type counterparts. Both of these factors result in flares of a given energy appearing larger in a given filter and more easily detectable on low-mass stars. Stars with more active magnetic fields have also shown greater flaring activity than their inactive counterparts (Hilton et al. 2011).

Due to their likelihood to flare at least once in a given night, early ground-based observations of white-light flares typically focused on nearby magnetically active M dwarfs such as AD Leo (e.g. Oskanian 1969; Pettersen, Coleman & Evans 1984). Such campaigns have been used to study statistical distributions of flare properties (e.g. the flare energy occurrence rate Lacy et al. 1976; Gershberg &

\* E-mail: jamesjackman@asu.edu

† Winton Fellow

‡ Juan Carlos Torres Fellow

**Table 1.** NGTS fields used in this work, along with the number of nights and sources observed per field.

Field	First night	Last night	Nights	N <sub>Sources</sub>
NG0522-2518	2015 Sep 21	2016 May 03	149	7435
NG0531-0826	2015 Sep 23	2016 Apr 20	138	10178
NG0549-3345	2016 Aug 06	2017 Jan 31	126	9074
NG0612-2518	2015 Sep 21	2016 May 14	157	13 452
NG0618-6441	2015 Sep 21	2016 May 24	159	11 689
NG1135-2518	2015 Nov 26	2016 Aug 03	126	6407
NG1213-3633	2015 Nov 28	2016 Aug 04	116	10 625
NG1315-2807	2016 Jan 05	2016 Aug 30	105	8176
NG1340-3345	2016 Apr 18	2016 Aug 31	67	11 433
NG1416-2518	2016 Jan 05	2016 Sep 11	100	8589
NG1421+0000	2016 Jan 05	2016 Aug 31	100	4279
NG1428-2518	2016 Jan 05	2016 Sep 15	109	8840
NG1444+0537	2016 Jan 13	2016 Sep 03	99	4181
NG2025-1941	2016 Apr 20	2016 Nov 28	89	12 862
NG2028-2518	2016 Apr 20	2016 Nov 30	121	11 496
NG2047-0248	2016 Apr 20	2016 Nov 28	122	12 565
NG2058-0248	2016 Apr 20	2016 Dec 01	126	10 491
NG2126-1652	2016 Apr 20	2016 Dec 12	126	6766
NG2132+0248	2016 Apr 20	2016 Dec 02	119	8537
NG2142+0826	2016 Apr 20	2016 Dec 02	121	9077
NG2145-3345	2016 Apr 20	2016 Sep 02	60	5615
NG2150-3922	2016 Apr 20	2016 Dec 21	145	5905
NG2331-3922	2016 May 03	2017 Jan 10	129	3433
NG2346-3633	2016 May 03	2017 Jan 10	150	3495

Shakhovskaia 1983) on these single stars, along with the temporal morphology of flares (e.g. Rodono 1974). In addition to these studies, targeted spectroscopic campaigns have been used to study the spectral signature of flares (e.g. Kowalski et al. 2013). However, as these works are typically limited to nearby active M stars it is difficult to scale up the highly detailed observations from these studies to large sample sizes.

Wide-field exoplanet surveys have helped to partially overcome this limitation, by providing simultaneous observations of thousands of stars for long periods of time. Most notable for white-light flare studies was the *Kepler* space telescope, which provided continuous observations of over 150 000 target stars for 4 yr in its primary mission (Borucki et al. 2010). Stars were observed in either long or short cadence modes of 30 and 1 min, respectively. Results from *Kepler* include the detection of ‘superflares’ on solar-type stars (Maehara et al. 2012; Shibayama et al. 2013) that before *Kepler* were rarely observed events (Schaefer 1989; Schaefer et al. 2000), studies of quasi-periodic pulsations (Balona et al. 2015; Pugh et al. 2016), and improved understanding of flare occurrence rates (e.g. Hawley et al. 2014). This work is now being carried forward with the Transiting Exoplanet Survey Satellite (*TESS*; Ricker et al. 2015), in particular with the 2-min short-cadence mode during the primary mission and more recently, the 20-s fast-cadence mode. The 2-min cadence *TESS* observations have already been used to find flares down to late-M stars (e.g. Günther et al. 2020; Medina et al. 2020) and to study the relation between flare occurrence and star-spot phase (Doyle et al. 2019; Feinstein et al. 2020). However, it should be noted that the short cadence mode observations of both *Kepler* and *TESS* represent only a small fraction of the total data collected (e.g. approximately 0.3 percent for the *Kepler* Q1 data; Gilliland et al. 2010). The vast majority of observations for both missions are instead at the 30-min long-cadence mode, of pre-selected targets for *Kepler* and full-frame images for *TESS*. As the typical time-scale of stellar flares is on the order of minutes to hours (e.g. Hawley

et al. 2014), longer cadence modes can have the effect of smearing out their appearance in light curves and can result in inaccurate measurements of flare parameters (Yang et al. 2018). Along with this, substructure can be removed and short duration flares are missed entirely. Consequently, for precise measurements of flare parameters on a large sample of stars, high-cadence observations with a wide field of view are required.

Such observations have become possible in recent years with ground-based optical surveys such as the Next Generation Transit Survey (NGTS). NGTS is a ground-based transiting exoplanet survey located at the ESO Paranal Observatory in Chile. NGTS is formed of 12, 20 cm f/2.8, optical telescopes operating with an exposure time of 10 s and total field of view of  $\approx 96 \text{ deg}^2$  (Wheatley et al. 2018). NGTS is designed to be sensitive to the bright K and M stars ( $I \leq 16$ ), so each telescope is fitted with a custom Andor iKon-L 936 camera, with a back-illuminated deep-depletion CCD. The bandpass is between 520 and 890 nm. Unlike the *Kepler* and *TESS* short cadence modes, NGTS observes without a set target list, meaning that high cadence light curves can be generated for all stars in our field of view.

Results to date from searching for stellar flares with NGTS include G star superflares (Jackman et al. 2018), the detection of quasi-periodic pulsations from a flare on a pre-main-sequence (pre-MS) M star (Jackman et al. 2019a) and the first detection of a white-light flare from an L2.5 dwarf (Jackman et al. 2019b). These are single flare events, and it is clear that the wide field of view full-frame images of NGTS can also lend themselves to larger scale studies of flare parameters.

In this work, we present a survey of stellar flares detected in the first data release of NGTS (NGTS DR1<sup>1</sup>). We describe our detection method and our method for measuring stellar and flare parameters, then proceed to present our results. We discuss the properties of the measured flare parameters, and compare with previous results from *Kepler* and from the ground. We also discuss how our observations complement those from previous surveys.

## 2 OBSERVATIONS

The 24 fields of data used in NGTS DR1 span observation dates between a earliest date of 2015 September 21 and a latest date of 2017 January 31. A full list of the fields used and information about them can be found in Table 1. Throughout this work, we use the standard NGTS apertures that are generated for all stars within our field of view with  $I < 16$ , as described by Wheatley et al. (2018).

### 2.1 Characterizing stars

The NGTS pipeline is designed to cross-match all input catalogues with a range of catalogues to provide broad-band photometric information for all observed sources. This pipeline currently performs matching with *Gaia* DR2 (Gaia Collaboration 2016, 2018a), 2MASS (Skrutskie et al. 2006), APASS (Henden & Munari 2014), UCAC4 (Zacharias et al. 2013), RAVE (Kunder et al. 2017), *GALEX* (Martin et al. 2005), and ALLWISE (Wright et al. 2010, Cutri et al. 2014). Along with this, distance information is obtained from *Gaia* DR2 using the catalogue of Bailer-Jones et al. (2018).

To obtain the stellar parameters for each source, flaring or non-flaring, we used the revised *TESS* Input Catalog (TIC; Stassun et al. 2019). The TIC is a collection of cross-matched catalogue photometry and stellar parameters for sources across the whole sky.

<sup>1</sup><http://eso.org/rm/api/v1/public/releaseDescriptions/122>

Importantly for our analysis the TIC provides effective temperatures and radii (using *Gaia* DR2 parallaxes) for effectively all sources in our input catalogue. For each NGTS field, we matched the results of our cross-matching with the TIC using the *Gaia* DR2 source ID, or the 2MASS source ID where *Gaia* DR2 did not provide a match. For this work, we have used the stellar parameters provided in the TIC. We found that many radii values in the TIC did not have an associated uncertainty. In these cases, we assumed a 10 per cent uncertainty on the quoted radius. We associated stars with a spectral type using their TIC effective temperature and the effective temperature-spectral type tables from Pecaut & Mamajek (2013).<sup>2</sup>

At times in our analysis, it was useful to separate main-sequence and non-main-sequence stars. To do this, we used a Hertzsprung–Russell (HR) diagram, as shown in Fig. 1. We calculated a median absolute magnitude curve using all astrometrically and photometrically ‘clean’ *Gaia* sources (i.e. those passing the filters of Arenou et al. 2018) within 100 pc (e.g. Gaia Collaboration 2018b). This curve was measured as a function of the TIC v8 effective temperature. We then defined empirical upper and lower limits in absolute *Gaia* G magnitude as  $\Delta M_G = +0.65$  and  $\Delta M_G = -0.55$  from the median curve. We used  $\Delta M_G = -0.55$  as a lower limit to remove the majority of equal mass binaries and pre-main-sequence stars and ensure our main-sequence sample is dominated by field age stars. It has previously been noted that binary systems pervade throughout the entirety of the HR diagram (e.g. Gaia Collaboration 2019; Winters et al. 2019), making it difficult to choose a limit that will remove them without removing large numbers of legitimate isolated main-sequence stars. To analyse pre-main-sequence stars alone, we use a separate limit of  $\Delta M_G = -0.753$ . This is something we discuss further in Section 4. The  $\Delta M_G = +0.65$  upper limit is chosen empirically in order to remove cool subdwarfs while keeping most main-sequence stars. Our chosen limits and how they compare to the sample of stars observed with NGTS can be seen in Fig. 1.

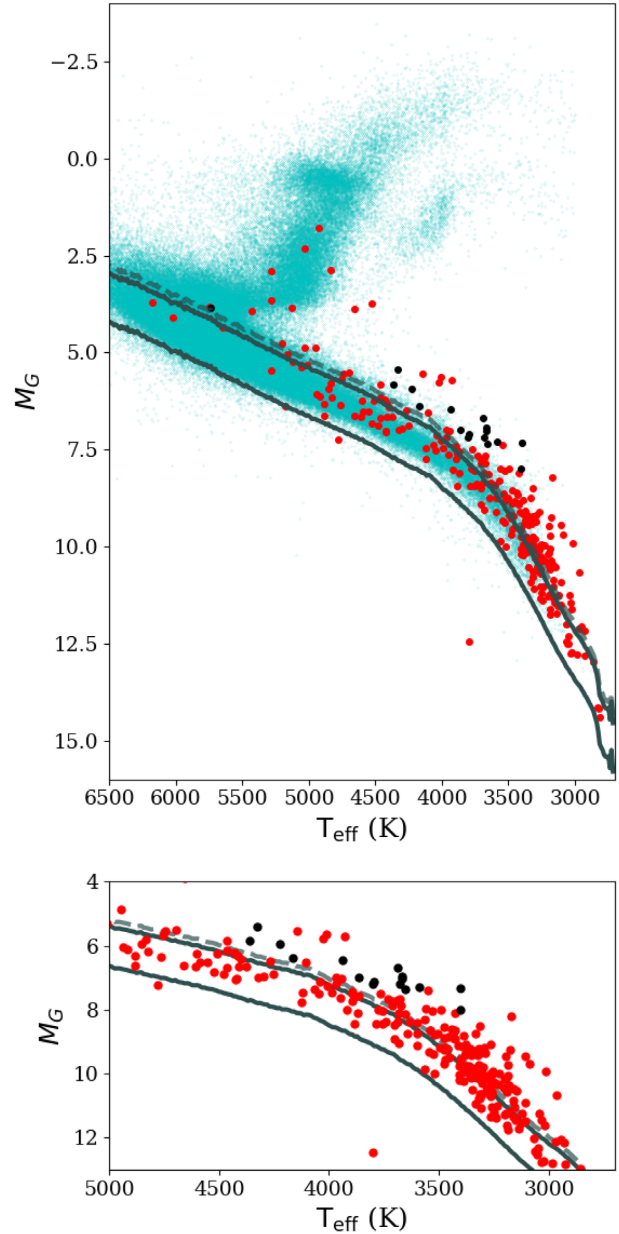
From this analysis, we identified 35612 K0V–K8V stars and 4135 M0–M5 stars that are consistent with our definition of being isolated and on the main sequence. Of the K stars, 27694 are K0V–K4V and 7918 are K5V–K8V. Of the M stars, 3757 are M0V–M2V and 378 are M3V–M5V.

## 2.2 Flare detection

To search for stellar flares in each NGTS field, we started with the raw light curves. We detrended these light curves using a custom version of the NGTS SYSREM process (Wheatley et al. 2018). During this process we applied an extra filter to remove images with excess variance (e.g. clouds across the image).

This filter was defined empirically using the airmass,  $X$ , and  $\sigma_{\text{excess}}$  values calculated by the NGTS pipeline and SYSREM.  $\sigma_{\text{excess}}$  is a measure of the per-image variance introduced by external factors (e.g. clouds) and is expressed in units of magnitude (Collier Cameron et al. 2006).  $\sigma_{\text{excess}}$  used by the SYSREM pipeline to down-weight images where there are lots of outlying points across multiple stars. We removed all observations with  $\sigma_{\text{excess}} > 0.02$  for airmasses below or equal to 1.2. For airmasses above 1.2, we removed all points with  $\sigma_{\text{excess}} > 0.008 + 0.01X$ . This filter limited the effects of clouds without penalising images at higher airmasses that tend to have inherently higher variance, along with not removing outlying data points due to large flares.

<sup>2</sup>[http://www.pas.rochester.edu/~emamajek/EEM\\_dwarf\\_UBVIJHK\\_color\\_s\\_Teff.txt](http://www.pas.rochester.edu/~emamajek/EEM_dwarf_UBVIJHK_color_s_Teff.txt)



**Figure 1.** Hertzsprung–Russell diagram of the stars observed with NGTS and cross-matched with *Gaia* DR2 and the TIC v8 catalogue, using the TIC v8 effective temperatures. All stars from the NGTS fields are shown in cyan, while the flaring stars are depicted with the red circles. The black circles are sources associated with Orion, using the best-fitting SEDs from Jackman et al. (2020). The dark grey lines indicate the upper and lower main-sequence limits specified in Section 2.1, and the dashed line is the pre-main-sequence cut-off. The bottom plot shows a zoomed section with only the flare stars and main-sequence limits displayed.

When searching for flares in the detrended light curve for each star, we split the light curve into individual nights. The majority of reported stellar flare durations are on the order of minutes to hours (e.g. Hawley et al. 2014; Maehara et al. 2016), with only the largest exceeding this. This means almost all of the observed flares in the NGTS light curves should be confined within single nights. Splitting the data into individual nights also reduces the effect of longer time-scale variations.



