Flares in Open Clusters with K2.

I. M45 (Pleiades), M44 (Praesepe) and M67

Ekaterina Ilin[1], Sarah J. Schmidt[1], James R. A. Davenport[2] and Klaus G. Strassmeier[1]

[1] Leinbiz Institut für Astrophysik Potsdam
   e-mail: eilin@aip.de
[2] University of Washington e-mail: jrad@uw.edu

ABSTRACT

Context. The presence and strength of a stellar magnetic field and activity is rooted in a star's fundamental parameters such as mass and age. Can flares serve as an accurate stellar "clock"?
Aims. To explore if we can quantify an activity-age relation in the form of a flaring-age relation we measured trends in the flaring rates and energies for stars with different masses and ages.
Methods. We investigated the time-domain photometry provided by Kepler's follow-up mission K2 and searched for flares in three solar metallicity open clusters with well-known ages, M45 (0.125 Gyr), M44 (0.63 Gyr) and M67 (4.3 Gyr). We employed the automated flare finding and analysis pipeline Appaloosa, originally designed for Kepler by James R. A. Davenport. We introduced a synthetic flare injection and recovery subroutine to ascribe detection and energy recovery rates for flares in a broad energy range for each light curve.
Results. We collected a sample of 1,638 stars, mostly late-K to mid-M dwarfs and found 804 flare candidates with flare energies ranging from $5 \cdot 10^{31}$ erg to 10^{34} erg, of which 647 belong to M45, 157 to M44 and none to M67. On one hand, flaring activity depends both on $T_{\rm eff}$, and age. On the other hand, all flare frequency distributions have similar slopes with $\alpha \approx 2.0 - 2.4$, supporting a universal flare generation process. We discuss implications for the physical conditions under which flares occur, and how the sample's metallicity and multiplicity affect our results.

Key words. Methods: data analysis, Stars: activity, Stars: flare, Stars: low-mass

1. Introduction

As stars age, their magnetic activity evolves. A magnetic dynamo should be at work within all stars with outer convection zones (Schatzman 1962). When the dynamic magnetic field in the interior reaches the stellar surface and interacts with the atmosphere, a variety of phenomena arise, including star spots, chromospheric emission, and flares (Parker 1979). Flares are magnetic reconnection events that lead to a change in field line topology and subsequent energy release (Priest & Forbes 2002). They couple the chromosphere to the corona and emit electromagnetic energy in the form of both thermal and non-thermal emission in radio, hard and soft X-rays (> and $\leq$ 10 keV, respectively), UV and white light (see Benz & Güdel 2010; Benz 2016, for detailed reviews). The amount of energy released during a stellar flare relative to the star's luminosity can exceed the strongest solar flares and even temporarily amplify the optical stellar flux by orders of magnitude for cool dwarfs (Schaefer et al. 2000; Candelaresi et al. 2014). Flaring activity's prospective observational availability and strong link to magnetic evolution has incited us to attempt to quantify it as a function of stellar age and mass.
Age, rotation and magnetic activity are tightly interrelated by the concept of stellar dynamos (Noyes et al. 1984). Observational evidence for a relation between stellar rotation, mass and age (Radick et al. 1987; Patten & Simon 1996; Queloz et al. 1998, to name a few), and its theoretical backing (Mestel 1984; Kawaler 1988) gave rise to gyrochronology (Barnes 2003) – the