For each night, we performed a two-step flare detection method. This is the same method as was used by Jackman et al. (2018, 2019a, 2020). For clarity, we outline the method here as well. First, we look for discontinuities within the night, by searching for regions where there are at least three consecutive points above six median absolute deviations (*MAD*) from the median of the night. The *MAD* is a measure of the variance in a sample and is calculated as the median of the absolute deviations from the sample median. We use *MAD* over the standard deviation as it is a robust parameter, so is less likely to be altered by the presence of single point outliers, such as cosmic rays. Regions with at least three consecutive points above six *MAD* are flagged as flare candidates. As a second independent step, we compare the median of the night against the median of the entire light curve for that object. If the median of the night is at least five *MAD* above the median light curve flux, then it is flagged. We do this as this is indicative of a long duration and high-amplitude flare that may not be fully observed within one night (such as the flare discussed in Jackman et al. 2019a, that dominated the entire night of observation). Once this automated flagging procedure was complete, we inspected each flagged night visually and removed any false positives. Such false positives can be due to astrophysical events (e.g. short-period variable stars), human-made (e.g. satellites passing through our aperture), or systematics (e.g. overlapping apertures).

Events that passed our checks and were classified as flares were also inspected for signs of complexity. Flares in previous works have been classified as ‘simple’ and ‘complex’. Simple flares typically follow the classical fast rise, exponential decay shape (e.g. Moffett 1974). Complex flares deviate from this template, showing multiple peaks (e.g. Davenport et al. 2014), oscillations (e.g. Anfinogentov et al. 2013) or other substructure that might complicate their analysis. Complex flares may also arise from the superposition of two simultaneous independent flare events. Each flare light curve was visually inspected for signs of multiple peaks and oscillations, which if present resulted in a complex flag being applied to that event.

### 2.2.1 Centroiding

We have performed a centroiding analysis for each of our flare candidates. This was done in order to confirm the true source of each flaring event. If a star has a flaring companion nearby, we would expect the centroid position to shift relative to the centre of the aperture for the duration of the flare. Using this shift, we can identify cases where the detected flare is not coming from the expected source and where possible assign the correct *Gaia* source to the flaring star (e.g. Günther et al. 2017).

Flux from nearby stars within an aperture will dilute the observed light curves. This will result in reduced measured flare amplitudes, hindering their detection. The reduced flare amplitudes will also result in decreased measured flare energies. Close stars may also affect the quality of the stellar characterization discussed in Section 2.1. Accurate stellar temperatures and radii are essential for calculating flare energies. Therefore, for each NGTS target, we searched for neighbouring *Gaia* DR2 or 2MASS sources within 15 arcsec, the radius of an NGTS aperture. Stars with nearby neighbours that could cause dilution are flagged during our analysis and only used in sections that do not require accurate stellar characterization.

### 2.3 Flare amplitude

We calculated our flare amplitudes using  $\frac{\Delta F}{F_q}$ .  $\Delta F$  is the change in observed flux due to the flare, while  $F_q$  is the quiescent flux. An

amplitude of 1 indicates that the flare is emitting as much flux in the NGTS filter as the quiescent star. We measured the quiescent flux for each flare event by taking the median of the 20 min preceding the flare. If the flare dominated the entire night, we used the median value of the previous night. We confirmed the calculated quiescent flux and amplitude for each flare through visual inspection. For events where the quiescent flux calculation was incorrect (e.g. due to including another flare), we manually corrected the values.

### 2.4 Flare energy

To calculate our flare energies we have followed the method of Shibayama et al. (2013), which assumes that both the star and the flare behave as blackbody radiators in order to estimate the flare luminosity. The flare energy is calculated from the time integral of the luminosity over the total flare duration, which we discuss in Section 2.6. In this method, it is assumed that the flare spectrum can be broadly approximated as a blackbody with a temperature of  $9000 \pm 500$  K. This assumption is based on spectroscopic and multicolour photometric observations (Hawley & Fisher 1992; Kowalski et al. 2013). However, while these studies have shown that the optical spectrum of many flares can be broadly approximated with a 9000–10 000 K blackbody, the flare temperature has been observed to change both between and within individual flare events to values outside this range (e.g. Howard et al. 2020a). We note that while this limits the use of a single temperature model, we currently cannot predict these changes in the flare temperature, preventing the use of more complex models. When calculating the energy we assumed the underlying variation of the stellar flux is approximately linear during the flare. For each flare, a baseline of the stellar flux was estimated by linearly interpolating the light curve between just before and just after the flare, as determined by visual inspection. To calculate the uncertainty on the flare energy, we performed our flare energy calculation using a Monte Carlo process of 10 000 runs. This enabled us to propagate the uncertainties in the stellar effective temperature, radius, and the flare temperature to the final measured flare energy.

### 2.5 Flare occurrence rate

Previous studies (e.g. Lacy et al. 1976; Hilton et al. 2011; Hawley et al. 2014) have shown that flares occur with a power-law distribution in energy. This is typically written as

$$dN(E) = kE^{-\alpha}dE, \quad (1)$$

where  $N$  is the number of flares that occur in a given duration,  $E$  is the flare energy,  $k$  is a constant of proportionality, and  $\alpha$  is the power-law index. The flare frequency distribution, the number of flares per unit time with an energy of  $E_f$  or greater, typically denoted by  $\nu$  is obtained by integrating equation (1) from  $E_f$  to infinity, resulting in

$$\log \nu = C + \beta \log E_f, \quad (2)$$

where  $C = \log(\frac{k}{1-\alpha})$  and  $\beta = 1 - \alpha$  (Hawley et al. 2014).

The value of  $\alpha$  dictates how often high-energy flares occur relative to those of lower energy and has been observed to vary across different spectral types. For example, the Sun has  $\alpha \approx 1.75$  (e.g. Crosby, Aschwanden & Dennis 1993; Shimizu 1995; Aschwanden et al. 2000), while ultracool dwarfs have had measured  $\alpha$  values from 1.4 to 2 (e.g. Paudel et al. 2018).

We have used the NGTS data to study flare occurrence rates and where possible measure  $k$  and  $\alpha$  (and in turn  $C$  and  $\beta$ ) for subsets of our sample. In particular, we are aiming to study how the average occurrence rate changes as a function of spectral type, by combining

observations of flaring and non-flaring stars in our sample. We have done this for our main-sequence and pre-main-sequence samples.

### 2.5.1 Fitting the average flare occurrence rate

When measuring the average occurrence rate in each subset it was important to ensure the sample was complete across both flaring and non-flaring stars. Incompleteness in measured occurrence rates typically manifests itself as a flattening at low energies (in log–log space), where the measured distribution no longer matches the expected power-law behaviour. At this turnover, the chosen detection method is no longer sensitive to all flares and fewer are detected. Many studies have gotten around this incompleteness by fitting only to data with flare energies above some limit above the turnover (e.g. Hawley et al. 2014). A common way of determining this energy limit is to inject flares of known amplitude, duration, and energy and to measure the fraction of flares recovered as a function of energy,  $R(E)$ . The limiting energy can then be the minimum value recovered (e.g. Paudel et al. 2018), or the energy at which some percentage of flares are recovered (e.g. 68 per cent; Davenport 2016; Jackman et al. 2020). This has the advantage of obtaining a sample of flares that are mostly free of detection-sensitivity related effects, but will result in smaller samples due to the rarer nature of higher energy flares.

Other studies have used the measured flare recovery rate to correct the observed flare frequency distribution, to which a power law is then fit (e.g. Ilin et al. 2019; Kővári et al. 2020). This has the benefit of keeping lower energy flares in the fitted distribution while still accounting for the detection efficiency. However, care must be taken with this method, as the measured flare recovery rate from flare injection is often for individual energies and not for all energies above some value. Consequently, the recovery fraction as a function of energy cannot be directly applied to equation (2). Instead, it must be applied to equation (1) and the new ‘observed’ differential distribution of flare energy should be integrated from some energy  $E_f$  to infinity. We do this for a generalized manner in Appendix A to obtain the observed flare frequency distribution, accounting for detection sensitivity effects:

$$N(E > E_f) = \frac{k}{\alpha - 1} \left( R(E_f) E_f^{-\alpha+1} + \int_{E_f}^{E_{\max}} R'(E) E^{-\alpha+1} dE \right), \quad (3)$$

where  $R'(E)$  is the differentiated flare recovery fraction and  $E_{\max}$  is the energy at which the recovery fraction saturates [ $R'(E) = 0$ ]. At low energies this will give the observed turnover effect, while at high energies it will be equivalent to equation (2). Medina et al. (2020) recently presented a similar analysis specifically for *TESS* observations, which highlights the necessity of taking the recovery fraction into account in flare studies.

### 2.6 Flare duration

In our analysis we calculated both the full flare duration and the  $t_{1/2}$  time-scale, defined as the time where the flare is above half of its maximum flux (e.g. Davenport et al. 2014). The full flare duration was measured by eye for each flare and was used in our energy calculation. We have used the  $t_{1/2}$  time-scale when comparing our flare durations with the flare amplitude and energy and with other studies. We have chosen the  $t_{1/2}$  time-scale for this analysis over the e-folding time as used in, e.g. Maehara et al. (2016) or the full flare duration as it uses the brightest, and best resolved, part of the flare to provide a precise measurement of the time-scale of

both the flare rise and decay. We used the 13-s cadence NGTS data when determining the  $t_{1/2}$  value. Each flare was visually inspected to make sure the measured  $t_{1/2}$  values were sensible (e.g. not thrown off due to outliers). Some flares are known to exhibit multiple peaks, particularly during the flare decay, where the flux may rise back above the threshold value. When this occurs we measure the first and last point where the light curve drops below half of the flux maximum. This way we can obtain a minimum and maximum time-scale for the more complex flares.

### 2.7 Completeness

An important factor to consider when analysing the distribution of flare stars is the completeness of our sample. This is particularly vital when comparing our sample of flaring stars to non-flaring stars, such as when probing the fraction of stars to show a flare of a given energy. For fainter stars, we expect our detection threshold to be at higher flare amplitudes than for brighter stars, due to the increased relative variance in their light curves. This will result in us only detecting higher energy flare events for a faint star than for a bright star of the same spectral type.

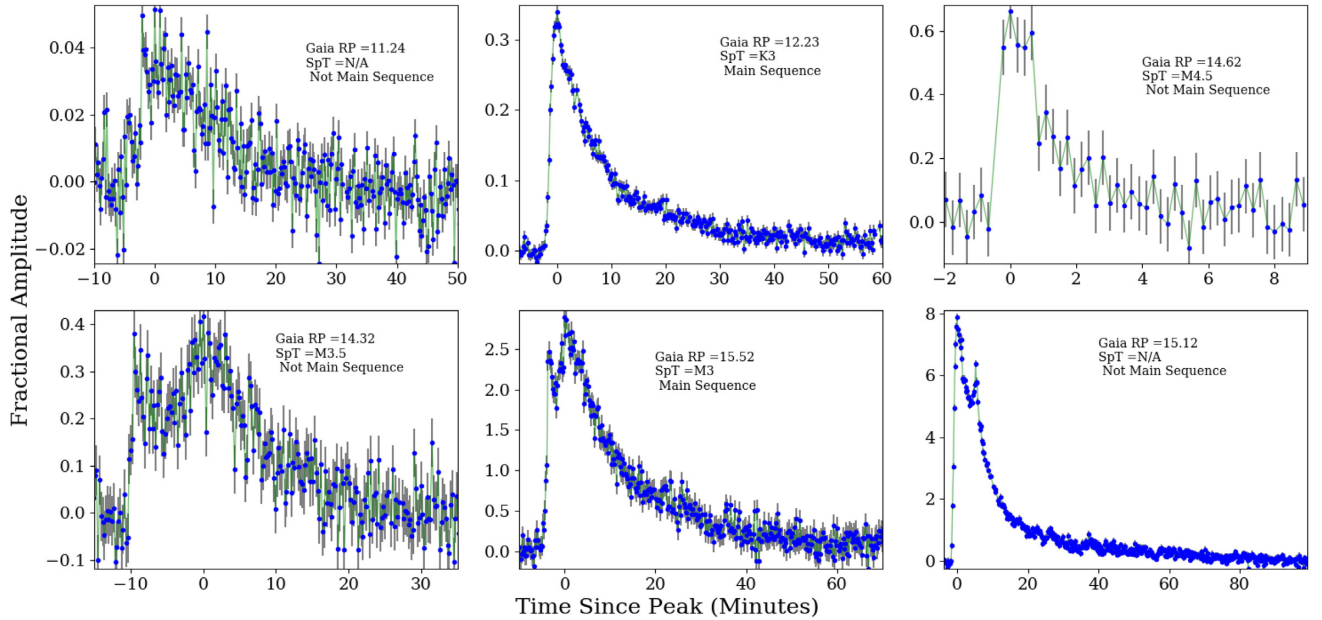
To parametrize the completeness of our sample, we performed flare injection and retrieval tests. These tests were performed on all K and M stars in our sample, whether or not they are on the main sequence. We followed the method outlined by Jackman et al. (2020), which itself was based on the flare injection and retrieval method used by Davenport (2016). We injected 50 artificial flares into each night of every K and M star light curve in our sample. For the average light curve in our sample, this resulted in roughly 6000 flares being injected and tested.

These flares were created using the Davenport et al. (2014) empirical flare model. Amplitudes for the artificially generated flares were chosen randomly from a uniform distribution between 0.01 and 4 times the light-curve median. The  $t_{1/2}$  durations were chosen randomly from a uniform distribution between 30 s and 70 min. Every flare had its energy calculated using the method given in Section 2.4, using the TIC v8 effective temperatures, radii, and associated uncertainties from Section 2.1.

We then used our detection method from Section 2.2 with our light curves with injected flares, to see how many flares we could retrieve. If an injected flare was flagged as a candidate by our detection method, it was given a flag of 1. If it was not detected, it was given a flag of 0 for that star. The end result for each star was a catalogue of flares, with known energies, and a corresponding array for whether each flare, if it had occurred on that star, would have been detected in our survey. We calculated how many flares were retrieved as a function of energy for each star. We used 20 logarithmically spaced bins in energy which, following Davenport (2016), were smoothed using a Wiener filter of three bins. The flare energies and recovery fractions were then recorded for later use.

### 2.8 Stellar rotation

Many active stars show periodic modulation in their light curves. This is due to the presence of dark star-spots on their surface that appear and disappear from view as the star rotates (e.g. Maehara et al. 2017; Doyle et al. 2018). We searched the light curves of all flare stars for signs of periodic modulation consistent with rotation. To measure rotation periods, we used the *ASTROPY* package *LombScargle* (Astropy Collaboration 2013). We initially masked any identified flares in our light curves to remove their effect on the measured power spectrum. We then constructed a frequency grid formed of 2000 periods spaced



**Figure 2.** A selection of flares detected in the first data release from NGTS. The blue circles are individual NGTS data points, at the standard 13-s cadence. The green line joins these points to aid the eye. The bottom row shows examples of complex flares. Inset in each panel is the *Gaia* RP magnitude, spectral type used in our analysis and whether the source lies on the main sequence. In this plot, all sources marked as not being on the main sequence are all above the main sequence. Note the short duration of the flare on the top right, which would not have been resolved in either *Kepler* 1 minute or *TESS* 2-min cadence modes.

linearly between 13 s and 1 d, then 20 000 periods spaced linearly between 1 and 30 d. For each flare, star light curve we used the full time resolution data to calculate a generalized Lomb–Scargle periodogram, which returns normalized power values between 0 and 1 (corresponding to a perfect straight line and a perfect sine wave, respectively). To determine the best-fitting period of our sample, we used a similar method as Oelkers et al. (2018). First, we remove the effects of aliases, by rejecting periods within 5 per cent of single day aliases and those due to the lunar month ( $29 \pm 3$  d) and ignoring the associated signals in our analysis. This is done iteratively for each star until the best remaining period is not due to these aliases. We then checked the remaining periods to see whether their harmonics fall into any of our identified alias regions. We checked 0.5P, 1.5P, 2P, and 3P for this and rejected any period that fails this test. Again, this is done iteratively until the best remaining period does not fail this test. Once we were satisfied the best period was not due to a systematic alias, we checked whether its normalized power was above 0.1 (Oelkers et al. 2018). This was done to remove spurious periods with weak powers in the periodogram and to keep only those where we were confident a true astrophysical signal was present.

### 3 RESULTS

From searching the 196 600 stars in 24 fields in NGTS DR1, we have identified 610 flares from 339 stars. A selection of the detected flares is shown in Fig. 2. Those on the main sequence have spectral types between F8V and M6V. We note that the cut-off at spectral type of M6V is due to the limiting magnitude of  $I = 16$  for the standard NGTS light curves.

#### 3.1 Distribution

To probe the distribution of stars in our sample, we use the results of our cross-matching pipeline and TIC v8 from Section 2.1. For sources

with centroid shifts, we have manually corrected the cross-matching to give the correct *Gaia* DR2 source where possible.

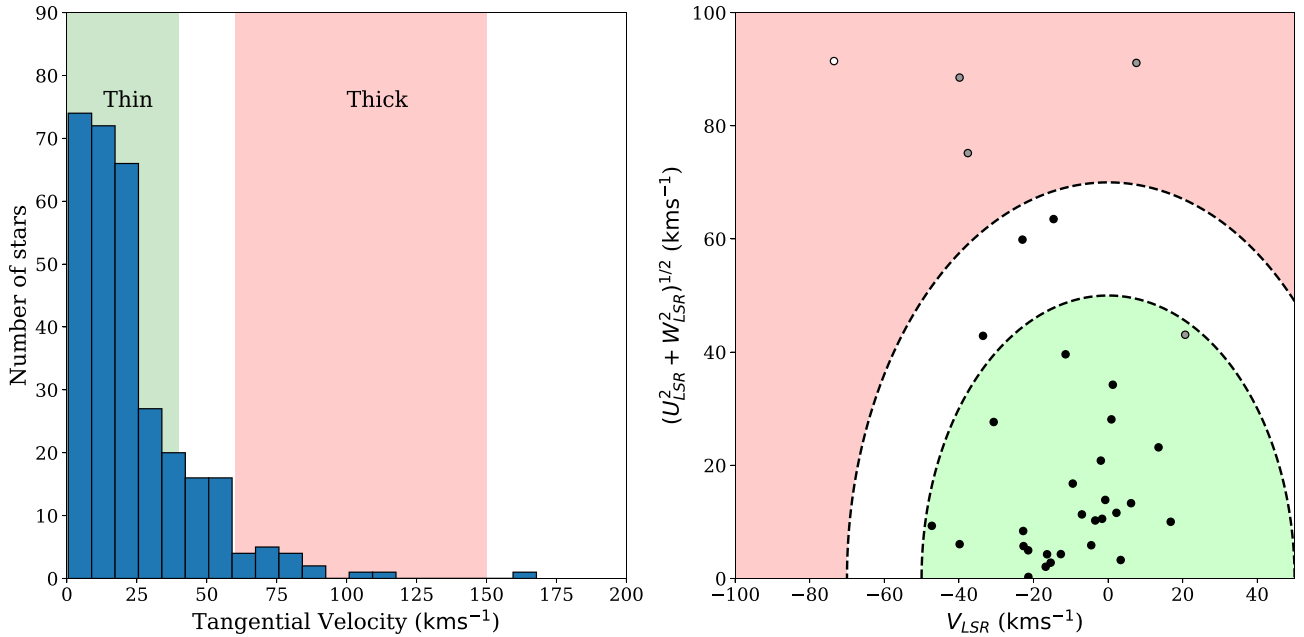
##### 3.1.1 Centroiding

Through centroiding, we found that 11 per cent of flares in our final sample were due to a star not at the centre of the NGTS aperture. This is similar to the 10 per cent of ‘false sources’ identified by Shibayama et al. (2013) when analysing *Kepler* pixel level data for G dwarf superflares. As *Kepler* and NGTS operate with pixel scales of 4 and 5 arcsec, respectively, the similar level of blending is expected. These percentages highlight the importance of centroiding for accurate flare analysis. This is particularly relevant for *TESS*, which has a pixel scale of 21 arcsec per pixel (Ricker et al. 2014). We therefore expect *TESS* to suffer from a higher fraction of blended flare sources, making careful centroid analysis vital.

An alternative solution that has been used previously (e.g. Shibayama et al. 2013; Yang et al. 2017) is to choose isolated stars, avoiding the need for centroiding of every source. While this has the benefit of providing a cleaner input catalogue, it can limit the number of observed systems and remove potentially interesting events (e.g. the quasi-periodic pulsation, QPP, exhibiting flare from Jackman et al. 2019a). Along with this, it will remove serendipitous observations of flares from stars too faint to see in quiescence that may be located close to a brighter source.

##### 3.1.2 Galactic distribution

Fig. 1 shows the position of our flaring sources on a HR diagram. We calculate that 56 per cent of stars observed with NGTS are consistent with being an isolated main-sequence star, according to the criteria from Section 2.1. We calculate that approximately 3 per cent of stars observed with NGTS are pre-main sequence. Of the flare sources, 32 per cent of our sample are consistent with being an isolated



**Figure 3.** Left: The tangential velocity,  $V_{\text{tan}}$ , distribution for flare stars with *Gaia* DR2 matches. The green- and red-shaded regions represent approximations for the thin and thick disc, respectively (Gaia Collaboration 2018b). Right: Toomre diagram for flare stars with radial velocity information. The black, grey, and white dots indicate thin disc stars, thick disc candidates, and thick disc stars, respectively, according to the probability classes from Bensby, Feltzing & Lundström (2003). The green and red regions represent the expected  $v_{\text{total}}$  distributions for the thin and thick disc from (Bensby, Feltzing & Oey 2014). We can see that the majority of our flare stars are consistent with the thin disc distribution.

main-sequence star according to the criteria from Section 2.1. In total, 63 percent are consistent with being pre-main sequence and the remaining 5 percent either reside in the gap between main-sequence and pre-main-sequence sources (making them ‘non-main sequence’, a mixture of young and binary systems that cannot be easily distinguished) or are on the subgiant branch in Fig. 1.

When investigating the distance distribution of our flare stars, we identified an apparent excess in the number of observed flaring stars between 350 and 400 pc. These additional objects come from a single field (NG0531-0826) and are associated with the Orion Complex. The Orion Complex is vast collection of multiple clusters with many stellar populations (e.g. Kounkel et al. 2018). This sample of pre-main-sequence stars was analysed separately in Jackman et al. (2020) and underwent Spectral Energy Distribution (SED) fitting to account for the specific extinction and reddening properties from dust in the Orion region. We use the results of this SED fitting for these stars in the rest of this work. For reference these stars are depicted alongside the rest of the flare survey in Fig. 1.

To estimate the age of the remaining pre-main-sequence stars, we compared the pre-main-sequence cut-off defined in Section 2.1 ( $-0.753$  mag above the fiducial median absolute *Gaia* G magnitude) to the PARSEC isochronal models (Bressan et al. 2012). We found that, from comparison to the PARSEC isochrones, our imposed cut-off in absolute *Gaia* G magnitude results in a mass-dependence in the limiting age. This is because the time taken to reach the zero-age main sequence is mass dependent, taking longer for lower mass stars (Hayashi 1961; Bressan et al. 2012). We found that late K stars (K6-K8) designated as pre-main sequence are younger than approximately 15 Myr. For M0-M2 and M3-M5 stars, this increases to 30 and 40 Myr, respectively.

To probe the Galactic distribution of our flare stars, we have calculated their kinematics. For all stars with distance and proper motion information from a *Gaia* DR2 match, we calculate the

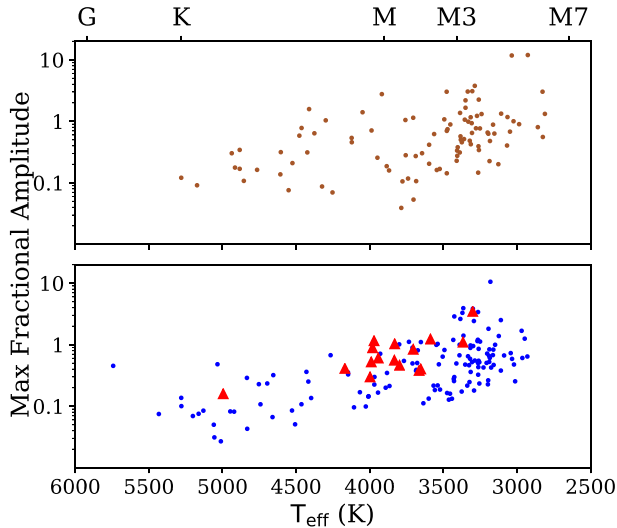
tangential velocity  $V_{\text{tan}}$ . *Gaia* DR2 also provides radial velocity measurements for selected stars brighter than  $G_{\text{RVS}} = 12$  (Katz et al. 2018). Consequently, where a measurement of the radial velocity is available from our matching with *Gaia* or RAVE, we also calculate the velocities relative to the LSR  $U_{\text{LSR}}, V_{\text{LSR}}, W_{\text{LSR}}$ .

We are able to calculate full 6D coordinates for 34 stars in our sample and tangential velocities for 309 stars. The distribution of the tangential velocities and kinematics for our sample is shown in Fig. 3. We use the kinematic selection criteria of Bensby et al. (2003) to calculate the thick-disc-thin-disc (TD/D) relative probability for each of the stars with a full 6D solution. We find that 29 of the 34 stars (85 per cent) with 6D kinematics in our sample are most likely to be thin disc members [ $P(\text{thick}) < 0.1 \times P(\text{thin})$ ]. Of the remaining stars, four are thick disc candidates [ $P(\text{thick}) > 0.1 \times P(\text{thin})$ ] and one is a most likely a member of the thick disc [ $P(\text{thick}) > 10 \times P(\text{thin})$ ]. This likely thick disc source is TYC 5181-713-1, a bright star located at the turn-off to the subgiant branch in Fig. 1. We believe this source is a subgiant, something we discuss further in Section 4.5.

When taking the  $V_{\text{tan}}$  sample separately, we again find the majority of stars reside within the regions expected for a thin disc. Thin disc members are typically characterized as younger than those of the thick disc (e.g. Bensby, Feltzing & Lundström 2004). Consequently, we might expect a greater number of flare stars from the thin disc, due to their relative youth (and increased activity).

In Fig. 3, we can see that two stars have  $V_{\text{tan}}$  greater than  $150 \text{ km s}^{-1}$ , the upper velocity limit for a thick disc. Investigating these stars further reveals that one is LHS 2802, previously classified by Reid & Gizis (2005) as an sdM4.5 subdwarf. This source can also be seen as an outlier from the main sequence in Fig. 1 and will be analysed further in Jackman et al. (in preparation). The other is a subgiant candidate NGTS J060634.4-250020. We discuss subgiant candidates in our sample in Section 4.5. While these stars have higher velocities than expected for the thick disc, they are not high enough





**Figure 4.** Maximum observed fractional flare amplitude against effective temperature. Top: Main-sequence isolated stars. Bottom: Non-main-sequence stars. Stars associated with Orion are shown in the bottom panel as the red triangles. Note how M stars of all ages can flare with amplitudes equivalent to 10 times the quiescent flux.

to be within the realm of possible halo stars ( $V_{\text{tan}} > 200 \text{ km s}^{-1}$ ; Gaia Collaboration 2018b).

### 3.2 Flare amplitude

We have calculated the amplitude for all flares in our sample, using the method described in Section 2.3. In the case of blended sources, this amplitude corresponds to a lower limit. Our measured fractional flare amplitudes range between 0.027 and 12 ( $\Delta m_{\text{NGTS}} = -2.8$ ), with our highest amplitude flare coming from an M6 star in our main-sequence sample. Fig. 4 shows the distribution of maximum observed flare amplitudes with effective temperature for the non-blended sources in our sample. As we progress to cooler stars in our sample we observe the expected increase in maximum flare amplitude due to the changing ‘contrast ratio’ between the flare blackbody and stellar photosphere within the NGTS filter.

Splitting our sample into main-sequence and non-main-sequence stars (which includes both binaries and pre-main-sequence stars) reveals that some sources in our main-sequence sample exhibit higher maximum flare amplitudes than their non-main-sequence counterparts of similar temperature. While these sources appear to be consistent with isolated main-sequence sources in the HR diagram, they may be binaries where one component has been spun up through tidal interaction and is magnetically active (e.g. Jackman et al. 2019c). Alternatively, the stars non-main sequence include binary stars. The extra flux from the second star in these systems will dilute the observed signal and result in a smaller measured flare amplitude. Spectroscopic follow-up of these sources will help determine whether this is the case. Of the non-main-sequence stars, we can see that the Orion-associated sample from Jackman et al. (2020) appears to flare with the largest maximum amplitudes, which is to be expected for their very young ages.

### 3.3 Flare energy

Using the method outlined in Section 2.4, we have measured the bolometric energy for the 428 flares from stars with available

effective temperatures and radii from TIC v8. These energies range between  $1.5 \times 10^{31}$  and  $3.4 \times 10^{36}$  erg. The stars with calculated flare energies have effective temperatures between 6170 and 2800 K, equivalent to between F8 and M6 spectral type.

#### 3.3.1 Flare occurrence rate

We fit equation (3) to the observed average flare occurrence rates for a series of bins in spectral type, shown in Table 2. We did this for both our main-sequence and pre-main-sequence samples. To calculate an average flare occurrence rate for a specific bin, we assumed an ergodic principle that all stars taken together were equivalent to a single, average, star. This average star was observed for the same duration as the sum of the observing times for all stars within that bin (e.g. 10 stars observed for 1 h were equal to one average star observed for 10 h). This method included both the flaring and the non-flaring stars, to avoid biasing our samples to more active stars and to obtain the true average behaviour. We used equation (3) and the method outlined in Section 2.5.1 to account for the decreased detection efficiency of low-energy flares from fainter stars relative to brighter stars of the same spectral type. If we did not account for this, then the average flare occurrence rate would be diminished at lower energies, where we can only detect flares from the brighter star. We note that one way around this, as mentioned in Section 2.5.1, would be to select a low-energy cut-off where we are confident we would detect flares from all the stars in our sample. However, due to the wide range of magnitudes of stars used in our sample ( $I \approx 9\text{--}16$ ), we found that only the highest energy flares would be detected on both the bright and faint stars. As these are the rarest flare events, this resulted in low numbers of flares to use when fitting our occurrence rates. By directly including the recovery fraction in our fitting, we were able to include both the high-energy flares from faint stars and low-energy flares from bright stars, maximizing our sample size.