age-dating of a single main sequence star from its mass and rotation period. Barnes (2010), Meibom et al. (2015) and Barnes et al. (2016) have found an unambiguous relation to hold for stellar ages from 0.6 Gyr (Hyades) up to about 4.3 Gyr (M67) for solar type stars. However, this relation varies strongly with stellar mass (Barnes 2010; Van Saders et al. 2016). Additionally, the detectability of rotation periods in photometric data drops rapidly beyond solar ages, as noted by Esselstein et al. (2018). Secondary magnetic stellar age tracers such as chromospheric activity (Soderblom et al. 1991; Pace 2013; Lorenzo-Oliveira et al. 2016) or magnetic activity from Zeeman Doppler Imaging (ZDI) and Zeeman Broadening (ZB) (Vidotto et al. 2014) are being explored but have not yet been developed into full blown age dating techniques.
Among these more or less expedient magnetic age tracers flaring activity prospectively stands out with respect to practicability. In cool dwarfs, flares do not suffer from a lack of contrast to quiescent luminosity due to their $\sim$ 10,000 K blackbody spectrum (Hawley & Fisher 1992), and are observable in most common photometric bands (Benz & Güdel 2010). If flare processes are truly scale invariant (Lacy et al. 1976; Shakhovskaya 1989) and have a sufficiently long cosmic activity timescale, a flaring-age clock may be calibrated for a broad range of stellar masses down to ultra-cool dwarfs (UCDs, $T_{\rm eff} < 2,900$ K, Kirkpatrick et al. 1999) and take over when other age relations break down. In accordance to dynamo theories, the flaring-age relation is expected to depend on stellar mass. In fact, Walkowicz et al. (2011)

note that M dwarfs flare more often in white light (but are less energetic) than K dwarfs, and release more energy per time relative to their quiescent flux. However, at least for late-M dwarfs and UCDs, the relation seems weak (Paudel et al. 2018).

On the other hand, some constraints that limit the use of stellar age tracers may apply to a flaring clock as well: Assuming rotation is the main driver of any magnetic activity via the classical dynamo theory, a flaring clock cannot be more precise than gyrochronology on an individual star. Short term variations due to stellar activity cycles (Montet et al. 2017), breakdown of the relation in a certain age or mass range (e.g. Pace 2013; Paudel et al. 2018), or saturation of flaring activity above some critical Rossby number similar to what is observed in X-ray luminosity (Wright et al. 2011) could diminish the clock's utility. Furthermore, a gyrochronology-type convergence of the flaring-age relation may not occur or occur later than in the case of rotation. Multiplicity of stellar systems affects our implications about the rotation-age relation both drawn from large samples (Douglas 2016; Douglas et al. 2017) and regarding individual systems' evolution (Rebull et al. 2016). We expect similar arguments to apply to flares. Metallicity influences magnetic activity (Gray et al. 2006; Karoff et al. 2018) and we expect it to affect flaring behaviour. A technical constraint is currently set by a target's total available monitoring time relative to the flare frequency, in particular for ground based observations, but time-resolved photometric surveys like CoRoT (Auvergne et al. 2009), *Kepler* (Jenkins 2010) and the *Transiting Exoplanetary Survey Satellite* (*TESS*; Ricker et al. 2014) have begun to turn the tide.

Evidence for a measurable time scale for flaring activity in low-mass stars has accumulated during the last decade. Based on data from the Sloan Digital Sky Survey (SDSS; York et al. 2000), both Kowalski (2009) and Hilton et al. (2010) find that flaring M0-M6 dwarfs are preferentially found at lower Galactic heights than stars without notable flaring activity but with Hα emission, implying a flaring activity lifetime for these dwarfs that is shorter than the chromospheric activity lifetime. Doorsselaere et al. (2017) used *Kepler* photometry and rotation information from McQuillan et al. (2014) to find that flaring activity is a good predictor of rotation period, which also favors a typical flaring activity timescale. Moreover, Clarke et al. (2018) inferred a promisingly low intrinsic variability from a largely consistent relation between flaring activity level, rotation and age for pairs of coeval stars with similar masses in a sample of wide binaries found in *Kepler* archives.

In this work, we test if a mass-dependent flaring-age relation exists and can be inferred from flare statistics in open cluster (OC) photometry. OCs observed by *Kepler*'s follow up mission *K2* (Howell et al. 2014) allow us to study large cohorts of coeval stars with a low spread in metallicity. We cover three of these objects, Pleiades (M45, 0.125 Myr; Bell et al. 2012), Praesepe (M44, 0.63 Gyr; Boudreault et al. 2012) and M67 (4.3 Gyr; Dias et al. 2012), observed during the mission's campaigns C4 and C5 (section 2).