To calculate the recovery fraction  $R(E)$  for our average star, we used the results of the injection and recovery tests described in Section 2.7. We first generated a grid of flare energies, spaced logarithmically from energies an order of magnitude below the lowest observed flare energy to an order of magnitude above the highest. For each star, we multiplied the value of its own recovery fraction at each energy in our grid by its observing duration to calculate a series of ‘effective durations’. This had the effect of weighting the contribution of each star at different energies, e.g. downweighting the contribution of a faint star at very low flare energies. To calculate  $R(E)$  for our average star, we summed the effective durations at each flare energy and then divided it by the total observing duration. We used this average  $R(E)$  in our fitting of the observed average flare occurrence rate.

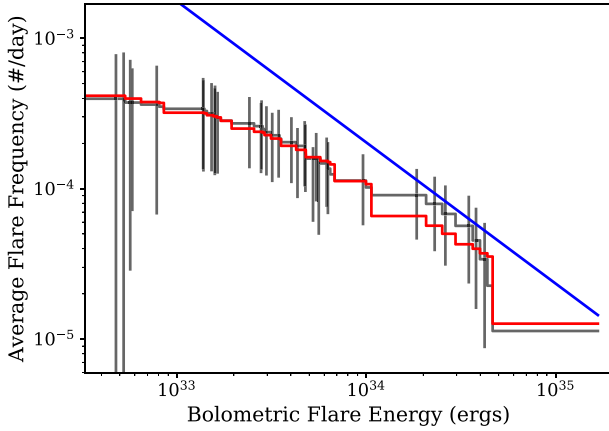
To fit equation (3) to our data, we used the EMCEE PYTHON package (Foreman-Mackey et al. 2013) to generate a Markov Chain Monte Carlo process. We used 32 walkers for 10 000 steps and used the final 2000 to sample the posterior distribution. To account for possible uncertainties in our injection and recovery of the smallest flares, we followed Ilin et al. (2019) and multiplied the uncertainties in our occurrence rate by  $R(E)^{-1/2}$ .

The results of our fitting and the marginalized 1D uncertainties are given in Table 2. Fig. 5 shows an example of the recovered average flare occurrence rate compared to the observed. We can see that as expected, the recovered rate is more frequent than the observed behaviour at all but the highest energies, where all flares are recovered. Fig. 6 shows the recovered rates for our main-sequence K and M stars together, while Fig. 7 shows the main-sequence

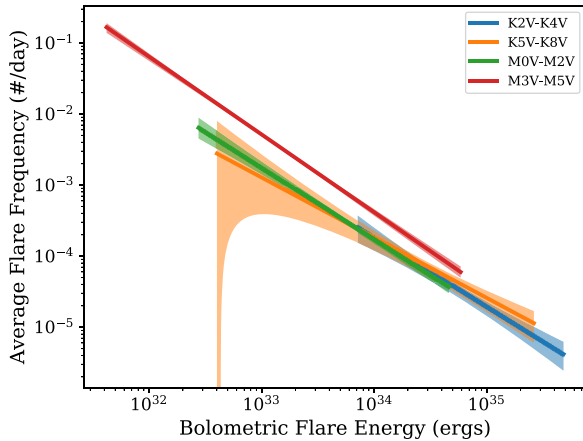


**Table 2.** Parameters from the power-law fits to main-sequence and pre-main-sequence subsets of our data.  $N_{\text{flares}}$  is the number of flares in each subset.  $N_{\text{stars,flare}}$  is the number of flaring stars in each subset, while  $N_{\text{stars,total}}$  is the total number of stars, both flaring and non-flaring. The uncertainties here are the marginalized 1D values and correspond to the uncertainties on the power-law fit for the average flare rate, rather than an explicit range of power laws in the input sample.

Class	$\log k$	$\alpha$	$C$	$\beta$	$N_{\text{flares}}$	$N_{\text{stars,flare}}$	$N_{\text{stars,total}}$
K2V-K4V	$29.4 \pm 7.2$	$1.97 \pm 0.20$	$29.4 \pm 7.1$	$-0.97 \pm 0.20$	18	17	15613
K5V-K8V	$25.0 \pm 7.4$	$1.85 \pm 0.21$	$25.0 \pm 7.3$	$-0.85 \pm 0.21$	18	14	6272
M0V-M2V	$30.4 \pm 3.4$	$2.01 \pm 0.10$	$30.4 \pm 3.4$	$-1.01 \pm 0.10$	34	26	4495
M3V-M5V	$33.9 \pm 1.6$	$2.09 \pm 0.05$	$33.8 \pm 1.5$	$-1.09 \pm 0.05$	91	39	626
K5-K8 (<15 Myr)	$25.4 \pm 4.8$	$1.82 \pm 0.14$	$25.5 \pm 4.8$	$-0.82 \pm 0.14$	19	14	405
M0-M2 (<30 Myr)	$20.9 \pm 1.8$	$1.69 \pm 0.05$	$21.0 \pm 1.7$	$-0.69 \pm 0.05$	39	33	705
M3-M5 (<40 Myr)	$29.5 \pm 1.2$	$1.94 \pm 0.04$	$29.5 \pm 1.2$	$-0.94 \pm 0.04$	121	60	311

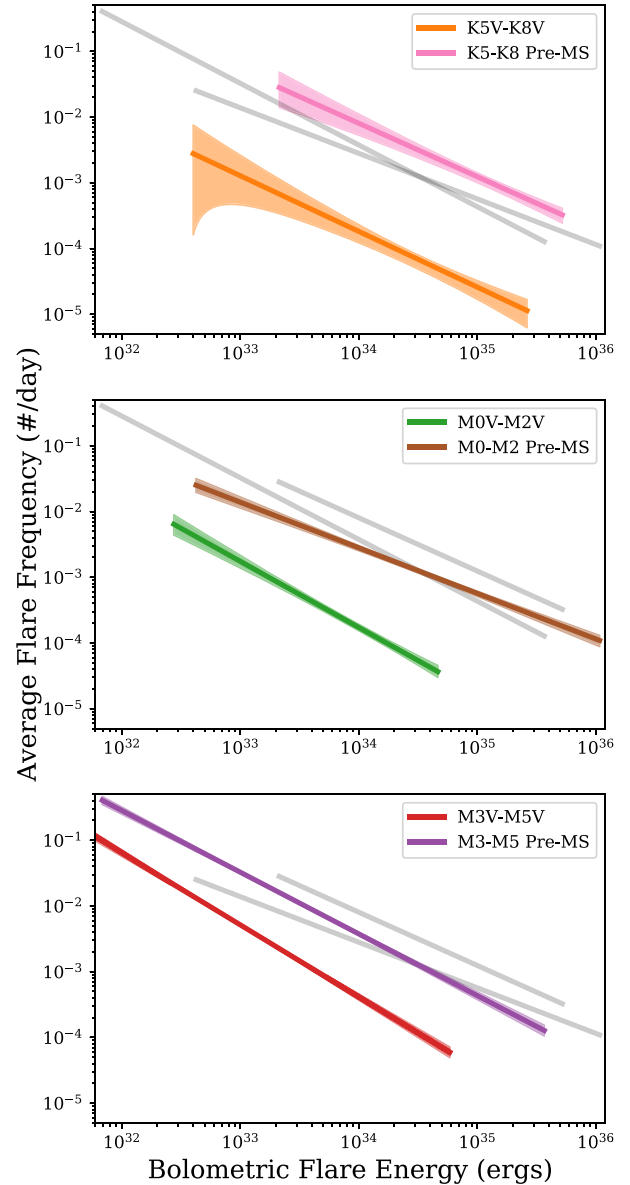


**Figure 5.** The results of our fitting of the average flare occurrence rate for M0V-M2V stars. The grey line show the observed occurrence rate, while the red line is our best-fitting model from using equation (1). The blue line is the power-law distribution required to give the red line and can be thought of as the intrinsic flare occurrence rate.



**Figure 6.** The best-fitting average flare occurrence rates for the main-sequence K and M subsets. The solid lines indicate the best-fitting power laws in the observed energy ranges, while the shaded areas are the  $1\sigma$  regions. The K and early M stars show similar flare rates, while the fully convective mid-M stars flare more frequently at all studied energies.

and pre-main-sequence samples together. We can see in Fig. 6 that the distributions for main-sequence K and early M stars overlap, suggesting that at field ages the average flare rate for partially convective stars does not have a strong mass dependence. We can



**Figure 7.** Average flare occurrence rates for the main-sequence and pre-main-sequence K5-K8, M0-M2, and M3-M5 subsets. The solid line is our best-fitting intrinsic flare model and the shaded areas are the corresponding  $1\sigma$  regions. The grey lines show the pre-main-sequence occurrence rates from the other panels for reference. Note the decrease in flare rate with age for all spectral types.

**Table 3.** Time in years for a flare of a given energy to occur on an average isolated main-sequence star in each of our late-type star classes. We can see that the mid-M stars flare more often on average in our sample than the early M stars. Note the pre-MS K stars are preferentially younger than their M star counterparts. We note that the isolated main-sequence sample does not include any of the stars in the pre-main-sequence sample and can be taken as representative of field age stars. The uncertainties correspond to the 16th and 84th percentiles propagated from our flare frequency power-law fits.

Class	$E > 10^{33}$ erg (yr)	$E > 10^{34}$ erg (yr)	$E > 10^{35}$ erg (yr)
Main sequence			
K2V-K4V	$1.7^{+2.1}_{-0.9}$	$15.2^{+7.4}_{-4.3}$	$144.9^{+43.2}_{-25.5}$
K5V-K8V	$2.2^{+2.6}_{-1.1}$	$15.9^{+6.5}_{-4.0}$	$112.7^{+44.4}_{-29.0}$
M0V-M2V	$1.6^{+0.4}_{-0.3}$	$15.9^{+2.0}_{-1.5}$	$160.8^{+50.9}_{-37.7}$
M3V-M5V	$0.5 \pm 0.1$	$6.6^{+0.8}_{-0.6}$	$82.4^{+20.8}_{-15.5}$
Pre-main sequence			
K5-K8 (<15 Myr)	$0.1 \pm 0.1$	$0.3^{+0.2}_{-0.1}$	$2.2^{+0.5}_{-0.3}$
M0-M2 (<30 Myr)	$0.2 \pm 0.1$	$1.0 \pm 0.1$	$4.8^{+0.5}_{-0.4}$
M3-M5 (<40 Myr)	$0.1 \pm 0.1$	$0.7 \pm 0.1$	$6.3^{+0.9}_{-0.7}$

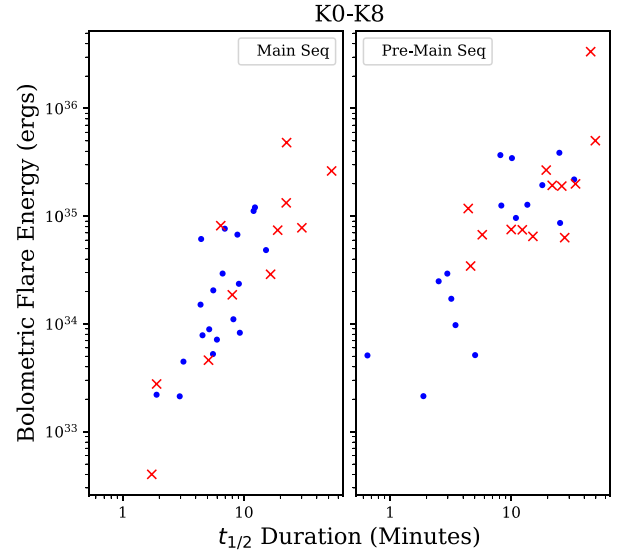
also see in Fig. 6 that fully convective mid-M stars are offset from the other distributions, showing a greater level of activity for field age stars compared to their higher mass counterparts. We believe this behaviour is due to the increased spin-down time and period of activity for fully convective stars relative to K and early M stars (e.g. Matt et al. 2015). We can see in Fig. 7 that the average flare occurrence rate for all spectral types has a clear dependence on age, with all pre-main-sequence samples flaring more often than their main-sequence counterparts. This, along with the behaviour of the main-sequence sample, is something we discuss further in Section 4.1.

For comparison with other works, we have also provided the values of  $C$  and  $\beta$ . During our analysis we noted that the uncertainties on  $k$  and  $\alpha$  were correlated and as such, the marginalized 1D uncertainties provided in Table 2 will not accurately reproduce our fitted power law distributions and associated uncertainties if used in future works. For this purpose, we have also provided the covariance matrices for each subset in Table A1.

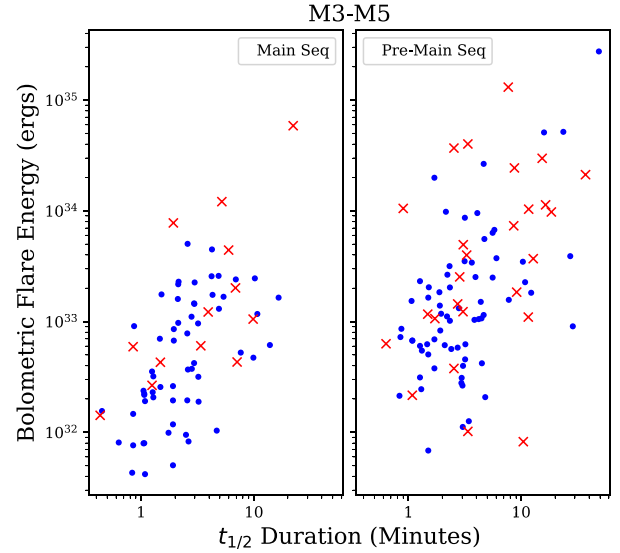
We have also used our fitted power laws to calculate the average waiting time of flares of different energies for main-sequence and pre-main-sequence K and M stars. These waiting times are shown in Table 3, showing that we expect a  $10^{35}$  erg flare to occur on a pre-main-sequence M star once every 5 yr.

### 3.4 Flare duration

We have measured the  $t_{1/2}$  time-scale of all observed flare events, using the high time cadence of NGTS. We measured  $t_{1/2}$  time-scales between 26 s and 50 min in our full sample. We have also split our sample into K, early M and mid-M spectral bins to see how the  $t_{1/2}$  time-scale varies with energy in each regime. Figs 8 and 9 show this distribution for our K and mid-M star samples. Our early M star sample sits between these distributions. We can see that both our K and mid-M star samples show a positive correlation between the  $t_{1/2}$  duration and measured flare energy in log-log space, however, there is notably more scatter in the mid-M samples. We measured the Pearson rank coefficients of each distribution in Figs 8 and 9, measuring 0.74 and 0.80 for the main-sequence and pre-main-sequence K star samples and 0.55 for both mid-M samples. At first glance, this would imply a strong relation between the flare energy and duration for K stars that breaks down towards later spectral types.



**Figure 8.** Duration and bolometric energy for flares detected from K stars in NGTS DR1. ‘Simple’ flares are shown as the blue circles, while ‘complex’ events are shown as the red crosses. Here, we have chosen the ‘shortest’ value of  $t_{1/2}$  as discussed in Section 2.6. The cut-off in detected flares for longer durations in each figure is consistent with a selection effect from our detection method. We can see that complex flares become more prevalent at longer durations and higher energies.



**Figure 9.** Duration and bolometric energy for flares detected from mid-M stars in NGTS DR1. Symbols are per Fig. 8.

However, we believe that the reason for this increased scatter at later spectral types is due to a greater number of detected lower amplitude, but longer duration, flares. These flares, due to their lower amplitudes are more easily detected on cooler stars due to the greater contrast between the flare and quiescent spectrum. To test this theory, we compared the observed distributions in Figs 8 and 9 to the results of our flare injection and recovery tests from Section 2.7. We found that the longer duration – lower energy cut-off in both the K and M star samples (and apparent linear relation) was consistent with a selection effect from our detection method, rather than a physical cut-off in  $t_{1/2}$  duration for low-energy flares.

### 3.5 Stellar rotation

We measured the periodicity in the light curves of our flaring stars in our sample using the method outlined in Section 2.8. From this, we detected and measured the periods of 102 stars in our sample. We vetted these stars to remove signals that may more likely be due to pulsation or eclipses than rotation. This left 97 stars, 36 of which were main sequence, 39 pre-main sequence, and 22 in between. Overall, these stars had spectral types between K0 and M5.5 and periods between 0.35 and 23.3 d. The main-sequence sample had spectral types ranging between K0 and M3.5 and periods between 0.35 and 22.3 d, while the pre-main-sequence sample had spectral types ranging between K0 and M5.5 and periods between 0.43 and 23.3 d. The sample in between the isolated main-sequence stars and the pre-main-sequence stars had periods between 0.2 and 12 d.

#### 3.5.1 Flares and star-spot phase

Star-spots are the result of local concentrations of the magnetic field that inhibit convection and create cooler regions of plasma on the photosphere. Correlations between spot groups and flares are frequently observed on the Sun (e.g. Zirin & Liggett 1987; Jiang et al. 2012; Yan et al. 2018), however, previous studies (e.g. Hawley et al. 2014; Doyle et al. 2018; Feinstein et al. 2020) have generally found no correlation between star-spot modulation and the timing of high-amplitude-stellar flares.

To search for any relation between flares and star-spot phase in our sample of spotted stars, we initially phase folded our light curves on the rotation period measured in Section 2.8. We then set all stars such that the minimum flux (the dominant star-spot or star-spot group is facing us) at phase 0.5. During this analysis, we assumed the modulation was due to a single large star-spot.

In order to avoid selection effects in our analysis, we required our flares to have fractional amplitude greater than the amplitude of variation. This is ensured we only used flares that would have been visible at all phases. To test the dependence of these large flares on the star-spot phase, we perform a single sample Kolmogorov–Smirnov (KS) test. We used a uniform distribution as our reference sample, based on the results of previous studies (e.g. Doyle et al. 2018). We tested individual stars in our sample that have more than three flares, along with testing the combined sets from Table 2. From our analysis no star, or combination of stars, seems to deviate significantly from a uniform distribution, consistent with previous results for large flares, and we discuss the reasons for this in Section 4.4.

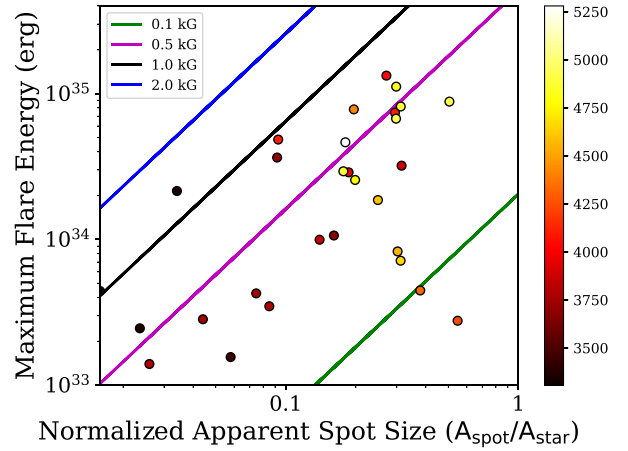
#### 3.5.2 Flares and star-spot size

Along with investigating how flares relate to star-spot phase, we can also see how flare energy is related to star-spot area. To do this we first calculate the star-spot area,  $A_{\text{spot}}$ , using the equation (e.g. Notsu et al. 2013)

$$\frac{A_{\text{spot}}}{A_{\text{star}}} = \left( \frac{\Delta F_{\text{spot}}}{F} \right) \left[ 1 - \left( \frac{T_{\text{spot}}}{T_{\text{star}}} \right)^4 \right]^{-1}, \quad (4)$$

where  $\frac{\Delta F_{\text{spot}}}{F}$  is the change in flux due to a spot normalized by the average flux and  $A_{\text{star}}$  is the apparent area of the star, equal to  $\pi R_s^2$ .  $T_{\text{star}}$  is the temperature of the stellar photosphere, which we have equated to  $T_{\text{eff}}$ . We calculate the star-spot temperature  $T_{\text{spot}}$  using the relation from Maehara et al. (2017):

$$\begin{aligned} \Delta T(T_{\text{star}}) &= T_{\text{star}} - T_{\text{spot}} \\ &= 3.58 \times 10^{-5} T_{\text{star}}^2 + 0.249 T_{\text{star}} - 808. \end{aligned} \quad (5)$$



**Figure 10.** Maximum observed flare energy against the normalized star-spot area. The coloured lines represent the maximum flare energy from a spot of a given magnetic field strength. The colour of the spot indicates its effective temperature. Note that all flares in this sample of stars with measured star-spot modulation can be explained with a field strength of 2 kG.

This relation was derived by Maehara et al. (2017) using a second-order polynomial fit to a compilation of stellar and spot temperatures from Berdyugina (2005), for stars with temperatures between 3300 and 5870 K. We also calculate the energy stored in the magnetic field,  $E_{\text{mag}}$ , using (e.g. Notsu et al. 2013)

$$E_{\text{mag}} \approx \frac{B^2}{8\pi} A_{\text{spot}}^{3/2}, \quad (6)$$

where  $B$  is the magnetic field strength of the star-spot. This represents the maximum flare energy from this star-spot, assuming all energy is converted from the magnetic field to the observed flare.

The results of our analysis are shown in Fig. 10. We can see in Fig. 10 that the majority of stars with detected flares and star-spot modulation have maximum observed flare energies that can be explained by a single large spot of 0.5 kG, across K and M stars alike. All flares and spots can be explained with a 2 kG field. We note the presence of some stars with large apparent spot sizes, reaching up to 60 per cent of the visible stellar surface. This can be explained by us overestimating the spot temperature in equation (5). Our magnetic field results agree with those of Howard et al. (2020b), who performed the same analysis using the EvryScope flare sample for cool stars. Our results are also consistent with previous observations of kilogauss magnetic field strengths for K and M stars (e.g. Saar, Linsky & Beckers 1986; Shulyak et al. 2017). We have assumed in this analysis that the spots are located at the equator. Previous studies (e.g. Notsu et al. 2019) have noted that the position of the lines in Fig. 10 will shift upwards for stars with lower inclination angles, or for spots at higher latitudes (e.g. near the pole) that cause smaller brightness variations. Consequently, those flares lying above the 1 and 2 kG lines may in fact be consistent with lower magnetic field strengths if their spots are located at high latitudes or the stars have low inclinations.

## 4 DISCUSSION

We have presented the results of searching for stellar flares in the first NGTS data release. We have identified 610 flares from 339 stars at a 13-s time cadence. We have been able to find flares from redder stars than the main *Kepler* survey was able to do and are currently probing a similar regime to *TESS*, but with a higher time resolution for all stars.



## 4.1 Flare occurrence rates

### 4.1.1 Main sequence

In Section 2.5.1, we measured the average flare occurrence rates for the K and M stars in our sample, the values of which are shown in Table 2. The best-fitting power laws of main-sequence stars are shown in Fig. 6. We can see in Fig. 6 that the average flare occurrence rates of K and early M stars overlap, while the fully convective mid-M stars are offset, showing an increased level of activity. As mentioned in Section 2.5.1, we believe that this offset is due to the increased spin-down time of fully convective stars relative to K and early M stars (e.g. Matt et al. 2015).

By  $\approx 650$  Myr isolated K stars have mostly spun down (Douglas et al. 2016), meaning our field age sample will likely be dominated by slow rotating stars. Newton et al. (2017) measured the rotation periods of field age M dwarfs with MEarth photometry. The majority of early M stars in their sample have periods longer than 10 d, showing they have mostly spun down. K and early M stars are believed to share a common dynamo. The strength of this dynamo and the relative magnetic activity (e.g.  $L_x/L_{\text{bol}}$ ) of field stars is known to be tied to the rotation period, with stars with periods above 10 d showing reduced activity (e.g. Pizzolato et al. 2003; Wright et al. 2011). Consequently, the field age K and early M stars having spun down to the same rotational period (and magnetic activity) distribution would explain our observed common average flare rates.

In contrast to early M stars, mid-M stars show significant scatter in their period distribution, showing that while some stars have spun down, many have not. Studies of the evolution of the quiescent magnetic activity of early and mid-M stars using ultraviolet (UV) emission have found results reflecting this behaviour. In a study of the near-ultraviolet (NUV) and far-ultraviolet (FUV) emission of open cluster and field age M dwarfs, Schneider & Shkolnik (2018) found that the UV emission of early M stars dropped sharply between  $\approx 150$  Myr and field ages (by a factor of 31 in the FUV), while the emission from mid-M stars retained high levels of UV activity into field ages. Therefore, while K and early M stars in our sample have on average spun down and dropped in magnetic activity, the longer spin-down of mid-M stars helps to sustain their average flaring behaviour into field ages. This is also in agreement with West et al. (2008), who measured the fraction of active M dwarfs, using  $H\alpha$  luminosity, as a function of spectral type and combined this with dynamical simulations to derive age-activity relations for M dwarfs. They found that the active fractions and activity lifetimes of mid-late M dwarfs were greater than for early M dwarfs, making it likely that our observed increased flare rates of mid-M stars is due to an extended activity lifetime.

Medina et al. (2020) recently studied the white-light flaring behaviour and rotation periods of nearby field age mid-M dwarfs using *TESS* short-cadence data and MEarth photometry. They found that these field stars fell into two distributions in flaring activity-rotation space – fast rotating and regularly flaring, and slow rotating ( $\approx 100$  d) and rarely flaring. This is consistent with some mid-M stars having spun down and decreased in flaring activity, while others have not. We note here that in order to measure the average white-light flare behaviour we will have combined stars from both regimes highlighted by Medina et al. (2020). Therefore, individual mid-M stars may show activity above or below this behaviour, depending on the regime they fall into. We intend to measure the average occurrence rate in each regime in future work.

In Table 3, we have used our fitted occurrence rates to calculate the average waiting time for flares with energies above  $10^{33}$ ,  $10^{34}$ , and  $10^{35}$  erg. We can see from Table 3 that on average an isolated

main-sequence M0-M2 star will flare with an energy of at least  $10^{34}$  erg once every 16 yr and  $10^{35}$  erg roughly once every 160 yr. Mid-M stars show a flare of  $10^{34}$  erg once every 7 yr and  $10^{35}$  erg roughly once every 82 yr. Our K and early M stars subsets have similar waiting times, especially for energies up to  $10^{34}$  erg, namely one flare every 15 yr. This illustrates our finding that the K and early M stars have spun down to a common state of activity. We note that what we have calculated is the average flare rates and waiting time, by combining large samples of active and inactive stars within chosen ranges of age and spectral type. The uncertainties given in Table 2 and 3 correspond to the error on the calculated average flare rate for the average star, rather than an explicit range of power laws and waiting times for every star in the sample. Therefore, we note that while these values can be used for the flare waiting time of individual stars, they are best suited for predicting the number of flares from large samples of stars and that some stars of interest (e.g. very active M stars) may show activity and values of  $C$  outside our calculated ranges.

Our measured values of  $\alpha$  for main-sequence M and late-K stars are consistent with those measured in previous surveys using ground and space-based observations (e.g. Hilton et al. 2011; Yang et al. 2017; Lu et al. 2019; Howard et al. 2019). A value of  $\alpha$  greater than 2 implies that, if the distribution given in equation (2) holds at all energies, then the lowest energy flares can dominate the total energy distribution (e.g. Güdel et al. 2003), possibly providing a way of heating the Solar and stellar corona through frequent nanoflares (e.g. Doyle & Butler 1985). Our measured values of  $\alpha$  for the K and early M main-sequence subsets are all consistent with  $\alpha = 2$ . Interestingly, our best-fitting power law for the fully convective mid-M stars has  $\alpha > 2$ , although it is consistent with  $\alpha = 2$  within  $2\sigma$ . Our fitted  $\alpha$  values suggest that for K and early M stars large flares dominate the total energy distribution, however, our uncertainties do not allow us to confidently rule out the alternative, while our results suggest that small flares dominate the distribution for the average mid-M star.

### 4.1.2 Pre-main-sequence stars

We have also measured the average flare occurrence rates for pre-main-sequence stars in our sample, the results of which are shown in Table 2. We defined our pre-main-sequence sample in Section 2.1 using the positions of the stars on the HR diagram in Fig. 1. As we discussed in Section 3.1.2 the limit for stars to be designated as pre-main sequence had a mass-dependent age cut-off. Stars with a K6-K8 spectral type are younger than 15 Myr, M0-M2 are younger than 30 Myr, and M3-M5 are younger than 40 Myr. Along with this, we note that these samples may also be contaminated by young (but older than the pre-main-sequence samples) binaries that are shifted upwards on the HR diagram. We initially used the PARSEC isochrones to limit our pre-main-sequence M star sample to 15 Myr, however, this resulted in too few flares to fit a meaningful power law. Consequently, we combine all pre-main-sequence stars in each subset. As these pre-main-sequence samples are a mixture of ages, we have not used our fitted values here to quantitatively probe how stellar flare activity changes with age as we would for an open cluster (e.g. Ilin et al. 2019; Ilin et al. 2021; Jackman et al. 2020). Instead, we use this sample for comparison with our main-sequence data.

We have plotted our the average flare occurrence rates for our main-sequence and pre-main-sequence late-K and M star samples in Fig. 7. We can see that, as expected, the pre-main-sequence stars show a greater level of magnetic activity than their main-sequence counterparts, flaring more often and with greater energies. As stars age they undergo magnetic braking (e.g. Booth et al. 2017) that carries

away angular momentum, reducing differential rotation and their dynamo and magnetic activity. This in turn is expected to reduce flaring rates, as we can see in Fig. 7. This also broadly matches the behaviour of flares in the FUV. Loyd et al. (2018a), in their study of the FUV flaring activity of 40 Myr and magnetically inactive field age M stars from *HST* observations, found that for FUV flare energies above  $10^{30}$  erg, young M stars flare on average 50 times more often than their field age counterparts. However, we find smaller values for the lowest energy flares in our sample, which would be most comparable to those studied by Loyd et al. (2018a), with young M stars flaring 5–10 times as much as their main-sequence counterparts. One possibility for this may be because Loyd et al. (2018a) used known magnetically inactive and isolated field-age M stars from Loyd et al. (2018b) to compare to their young sample, whereas our field age sample contains both inactive and active M stars, some of which may not have fully spun down, as discussed in Section 4.1.1. This could act to increase the observed gap in flare rates. Further vetting and classification of the quiescent activity of stars in future flare studies will help test whether this is the case, or whether our findings point towards a true discrepancy between the change in flare rates between the optical and the FUV.

We have calculated the average waiting times for flares of various energies from our pre-main-sequence samples, shown in Table 3. We find that, on average, late K stars with ages below 15 Myr flare with energy of  $10^{34}$  erg approximately 50 times more often than for field age stars. For the same flare energy, the difference between field age M3–M5 stars and those with ages below 40 Myr is a factor of 9. While these pre-main-sequence M3–M5 stars are older than the K5–K8 sample, the reduced gap in flare rates between young and field age stars highlights how the magnetic activity persists with age for the fully convective sample. As we discussed in Section 4.1.1, this is likely due to the longer spin-down time-scales of fully convective stars.