We detected flares as bright peaks in *K2* time-domain photometry, using light curves (LCs) de-trended by Aigrain et al. (2016), with the flare finding and analysis pipeline `Appaloosa` (Davenport 2016) (section 3). Additional validation of flare candidates is obtained from an artificial signal injection and recovery subroutine that characterizes individual light curves with respect to detection rate and energy recovery. The final flare distributions are used to construct flare frequency distributions (FFDs), determine flare activity indicators and analyse these as functions of T_{eff} and age (section 4). We discuss the role of the targets' mul-
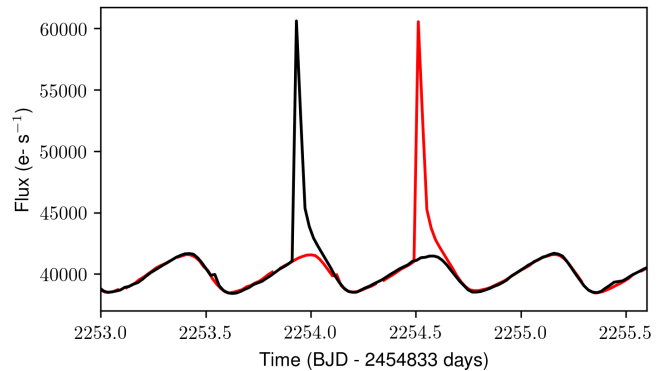

Fig. 1. $1.8 \cdot 10^{34}$ erg superflare observed on EPIC 211119999 (M45). Black: *K2* PDC_SAP flux. Red: K2SC residual model with periodic trends added. The light curves are offset by the dominant 0.58 d period.

tiplicity and metallicity, and put the results in the context of flare physics and coronal heating mechanisms in section 5.

2. Data

K2 time domain photometry is our source of flares, but the archived data need to be cleared of both systematic effects and intrinsic variability before analysis. We collected cluster membership information as well as multiband photometry to derive spectral type and T_{eff} for all targets that enabled us to group them by mass and age. We then estimated their luminosities to determine the energies that individual flares release. Each target was observed in 30 min cadence for up to 80 days each yielding light curves in which we searched for flare signatures.

2.1. De-trended Light Curves

The *Kepler* (Koch et al. 2010) mission has produced a bounty of high precision photometric observations in the Cygnus-Lyra region since its launch in 2009. In 2013, two reaction wheels failed, and the mission had to be redesigned. Since 2014, the follow-up mission *K2* (Howell et al. 2014) has been conducting ~ 80 days long observation campaigns nearby the ecliptic plane. Besides the reduced pointing accuracy, the main difference between *Kepler* and *K2* data are additional instrumental artifacts that come along with the challenging task of keeping the satellite in place. Systematic errors occur on different time scales, among which a 6 h trend, associated with the spacecraft roll, is most prominent (Van Cleve et al. 2016). If these are not properly removed, their shape occasionally resembles flare signatures.

Two de-trending approaches achieve restoration of *Kepler*'s former precision to a very similar degree: the pixel-level decorrelation method developed by Luger et al. (2016) (EVEREST) and the Gaussian Process (GP) de-trending performed by Aigrain et al. (2016) (K2SC). Because their released data products already include LCs with removed periodic signals for campaigns C3−C8 and C10, saving considerable computational effort, we opted for K2SC LCs (see example light curve in Fig. 1).

2.2. Open Clusters in K2

There are about 16 OCs observed during *K2* campaigns 0-18. They span a variety of ages from very young clusters (M21;

5 Myr; Piskunov et al. 2011) to some of the oldest we know about (M67; 4.3 Gyr; Howell et al. 2014). As a rule of thumb, we expect a solar type star to exhibit a flare with 10^{34} erg once every 800 years (Maehara et al. 2012), but, in general, later type stars flare more frequently throughout the whole flare energy spectrum (Doorsselaere et al. 2017). By choosing the most populous OCs, with $> 150 - 200$ members each, where every object features *2MASS* and/or *Pan-STARRS* band magnitudes (see section 2.3 below) as well as de-trended K2SC LCs, we maximize the observation time. Aiming for a broad age range, M45, M44, and M67 were selected, such that our sample covers ages from 125 Myr to 4.3 Gyr (Tab. 1).