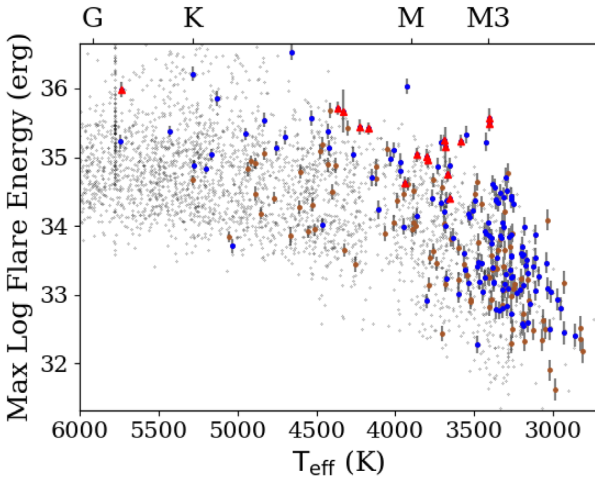
The best-fitting values of  $\alpha$  for our pre-main-sequence samples are shown in Table 2. Our value of late K stars is consistent with our measured value for the field age sample. However, we can see that both M star samples appear to have lower values of  $\alpha$  at young ages, which points towards a steepening of the average flare occurrence rate of M stars with age. The change of  $\alpha$  with age has been studied by various groups. From their study of ultracool dwarfs with different kinematic ages, Paudel et al. (2018) found evidence for an increase in the average  $\alpha$  with age, seemingly in agreement with our finding. However, the authors were unable to conclude any definite relation due to their small sample size. From a study of nearby stars and stars in open clusters, Shakhovskaia (1989) highlighted a dependence of  $\alpha$  on age. Their observations showed that  $\alpha$  appeared to decrease with age, in contrast with our results. In their study of flares from the Pleiades and Praesepe open clusters, Ilin et al. (2019) found no strict age dependence for  $\alpha$  but did appear to observe a decrease in  $\alpha$  between the clusters for early M stars. Ilin et al. (2019) discussed the evolution towards a shallower flare rate with age (decreasing  $\alpha$ ), suggesting that the magnetic field topology changes with age, in turn allowing longer build ups of magnetic stress and energy. This then results in fewer small flares and more large ones. Our results would suggest the opposite, that a change in the field topology instead acts in favour of smaller flares. Very young early M stars are expected to be fully convective, forming a radiative core after approximately 100 Myr (e.g. Baraffe et al. 2015). Fully convective M stars are believed to predominantly have strong dipolar poloidal fields (e.g. Morin et al. 2008, 2010; Shulyak et al. 2017), while partially convective stars have multipolar fields with a large scale toroidal element. For Solar flares, it is the toroidal fields that supply the

energy and the same is typically considered for flares from partially convective M dwarfs (e.g. Mullan & Bais 2018). However, in a study of the magnetic field and flares from the M8 planet host TRAPPIST-1, Mullan et al. (2018) noted that while a toroidal field would dominate small flares, the poloidal field may contribute in part to higher energy flares with a different power law. If this is the case, then young early M dwarfs (which are fully convective) may have a dipolar field that is more preferable for high-energy flares. As the star ages and the radiative zone and the interface dynamo are formed, a multipolar field is generated which would reduce the field contribution to high-energy flares and results in an apparent preference towards smaller flares. As the radiative zone for early M stars is formed by 100 Myr, this could explain why this behaviour was not observed in the 125 and 650 Myr samples of Ilin et al. (2019). However, this does not explain why the value of  $\alpha$  for our mid-M stars, which remain fully convective their entire lives, also appears to increase with age in our sample. This would point towards other elements at work, not considered here. Along with this, it does not explain why the opposite behaviour in  $\alpha$  with age was observed by Shakhovskaia (1989). On a similar note, Ilin et al. (2020) recently compiled measurements of  $\alpha$  from the literature for young and field age stars, finding a spread of measured values from 1.4 to 2.5. They noted that this spread of measured values, with various ways of determining the uncertainty, currently makes it difficult to assess whether  $\alpha$  truly is changing with mass and age. Future studies of flares from open clusters should aim to address the evolution of  $\alpha$  in a systematic way to tackle this problem.

Davenport et al. (2019) has investigated how the flare occurrence rate varies as a function of age by applying gyrochronology techniques to the Davenport (2016) *Kepler* sample. Using rotation period as a proxy for stellar age, they found that flaring rates decrease as stars age and they generated flare evolution models which vary as a function of age and stellar mass. Using this model for a  $0.5 M_{\odot}$  star flaring at an energy  $10^{35}$  erg gives a best-fitting age around 2 Gyr for the young samples in Fig. 7. This is not consistent with their position on the HR diagram, suggesting the model from Davenport et al. (2019) is overestimating the flare rate for these stars. Both Jackman et al. (2020) and Ilin et al. (2020) found similar discrepancies of this model with the measured average flare occurrence rates of low-mass stars in open clusters. We believe the overestimation of the flare rates is because the Davenport et al. (2019) model only used active flare stars as an input, whereas we have used both the flaring and non-flaring stars in our sample. Therefore, we note that this model is most applicable to active subsets and should not be used to model the average flaring behaviour of groups of stars.

#### 4.1.3 Refitting the Orion occurrence rate

We have also applied our new method to the average occurrence rate of Orion-associated 3400–3940 K (approximately M0–M3.5) stars from Jackman et al. (2020). In their analysis, Jackman et al. (2020) fit a power law according to equation (2) to flares above an energy of  $9 \times 10^{34}$  erg. Only stars with a 68 per cent completeness limit below this value, determined from injection and recovery tests, were used in the fitting of the average flare occurrence rate. This limited the number of flares and stars that could be used in the analysis, with 11 of 40 stars being excluded from the analysis. The best-fitting power law from Jackman et al. (2020) had  $\alpha = 2.63 \pm 0.62$  and  $C = 57.4 \pm 21.9$ . We note that  $C$  was measured in units of  $\text{year}^{-1}$  and in units of  $\text{day}^{-1}$  is  $C = 54.8 \pm 21.9$ . We ran our analysis using the flare injection and recovery results from Section 2.7. We measured



**Figure 11.** Maximum observed bolometric flare energy against effective temperature for our detected flaring stars. Stars consistent with being isolated and on the main sequence are shown as the brown circles. Stars not consistent with being on the main sequence, but with no known association, are shown as the blue circles. Stars associated with the Orion complex from Jackman et al. (2020) are shown as the red triangles. The underlying black points represent stars from Yang & Liu (2019), using long cadence *Kepler* observations. We believe the line of stars around 5800 K is due to an error in the KIC effective temperatures from the Yang & Liu (2019) sample. Note how the non-main-sequence stars appear towards higher energies.

a best-fitting power law with  $\alpha = 1.94 \pm 0.16$  and  $C = 30.7 \pm 5.7$ . These values are consistent with the previously measured values to well within  $2\sigma$ , but have smaller measured uncertainties. The inclusion of lower energy flares in our fitting is likely responsible for the decrease in our measured  $\alpha$ , while the decrease in the measured value of  $C$  is due to both the lower value of  $\alpha$  and the inclusion of stars which had 68 per cent completeness limits below the specified energy of  $9 \times 10^{34}$  erg. Many of these stars will have contributed non-negligible amounts to our calculated observing duration, bringing the measured rate down.

#### 4.2 Maximum flare energy

In Section 2.4, we calculated the energies of the flares in our sample, finding them to range between  $1.5 \times 10^{31}$  and  $3.4 \times 10^{36}$  erg. One aspect of interest in our study is whether we can use our sample to probe the maximum limit of flare energy. Fig. 11 shows the distribution of maximum measured flare energy with effective temperature for our sample. As we progress to cooler stars we see a decrease in the maximum flare energy, similar to what has been seen in other surveys of maximum flare energy (e.g. Davenport 2016). We can see that field age K and early M stars can flare with energies up to approximately  $10^{35}$ , which is 10 000 times that of the Carrington event on the Sun (Carrington 1859; Tsurutani et al. 2003). In Fig. 11, we have plotted the maximum flare energies from the *Kepler* long cadence sample of Yang & Liu (2019). The two samples seem to overlap in the energy space covered, in particular at the upper limit of flare energies for M dwarfs.

Recent observations of white light flares from M dwarfs with ASAS-SN have recorded V-band flare energies up to  $10^{36}$  erg and corresponding bolometric energies up to  $10^{37}$  erg (Schmidt et al. 2019; Rodríguez Martínez et al. 2020), beyond the apparent upper energy limit seen in Fig. 11. This implies that our observed upper limit is not due to a physical limitation in the amount of energy that

can be stored and released through magnetic reconnection, but due to a selection effect from our observing strategy. To test and confirm this, we used the best-fitting average flare occurrence rates measured in Section 3.3.1. For each subset, we applied our recovery fraction  $R(E)$  to the intrinsic power-law distribution using equation (1) to predict the observed flare distribution. We used this predicted flare distribution to calculate the maximum energy at which we should have seen one flare in our observations. We then compared this value to our measured energies for the subset. If the maximum energy distribution cuts off before this limit, then the cut-off would likely be due to a physical limitation. If the limit is below our observed maximum flare energies, then the apparent cut-off in our distribution is due to a selection effect, namely not having observed each subset for enough time. Applying this analysis to the main-sequence stars in Fig. 11 reveals that the upper energy limit in our sample of K and early to mid-M stars is indeed consistent with not having observed each subset for long enough, rather than a physical limit in the energy that can be released from flares. We predict that a main-sequence M0V-M2V or M3V-M5V star will flare with an energy of  $10^{37}$  erg approximately once every 17 000 and 12 500 yr, respectively. This drops to around 80 yr for our pre-main-sequence samples, for both spectral types. These long waiting times for individual stars highlight how the large sample sizes achieved with wide-field surveys are vital for studying the highest energy flare events.

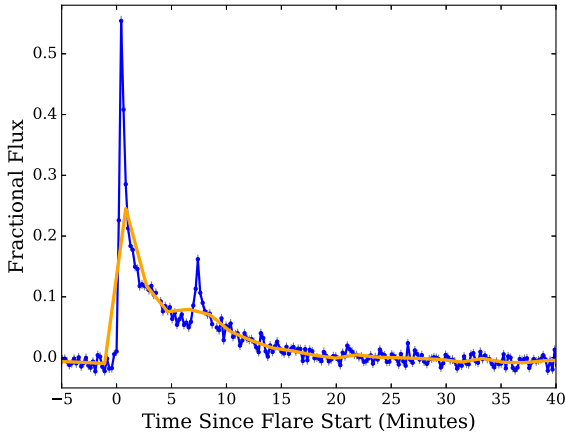
Splitting our sample of maximum flare energies into those consistent with being isolated main-sequence stars and those which are more likely to be binary or pre-main-sequence reveals that the non-main-sequence sources on average occupy regions of higher energy. This is expected as young stars are known to be more magnetically active than their older counterparts. This activity results in an increased rate of high-energy flares, making it more likely we would observe higher energies in our sample. Binary sources in this sample may have undergone tidal interactions that have prolonged this phase of strong magnetic activity, contributing to the increased maximum flare energy.

#### 4.3 Flare duration

In Section 3.4, we presented the range of flare  $t_{1/2}$  durations for our sample. The results for our K and mid-M stars are shown in Figs 8 and 9. In Section 3.4, we discussed that while there appears to be a strong positive correlation between the flare energy and the  $t_{1/2}$  duration for K stars that breaks down for later spectral types. However, after testing this distribution using our flare injection and recovery tests, we found that this behaviour was likely due to a selection effect. Something else we can see in Figs 8 and 9 is that complex flares appear to form the majority of the longest duration events. Long time-scale complex flares (particularly those with sharp multiple peaks) have previously been explained as multiple simple flare events (e.g. Davenport et al. 2014) occurring in quick succession. In this scenario, it is possible the successive flare events are a result of ‘sympathetic flaring’, where one large flare can trigger events in the same or nearby active regions, giving the appearance of multiple peaks in a single light curve. This trigger is provided through a form of physical link between flare sites. Multiple mechanisms (e.g. a direct link, a travelling disturbance, or the weakening/removal of magnetic fields elsewhere; Wang et al. 2001; Jiang et al. 2009; Zuccarello et al. 2009) have been proposed as the cause for this trigger, with the exact cause still up for debate. Therefore, it may be that our high energy complex flares with long durations are in fact formed of several simple flares.

From Figs 8 and 9 it also appears that, somewhat counter-intuitively, there are complex flares at the shortest  $t_{1/2}$  values. These





**Figure 12.** Example of a multi-peaked complex flare with a very short  $t_{1/2}$  duration ( $\approx 30$  s). NGTS data are shown by the blue line. Overlaid in orange is the simulated *TESS* 2-min short cadence mode showing the decrease in flare amplitude and resolved substructure.

flares are impulsive events with high-amplitude and short-duration peaks, followed by secondary smaller peaks in a long decay that do not reach the  $t_{1/2}$  limit. Therefore, when using the  $t_{1/2}$  value (which is primarily a measure of the width of the flare peak), these complex flares can appear very short in length. An example of such a complex flare is shown in Fig. 12. We have also presented how this flare would appear in *TESS* 2-min cadence observations, showing how longer cadence observations can smear out flare substructure and reduce the measured amplitude (e.g. Yang et al. 2018).

#### 4.4 Flares and star-spot phase

In Section 3.5.1, we used our sample of stars with detected flares and measured rotation periods to investigate whether the flare and star-spot phases were related. We found that, similar to results from previous studies, flares occurred with a uniform distribution in star-spot phase. This uniform distribution of stellar flares with star-spot phase has been explained in previous works through various ways. One possibility is that we are observing active regions that are permanently visible, or visible the majority of the time. This can be due to the stellar inclination relative to us, high-altitude spots, or a combination of the two. In this scenario, flares could still be associated with active regions but appear uniform in phase, while the observed flux modulation would arise from changing visible fractions of the spots.

An alternative explanation is that we are observing multiple active regions, with the observed light-curve modulation being due to their residual behaviour (e.g. more spots on one side Rackham, Apai & Giampapa 2018). This would mean a star-spot is effectively always in view, providing a relatively uniform distribution of active region associated flares.