M45 We use the sample determined by Rebull et al. (2016), where membership probabilities are primarily based on recent proper-motion studies (Bouy et al. 2015; Sarro et al. 2014; Lodieu et al. 2012). All ambiguous candidates were inspected on a case by case basis by Rebull et al. (2016). We use the same subset of those candidates except that they exclude the targets that have not been found to be periodic. This results in a set of high-confidence and lower-confidence members with $6 < K_s < 14.5$, yielding a sample of 826 stars with *K2* light curves. Lower-confidence means that all proper-motion studies confirmed it was an unambiguous member − as for high-confidence members − but single CMDs placed the star off the respective sequence. We include these, because we later construct our own CMDs and exclude certain targets by a similar criterion.

M44 We use a sample from Kraus & Hillenbrand (2007), who analysed photometry and proper motions of ~ 5 million objects to determine their memberships in different stellar populations. Douglas et al. (2014) selected 753 lower mass ($< 1.5 M_\odot$) and unsaturated ($K_p > 9$) stars from this survey with M44 membership and *K2* LCs, and added 41 known bright members not considered in the survey yielding a total of 794. We adopt their supplemented selection.

M67 The sample was drawn from Gonzalez (2016), who, similar to Douglas et al. (2014) and Rebull et al. (2016), studied stellar variability in *K2* light curves. They started with the sample used in the photometric survey by Nardiello et al. (2016), which they supplemented with other recent membership studies by Yadav et al. (2008) (proper motions) and Geller et al. (2015) (radial velocities). After checking for consistency in these works they assign, following Geller et al. (2015), a membership class to each object from which we only take those classified as M (members), SM (single members) and BM (binary members), resulting in a working sample of 272 stars with *K2* LCs.

From these three samples we finally removed targets that

- lacked multiband photometry and/or empirical spectra,
- fell off the main sequence in $g - r$, $r - i$, and/or $J - K$ color magnitude diagrams,
- or were assigned spectral types hotter than F4.

2.3. Multiband Photometry: $T_{\rm eff}$ and R_* and Excluded Objects

To consistently ascribe approximate stellar spectra and ultimately luminosities to the stars we retrieved photometric data from either the *Two Micron All Sky Survey* (*2MASS*; Skrutskie et al. 2006) or the *Panoramic Survey Telescope and Rapid Response System* (*Pan-STARRS*) Data Release 1 (Chambers et al.

Table 1. Cluster sample, partly adopted from Howell et al. (2014). LCs: number of LCs after matching the catalogs, without removed targets, i.e., outliers in the CMDs, stars that lack *Pan-STARRS* and *2MASS* multiband photometry, and stars with spectral type <F4. [Fe/H] values are retrieved from Netopil et al. (2016).

	M45	M44	M67
K2 campaign	4	5	5
age (Myr)	125	630	4,300
distance (pc)	135	160	908
distance, age (source)	(1)	(2)	(3)
LCs	759	768	258
removed targets	67	26	14
[Fe/H]	−0.01	0.16	0.03
membership	(4)	(5,6)	(7)

References. (1) Bell et al. 2012; (2) Boudreault et al. 2012; (3) Dias et al. 2012; (4) Rebull et al. 2016; (5) Kraus & Hillenbrand 2007; (6) Douglas et al. 2014; (7) Gonzalez 2016.

2016). For this, we matched the sample by position with the *Pan-STARRS* catalog, using the *CDS XMatch* tool[1] to collect the *Pan-STARRS grizy* measurements (*2MASS JHK* magnitudes are already matched in *EPIC*), which were then converted to SDSS *grizy* using Tab. 2 in Finkbeiner et al. (2016). We used all available color indices to derive $T_{\rm eff}$ and R_*, and eventually luminosities $L_{\rm Kp,*}$ that were required to determine individual released flare energies in the *Kepler* band ($E_{\rm Kp,flare}$, see section 2.5).

We employed the synthetic *SDSS/2MASS* colors, i.e. *JHK* and *grizy* bands, for solar metallicity standards (Pickles 1998) computed by Covey et al. (2007), and used the *Modern Mean Dwarf Stellar Color and Effective Temperature Sequence*, an up-to-date lookup table for spectral types, $T_{\rm eff}$, and R_* derived from $g - r$, $r - i$, $J - H$, and $H - K$ colors, compiled from the literature by Pecaut & Mamajek (2013)[2].