Due to the limits imposed on the flare amplitude, we are also currently only probing the behaviour of the highest energy events. Due to the intrinsic rarity of such events, it is possible that more subtle variations are present that we are not sensitive to. This regime was explored by Roettenbacher & Vida (2018) using long cadence *Kepler* observations of 111 main-sequence stars. They found that flares with amplitudes above 1 per cent occurred more often when a single large star-spot group was in view, suggesting a spot-flare relation was present for low-amplitude flares. Interestingly, they found a similar but less pronounced behaviour for a sample of 425 flares

**Table 4.** Parameters of flaring subgiant candidates. N/A indicates no rotation period was measured.

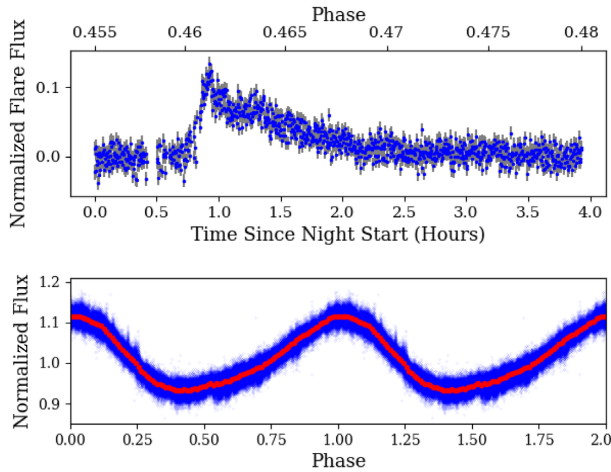
Object	Period (d)	Amplitude (per cent)
NGTS J060634.4-250020	1.6	4.8
NGTS J061739.4-651709	6.8	18.3
NGTS J134437.3-334737	N/A	N/A
NGTS J141137.1-255427	N/A	N/A
NGTS J21440.3-393121	N/A	N/A
TYC 5181-713-1	10.8	7.7
UCAC2 16087743	N/A	N/A

with amplitudes above 5 per cent, which we have not identified. This is likely due to our smaller sample of 82 flares from main-sequence K and M stars with measured rotation periods. More recently, Feinstein et al. (2020) explored both flare amplitude regimes using *TESS* short cadence observations of young stars. They found that for young stars, flares of all amplitudes has no correlation with the rotational modulation phase. This result is consistent with our observations for our pre-main-sequence sample.

#### 4.5 White-light flares from subgiants

In Section 3.1.2, we noted that one of our flaring sources was likely associated with the Galactic thick disc and was located at the turn-off to the subgiant branch on the HR diagram in Fig. 1. Subgiant stars have begun to evolve off of the main sequence and often show signs of variability. This variability can be due to tidally induced rotation (e.g. RS CVn) or pulsations (e.g. Delta Scuti stars) and can sometimes be accompanied by flares. Following our analysis of TYC 5181-713-1, we searched for other subgiant candidates within our flare star sample. We did this by selecting stars from our flare star sample based on their position on the HR diagram, specifically around the subgiant turn-off from the main sequence, with  $T_{\text{eff}}$  between 4700 and 5600 K and an absolute *Gaia* G magnitude less than 4. We identified six more sources in this region near this turn-off in Fig. 1. These objects are listed in Table 4, along with their measured periods and modulation amplitudes from our analysis in Section 2.8.

We performed a literature search for each object. Of note was NGTS J061739.4-651709 (2MASS J06173939-6517097), which has previously been classified by Watson, Henden & Price (2006) and Chen et al. (2018) as a Cepheid variable star using AAVSO and *WISE* photometry, respectively. The flare and phase-folded light curve of NGTS J061739.4-651709 is shown in Fig. 13. Cepheids are pulsating stars for which the period of pulsation is correlated to their luminosities (Leavitt & Pickering 1912). The presence of a white-light flare on a Cepheid variable would be of interest, due to the current debate about the causes of detected sudden UV and X-ray outbursts near their minimum and maximum radii, and whether this could be due to magnetic reconnection events (e.g. Engle et al. 2017; Moschou et al. 2020). In particular, we note that the flare from NGTS J061739.4-651709 occurs at the same phase as the X-ray flares observed by Engle et al. (2017), seemingly supporting the idea that these outbursts are due to magnetic reconnection events. However, Jayasinghe et al. (2018) used ASAS-SN photometry to classify the variability of NGTS J061739.4-651709 as due to rotation, rather than Cepheid-like variability. Further inspection of the NGTS light curve, shown in Fig. 13 shows that while the light curve does show asymmetric modulation, it is not consistent with the fast rise, slow decay expected for Cepheids, and is more likely due to rotation.



**Figure 13.** Top panel: The observed flare from NGTS J061739.4-651709 (2MASS J06173939-6517097). The top axis shows the modulation phase. Bottom panel: The phase folded light curve of NGTS J061739.4-651709, showing how the modulation is likely due to rotation rather than Cepheid-like variability. The blue points are the phase-folded NGTS light curve, and the red points are the same data binned to 100 points in phase.

From our literature search, we found that UCAC2 16087743 was identified as a spectroscopic binary candidate from RAVE data by Birko et al. (2019) and Jack (2019). It is possible that this source is an RS CVn. RS CVns are variable tidally locked close-binary stars in which one component can be a subgiant. RS CVns have previously been observed to show both flaring behaviour and modulation due to star-spots, if one of the components is a K or M dwarf (e.g. II Peg; Frasca et al. 2008). It is also possible that the remaining subgiant candidates are also RS CVns. However, we note subgiants can have a wide range of periods and amplitudes of variability, making it difficult to classify these sources based on this information alone (see Gaia Collaboration 2019, for a list of variables and associated periods).

#### 4.6 Role of flares in exoplanet habitability

In recent years, stellar flares have taken an increasingly important role in the discussions around exoplanet habitability. This is in part due to the discoveries of the Earth-sized exoplanets in or near the ‘habitable zones’ of M stars (e.g. Dittmann et al. 2017; Bonfils et al. 2018). Due to the cool temperatures of these M stars, their habitable zones reside at a close proximity to the host star, exaggerating the effects of flares. Detrimental effects from flares discussed in the literature include atmospheric erosion and ozone depletion due to flare-related X-ray and UV (XUV) irradiation (Segura et al. 2010; Tilley et al. 2019) and from bombardments of charged particles from associated coronal mass ejections (e.g. Lammer et al. 2007; Kay, Opher & Kornbluth 2016).

On the other hand, the UV irradiation from flares has been noted as a possible way of facilitating prebiotic chemistry, which may require quiescent NUV flux not available from cool M stars. For active stars, the UV irradiation could be provided by regularly flaring activity, with photochemistry occurring during flares and halting during stellar quiescence (Ranjan, Wordworth & Sasselov 2017; Rimmer et al. 2018). For tidally locked planets, Mullan & Bais (2018) have postulated that this activity could provide an analogue to the day/night cycle and a yearly cycle could even be simulated through stellar activity cycles. Through studying the rates of UV chemistry

and biomolecular reactions, Rimmer et al. (2018) identified minimum flaring limits required to bring a system into the ‘abiogenesis zone’. From this work, they found that this mechanism may only be viable for the most active stars.

To test whether the average flare rates measured in this work would be enough to reach the ‘abiogenesis zone’, we compared our best-fitting flare occurrence rates from Section 3.3.1 to the requirements from Rimmer et al. (2018), following the method outlined by Günther et al. (2020). We found that none of our average flare rates, for main-sequence or pre-main-sequence stars is enough to reach the abiogenesis zone and provide the required NUV flux. We found that a factor of approximately 2500 and 580 were required to bring the flare rates for main-sequence and pre-main-sequence mid-M dwarfs into the abiogenesis region. This changed to 66 000 and 3600 for main-sequence and pre-main-sequence early M stars. Spectroscopic studies have shown that the 9000 K flare model can often underestimate the flux in the *U* and NUV regions, due to not including the Balmer lines and the associated jump (Kowalski et al. 2013, 2019). However, simultaneous *HST* NUV and ground based optical spectroscopy observations have only measured the discrepancy to be a factor of two (Kowalski et al. 2019) and cannot provide the required NUV flux. This suggests then, at least for an average star, that flares cannot provide the required NUV flux, as determined by Rimmer et al. (2018), required for surface abiogenesis on habitable zone planets around M stars. However, we cannot rule out the possibility of individual highly active stars reaching this region (e.g. Günther et al. 2020).

## 5 CONCLUSIONS

We have presented the results of a search for stellar flares from the 196 600 stars in the first 24 fields observed with the NGTS. We identified 610 flares from 339 stars. These flaring stars have measured spectral types between F8 and M6. We find the majority of our flares come from stars that young or in binary systems. We have combined our flare observations with *Gaia* DR2 to identify that the majority of our sources have kinematics consistent with the Galactic thin disc. We have identified seven likely subgiants that showed flaring activity.

We have used the NGTS observations of flaring and non-flaring stars to measure the average flare occurrence rates for subsets of K and M stars. We presented a method to measure flare occurrence rates while accounting for how the detection efficiency changes with energy, removing the need to fit only to the highest energy flare events. We found that field age main-sequence K and early M stars showed similar average white-light flare occurrence rates, while fully convective field age mid-M stars were more active. We believe this is due to the increased spin-down time-scale of fully convective M stars, which causes them to remain active into field ages. We have compared the average flare occurrence rates for main-sequence and pre-main-sequence stars. We found that, as expected, for all spectral types pre-main-sequence stars flare more often.

We measured the rotation period of 105 stars in our sample which showed visible star-spot modulation. We compared the star-spot phase to flare occurrence and found no correlation between the two, similar to previous studies, and discussed possible reasons for this. We also used the observed light-curve modulation to estimate star-spot size and combined this with the measured flare energies to estimate the magnetic field strengths for stars in our sample. We found that all flares and spots for these stars could be explained using a 2 kG magnetic field strength.

We also discussed the implications of results for the effect of stellar flares on exoplanet habitability, in particular for exoplanets

around M stars. We compared our average flare occurrence rates to the abiogenesis zone from Rimmer et al. (2018). We found that no spectral sample, at any age, in our sample had an average flare occurrence rate which reached the abiogenesis zone at our observed energies. However, we note this does not rule out the possibility of individual active stars flaring at the required rates as found in other surveys.

## ACKNOWLEDGEMENTS

We are thankful to the referee, Manuel Güdel, for their helpful comments and suggestions. This research is based on data collected under the NGTS project at the ESO La Silla Paranal Observatory. The NGTS facility is funded by a consortium of institutes consisting of the University of Warwick, the University of Leicester, Queen's University Belfast, the University of Geneva, the Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR; under the 'Großinvestition GI-NGTS'), the University of Cambridge, together with the UK Science and Technology Facilities Council (STFC; project references ST/M001962/1 and ST/S002642/1). JAGJ was supported by STFC PhD studentship 1763096 for this work and now acknowledges support from grant *HST*-GO-15955.004-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. PJW, RGW, SG, and BTG acknowledge support by STFC under consolidated grants ST/P000495/1 and ST/T000406/1. BTG also acknowledges support from a Leverhulme Trust Research Fellowship. CEP acknowledges support from the European Research Council under the SeismoSun Research Project No. 321141. JSJ acknowledges support by FONDECYT grant 1201371, and partial support from CONICYT project Basal AFB-170002. EG gratefully acknowledges support from the David and Claudia Harding Foundation in the form of a Winton Exoplanet Fellowship.

This publication has used 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 publication has used 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. This work has used 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.

## DATA AVAILABILITY

Derived stellar and measured flare properties are available upon request to JAGJ.

## REFERENCES

Alekseev I. Y. et al., 1994, *A&A*, 288, 502  
 Andrews A. D., Panagi P. M., 1996, *Ir. Astron. J.*, 23, 171  
 Anfinogentov S., Nakariakov V. M., Mathioudakis M., Van Doorselaere T., Kowalski A. F., 2013, *ApJ*, 773, 156  
 Arenou F. et al., 2018, *A&A*, 616, A17

Aschwanden M. J., Tarbell T. D., Nightingale R. W., Schrijver C. J., Title A., Kankelborg C. C., Martens P., Warren H. P., 2000, *ApJ*, 535, 1047  
 Astropy Collaboration, 2013, *A&A*, 558, A33  
 Bailer-Jones C. A. L., Rybizki J., Fousneau M., Mantelet G., Andrae R., 2018, *AJ*, 156, 58  
 Balona L. A., Broomhall A.-M., Kosovichev A., Nakariakov V. M., Pugh C. E., Van Doorselaere T., 2015, *MNRAS*, 450, 956  
 Baraffe I., Homeier D., Allard F., Chabrier G., 2015, *A&A*, 577, A42  
 Bensby T., Feltzing S., Lundström I., 2003, *A&A*, 410, 527  
 Bensby T., Feltzing S., Lundström I., 2004, *A&A*, 421, 969  
 Bensby T., Feltzing S., Oey M. S., 2014, *A&A*, 562, A71  
 Benz A. O., 2017, *Living Rev. Sol. Phys.*, 14, 2  
 Benz A. O., Güdel M., 2010, *ARA&A*, 48, 241  
 Berdyugina S. V., 2005, *Living Rev. Sol. Phys.*, 2, 8  
 Birko D. et al., 2019, *AJ*, 158, 155  
 Bonfils X. et al., 2018, *A&A*, 613, A25  
 Booth R. S., Poppenhaeger K., Watson C. A., Silva Aguirre V., Wolk S. J., 2017, *MNRAS*, 471, 1012  
 Borucki W. J. et al., 2010, *Science*, 327, 977  
 Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, *MNRAS*, 427, 127  
 Carrington R. C., 1859, *MNRAS*, 20, 13  
 Chen X., Wang S., Deng L., de Grijs R., Yang M., 2018, *ApJS*, 237, 28  
 Collier Cameron A. et al., 2006, *MNRAS*, 373, 799  
 Crosby N. B., Aschwanden M. J., Dennis B. R., 1993, *Sol. Phys.*, 143, 275  
 Cutri R. M., et al., 2014, *VizieR Online Data Catalog*, p. 2328  
 Davenport J. R. A., 2016, *ApJ*, 829, 23  
 Davenport J. R. A., Covey K. R., Clarke R. W., Boeck A. C., Cornet J., Hawley S. L., 2019, *ApJ*, 871, 241  
 Davenport J. R. A. et al., 2014, *ApJ*, 797, 122  
 Dittmann J. A. et al., 2017, *Nature*, 544, 333  
 Douglas S. T., Agüeros M. A., Covey K. R., Cargile P. A., Barclay T., Cody A., Howell S. B., Kopytova T., 2016, *ApJ*, 822, 47  
 Doyle J. G., Butler C. J., 1985, *Nature*, 313, 378  
 Doyle L., Ramsay G., Doyle J. G., Wu K., 2019, *MNRAS*, 489, 2115  
 Doyle L., Ramsay G., Doyle J. G., Wu K., Scullion E., 2018, *MNRAS*, 480, 2153  
 Engle S. G., Guinan E. F., Harper G. M., Cuntz M., Remage Evans N., Neilson H. R., Fawzy D. E., 2017, *ApJ*, 838, 67  
 Feinstein A. D., Montet B. T., Ansdell M., Nord B., Bean J. L., Günther M. N., Gully Santiago M. A., Schlieder J. E., 2020, *AJ*, 160, 219  
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306  
 Frasca A., Biazzo K., Taş G., Evren S., Lanzafame A. C., 2008, *A&A*, 479, 557  
 Gaia Collaboration, 2016, *A&A*, 595, A2  
 Gaia Collaboration, 2018a, *A&A*, 616, A1  
 Gaia Collaboration, 2018b, *A&A*, 616, A10  
 Gaia Collaboration, 2019, *A&A*, 623, A110  
 Gershberg R. E., Shakhovskaia N. I., 1983, *Ap&SS*, 95, 235  
 Gilliland R. L. et al., 2010, *ApJ*, 713, L160  
 Gizis J. E., Paudel R. R., Schmidt S. J., Williams P. K. G., Burgasser A. J., 2017, *ApJ*, 838, 22  
 Güdel M., Audard M., Kashyap V. L., Drake J. J., Guinan E. F., 2003, *ApJ*, 582, 423  
 Günther M. N. et al., 2017, *MNRAS*, 472, 295  
 Günther M. N. et al., 2020, *AJ*, 159, 60  
 Hawley S. L., Davenport J. R. A., Kowalski A. F., Wisniewski J. P., Hebb L., Deitrick R., Hilton E. J., 2014, *ApJ*, 797, 121  
 Hawley S. L., Fisher G. H., 1992, *ApJS*, 78, 565  
 Hayashi C., 1961, *PASJ*, 13, 450  
 Henden A., Munari U., 2014, *Contrib. Astron. Obs. Skalnat Pleso*, 43, 518  
 Hilton E. J., Hawley S. L., Kowalski A. F., Holtzman J., 2011, in Johns-Krull C., Browning M. K., West A. A., eds, *ASP Conf. Ser. Vol. 448, 16th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun*. Astron. Soc. Pac., San Francisco, p. 197



- Howard W. S., Corbett H., Law N. M., Ratzloff J. K., Glazier A., Fors O., del Ser D., Haislip J., 2019, *ApJ*, 881, 9
- Howard W. S. et al., 2020b, *ApJ*, 895, 140
- Howard W. S. et al., 2020a, *ApJ*, 902, 115
- Ilin E., Schmidt S. J., Davenport J. R. A., Strassmeier K. G., 2019, *A&A*, 622, A133
- Ilin E., Schmidt S. J., Poppenhäger K., Davenport J. R. A., Kristiansen M. H., Omohundro M., 2021, *A&A*, 645, A42
- Jack D., 2019, *Astron. Nachr.*, 340, 386
- Jackman J. A. G. et al., 2018, *MNRAS*, 477, 4655
- Jackman J. A. G. et al., 2019a, *MNRAS*, 482, 5553
- Jackman J. A. G. et al., 2019b, *MNRAS*, 485, L136
- Jackman J. A. G. et al., 2019c, *MNRAS*, 489, 5146
- Jackman J. A. G. et al., 2020, *MNRAS*, 497, 809
- Jayasinghe T. et al., 2018, *MNRAS*, 477, 3145
- Jiang Y., Zheng R., Yang J., Hong J., Yi B., Yang D., 2012, *ApJ*, 744, 50
- Jiang Y.-C., Bi Y., Yang J.-Y., Zheng R.-S., Wang J.-X., 2009, *Res. Astron. Astrophys.*, 9, 603
- Katz D. et al., 2018, *A&A*, 616, A11
- Kay C., Opher M., Kornbleuth M., 2016, *ApJ*, 826, 195
- Kounkel M. et al., 2018, *AJ*, 156, 84
- Kowalski A. F., Hawley S. L., Wisniewski J. P., Osten R. A., Hilton E. J., Holtzman J. A., Schmidt S. J., Davenport J. R. A., 2013, *ApJS*, 207, 15
- Kowalski A. F. et al., 2019, *ApJ*, 871, 167
- Kunder A. et al., 2017, *AJ*, 153, 75
- Kóvári Z. et al., 2020, *A&A*, 641, A83
- Lacy C. H., Moffett T. J., Evans D. S., 1976, *ApJS*, 30, 85
- Lammer H. et al., 2007, *Astrobiology*, 7, 185
- Leavitt H. S., Pickering E. C., 1912, *Harv. Coll. Obs. Circ.*, 173, 1
- Loyd R. O. P., Shkolnik E. L., Schneider A. C., Barman T. S., Meadows V. S., Pagano I., Peacock S., 2018a, *ApJ*, 867, 70
- Loyd R. O. P. et al., 2018b, *ApJ*, 867, 71
- Lu H.-P., Zhang L.-y., Shi J., Han X. L., Fan D., Long L., Pi Q., 2019, *ApJS*, 243, 28
- Maehara H., Notsu Y., Notsu S., Namekata K., Honda S., Ishii T. T., Nogami D., Shibata K., 2017, *PASJ*, 69, 41
- Maehara H., Shibayama T., Notsu Y., Notsu S., Honda S., Nogami D., Shibata K., 2016, in Kosovichev A. G., Hawley S. L., Heinzel P., eds, *Solar and Stellar Flares and their Effects on Planets*, Proc. IAU Symp. 320, Proc. IAU, United Kingdom, p. 144
- Maehara H. et al., 2012, *Nature*, 485, 478
- Martin D. C. et al., 2005, *ApJ*, 619, L1
- Matt S. P., Brun A. S., Baraffe I., Bouvier J., Chabrier G., 2015, *ApJ*, 799, L23
- Medina A. A., Winters J. G., Irwin J. M., Charbonneau D., 2020, *ApJ*, 905, 107
- Moffett T. J., 1974, *ApJS*, 29, 1
- Morin J., Donati J.-F., Petit P., Delfosse X., Forveille T., Jardine M. M., 2010, *MNRAS*, 407, 2269
- Morin J. et al., 2008, *MNRAS*, 390, 567
- Moschou S.-P., Vlahakis N., Drake J. J., Evans N. R., Neilson H. R., Guzik J. A., Zuhone J., 2020, *ApJ*, 900, 157
- Mullan D. J., Bais H. P., 2018, *ApJ*, 865, 101
- Mullan D. J., MacDonald J., Dieterich S., Fausey H., 2018, *ApJ*, 869, 149
- Newton E. R., Irwin J., Charbonneau D., Berlind P., Calkins M. L., Mink J., 2017, *ApJ*, 834, 85
- Notsu Y. et al., 2013, *ApJ*, 771, 127
- Notsu Y. et al., 2019, *ApJ*, 876, 58
- Oelkers R. J. et al., 2018, *AJ*, 155, 39
- Oskanian V., 1969, in *Information Bulletin on Variable Stars*, No. 345, p. 1
- Paudel R. R., Gizis J. E., Mullan D. J., Schmidt S. J., Burgasser A. J., Williams P. K. G., Berger E., 2018, *ApJ*, 858, 55
- Pecaut M. J., Mamajek E. E., 2013, *ApJS*, 208, 9
- Pettersen B. R., Coleman L. A., Evans D. S., 1984, *ApJS*, 54, 375
- Pizzolato N., Maggio A., Micela G., Sciortino S., Ventura P., 2003, *A&A*, 397, 147
- Pugh C. E., Armstrong D. J., Nakariakov V. M., Broomhall A.-M., 2016, *MNRAS*, 459, 3659
- Rackham B. V., Apai D., Giampapa M. S., 2018, *ApJ*, 853, 122
- Ranjan S., Wordsworth R., Sasselov D. D., 2017, *ApJ*, 843, 110
- Reid I. N., Gizis J. E., 2005, *PASP*, 117, 676
- Ricker G. R. et al., 2014, in *Transiting Exoplanet Survey Satellite (TESS)*, Proc. SPIE, Vol. 9143, p. 914320
- Ricker G. R. et al., 2015, *J. Astron. Telesc. Instrum. Syst.*, 1, 014003
- Rimmer P. B., Xu J., Thompson S. J., Gillen E., Sutherland J. D., Queloz D., 2018, *Sci. Adv.*, 4, eaar3302
- Rodono M., 1974, *A&A*, 32, 337
- Rodríguez Martínez R., López L. A., Shappee B. J., Schmidt S. J., Jayasinghe T., Kochanek C. S., Auchettl K., Holoién T. W.-S., 2020, *ApJ*, 892, 144
- Roettenbacher R. M., Vida K., 2018, *ApJ*, 868, 3
- Saar S. H., Linsky J. L., Beckers J. M., 1986, *ApJ*, 302, 777
- Schaefer B. E., 1989, *ApJ*, 337, 927
- Schaefer B. E., King J. R., Deliyannis C. P., 2000, *ApJ*, 529, 1026
- Schmidt S. J., Cruz K. L., Bongiorno B. J., Liebert J., Reid I. N., 2007, *AJ*, 133, 2258
- Schmidt S. J. et al., 2016, *ApJ*, 828, L22
- Schmidt S. J. et al., 2019, *ApJ*, 876, 115
- Schneider A. C., Shkolnik E. L., 2018, *AJ*, 155, 122
- Segura A., Walkowicz L. M., Meadows V., Kasting J., Hawley S., 2010, *Astrobiology*, 10, 751
- Shakhovskaia N. I., 1989, *Sol. Phys.*, 121, 375
- Shibata K., 1999, *Ap&SS*, 264, 129
- Shibayama T. et al., 2013, *ApJS*, 209, 5
- Shimizu T., 1995, *PASJ*, 47, 251
- Shulyak D., Reiners A., Engeln A., Malo L., Yadav R., Morin J., Kochukhov O., 2017, *Nat. Astron.*, 1, 0184
- Skrutskie M. F. et al., 2006, *AJ*, 131, 1163
- Stassun K. G. et al., 2019, *AJ*, 158, 138
- Tilley M. A., Segura A., Meadows V. S., Hawley S., Davenport J., 2019, *Astrobiology*, 19, 64
- Tsurutani B. T., Gonzalez W. D., Lakhina G. S., Alex S., 2003, *J. Geophys. Res.*, 108, 1268
- Wang H., Chae J., Yurchyshyn V., Yang G., Steinegger M., Goode P., 2001, *ApJ*, 559, 1171
- Watson C. L., Henden A. A., Price A., 2006, in *The Society for Astronomical Sciences 25th Annual Symposium on Telescope Science*, Society for Astronomical Sciences, California, USA, p. 47
- West A. A., Hawley S. L., Bochanski J. J., Covey K. R., Reid I. N., Dhital S., Hilton E. J., Masuda M., 2008, *AJ*, 135, 785
- Wheatley P. J. et al., 2018, *MNRAS*, 475, 4476
- Winters J. G. et al., 2019, *AJ*, 157, 216
- Wright E. L. et al., 2010, *AJ*, 140, 1868
- Wright N. J., Drake J. J., Mamajek E. E., Henry G. W., 2011, *ApJ*, 743, 48
- Yang H., Liu J., 2019, *ApJS*, 241, 29
- Yang H., Liu J., Qiao E., Zhang H., Gao Q., Cui K., Han H., 2018, *ApJ*, 859, 87
- Yang H. et al., 2017, *ApJ*, 849, 36
- Yan X. L., Wang J. C., Pan G. M., Kong D. F., Xue Z. K., Yang L. H., Li Q. L., Feng X. S., 2018, *ApJ*, 856, 79
- Zacharias N., Finch C. T., Girard T. M., Henden A., Bartlett J. L., Monet D. G., Zacharias M. I., 2013, *AJ*, 145, 44
- Zirin H., Liggett M. A., 1987, *Sol. Phys.*, 113, 267
- Zuccarello F. et al., 2009, *A&A*, 493, 629

## APPENDIX A: DERIVATION OF THE OBSERVED OCCURRENCE RATE

Here, we will outline how we derived the observed flare occurrence rate used in Section 2.5.1. As outlined in Section 2.5, stellar flares are expected to occur with a power-law, or power-law-like, distribution in energy. The differential distribution of the number of flares in a given duration,  $N$ , with energy  $E$  is written as

$$\frac{dN(E)}{dE} = kE^{-\alpha}, \quad (\text{A1})$$

**Table A1.** Parameters and corresponding covariance matrices from the power law fits to main-sequence and pre-main-sequence subsets of our data.

Class	$\log k$	$\alpha$	Covariance matrix
K2V-K4V	$29.4 \pm 7.2$	$1.97 \pm 0.20$	$\begin{bmatrix} 51.3 & 1.46 \\ 1.46 & 0.04 \end{bmatrix}$
K5V-K8V	$25.0 \pm 7.4$	$1.85 \pm 0.21$	$\begin{bmatrix} 55.46 & 1.56 \\ 1.56 & 0.04 \end{bmatrix}$
M0V-M2V	$30.4 \pm 3.4$	$2.01 \pm 0.10$	$\begin{bmatrix} 11.7 & 0.34 \\ 0.34 & 0.009 \end{bmatrix}$
M3V-M5V	$33.9 \pm 1.55$	$2.09 \pm 0.05$	$\begin{bmatrix} 2.42 & 0.07 \\ 0.07 & 0.002 \end{bmatrix}$
K5-K8 (<15 Myr)	$25.4 \pm 4.8$	$1.82 \pm 0.14$	$\begin{bmatrix} 23.4 & 0.66 \\ 0.66 & 0.002 \end{bmatrix}$
M0-M2 (<30 Myr)	$20.9 \pm 1.8$	$1.69 \pm 0.05$	$\begin{bmatrix} 3.17 & 0.09 \\ 0.09 & 0.003 \end{bmatrix}$
M3-M5 (<40 Myr)	$29.5 \pm 1.2$	$1.94 \pm 0.04$	$\begin{bmatrix} 1.47 & 0.04 \\ 0.04 & 0.001 \end{bmatrix}$
Orion 3400-3940 K	$30.7 \pm 5.7$	$1.94 \pm 0.16$	$\begin{bmatrix} 31.2 & 0.88 \\ 0.88 & 0.02 \end{bmatrix}$

where  $k$  is a normalization constant and  $\alpha$  is the power-law index. When integrated from a flare energy  $E_f$  to infinity, we obtain the standard cumulative flare frequency distribution. The observed distribution of flares with energy also depends on the efficiency of the instrument being used to sample the flares. This efficiency, also called the recovery rate  $R(E)$ , is a function of energy. At high energies, we expect  $R(E)$  to be near or equal to 1, meaning we observe all flare events. At low energies,  $R(E)$  tends to zero, resulting in fewer flare detections. Therefore, we can write the observed number of flares with energy as

$$dN(E) = k R(E) E^{-\alpha} dE. \quad (\text{A2})$$

$R(E)$  can be sampled through flare injection and recovery tests. The cumulative flare frequency distribution is then

$$N(E > E_{\text{flare}}) = k \int_{E_{\text{flare}}}^{\infty} R(E) E^{-\alpha} dE. \quad (\text{A3})$$

Integrating by parts then gives

$$N(E > E_f) = k \left( \left[ \frac{R(E) E^{-\alpha+1}}{1-\alpha} \right]_{E_{\text{flare}}}^{\infty} - \int_{E_f}^{\infty} \frac{R'(E) E^{-\alpha+1}}{1-\alpha} dE \right), \quad (\text{A4})$$

where  $R'(E)$  is the differentiated  $R(E)$ . Assuming  $\alpha$  is always greater than 1, this results in

$$N(E > E_f) = k \left( - \frac{R(E_f) E_f^{-\alpha+1}}{1-\alpha} - \int_{E_f}^{\infty} \frac{R'(E) E^{-\alpha+1}}{1-\alpha} dE \right) \quad (\text{A5})$$

and then

$$N(E > E_f) = \frac{k}{\alpha-1} \left( R(E_f) E_f^{-\alpha+1} + \int_{E_f}^{\infty} R'(E) E^{-\alpha+1} dE \right). \quad (\text{A6})$$

Above some energy  $E_{\text{max}}$ , the recovery fraction will not change, i.e. it has reached its maximum efficiency, so  $R'(E)$  will equal zero. If this known (e.g from flare injection and recovery tests), then we can write the observed cumulative flare frequency as

$$N(E > E_f) = \frac{k}{\alpha-1} \left( R(E_f) E_f^{-\alpha+1} + \int_{E_f}^{E_{\text{max}}} R'(E) E^{-\alpha+1} dE \right). \quad (\text{A7})$$

We can see at high-flare energies that can all be detected, then  $R'(E) = 0$  and  $R(E) = 1$ , giving the standard cumulative flare distribution. At low energies, where  $R(E)$  equals zero, then the distribution plateaus to the value of the integral.

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