The temperature accuracy $\sigma_{T_{\rm eff}}$ in these two tables is heterogeneous for F4 to M9 dwarfs with median 75 K and ranging from 20 K to 220 K (see Tab. 2 for the correspondence of spectral types to $T_{\rm eff}$). The lowest accuracies are present for mid-K type stars. The majority targets had all magnitudes available. We adopted the median value from typically $3 - 4$ obtained $T_{\rm eff}$. Due to recent findings, e.g. by Jackson et al. (2018), who find that radii for fast rotating M dwarfs in the Pleiades are underpredicted in stellar models by $14 \pm 2\%$, we assume an uncertainty in R_* of $\sim 20\%$. This estimate also covers the spread in observed radii from Pecaut & Mamajek (2013).

We excluded all likely non-MS, early spectral type and foreground/background stars from the sample by examining the color-magnitude diagrams (CMDs) for $g - r$, $r - i$, and $J - K$ colors. The union of objects that fell off the main sequence in each CMD was rejected.

The final collection of targets mostly contains late-K to mid-M dwarfs and a few hotter stars. We divided our data in four sufficiently populated bins for stars below 4,000 K and otherwise considered the full sample.

---

[1] `http://cdsxmatch.u-strasbg.fr/xmatch`
[2] The updated version used here is 2018.03.22. It includes results from Kirkpatrick et al. (2011), Dieterich et al. (2014), Eker et al. (2015), Filippazzo et al. (2015), Benedict et al. (2016), Leggett et al. (2015), Schneider et al. (2015), Patel et al. (2014), West et al. (2011), Dahn et al. (2017) and Kaltcheva et al. (2017).

Table 2. Correspondence between spectral type and T_{eff}, as given by Pecaut & Mamajek (2013). T_{min}, T_{max} give the edges of considered T_{eff} bins.

T_{min} (K)	T_{max} (K)	spectral types
3,000	3,249	M3.5−M5.5
3,250	3,499	M2.5−M3
3,500	3,749	M1−M2
3,750	4,000	M0.5−K8
3,000	7,000	F4−M5.5

2.4. Excluded Data

A subset of data points from all LCs was excluded from further analysis to account for thruster firings and systematics which were not captured by the de-trending pipelines (see Tab. A.2 for an overview). CR flags were tracked but not removed. 24 severely saturated targets were removed from the sample. After the partial manual review of flare candidates we further excluded some LCs with anomalous variability and individual artifacts. A detailed description of excluded data and a full list of excluded LCs with high-energy artifacts is given in Appendix A.

2.5. Flare Energies

Several individual flare's parameters such as duration, amplitude, flavors of FWHM, or released energy are typically used in flare statistics (see e.g. Hawley et al. 2014; Yang et al. 2017; Yang et al. 2018). However, in long cadence LCs, duration and amplitude are subject to large errors arising mostly from low time sampling (Yang et al. 2018). The flare energy, i.e. the integration of the LC flux during a flare with the quiescent flux subtracted, is less severely affected by this.

Often, flares are described by a $T_{\text{flare}} \approx 9{,}000 - 10{,}000$ K (Hawley & Fisher 1992; Kretzschmar 2011; Shibayama et al. 2013) blackbody Spectral Energy Distribution (SED). We follow Shibayama et al. (2013), and define the projected stellar quiescent luminosity $L_{\text{Kp},*}$ and the flare luminosity in the *Kepler* band $L_{\text{Kp,flare}}$ as

$$L_{\text{Kp},*} = \pi R_*^2 \int d\lambda R_{\text{Kepler}}(\lambda) B_*(\lambda) \tag{1}$$

and

$$L_{\text{Kp,flare}} = A_{\text{flare}} \int d\lambda R_{\text{Kepler}}(\lambda) B_{\text{flare}}(\lambda), \tag{2}$$

respectively, with A_{flare} being the area covered by the flaring region on the stellar surface, R_* the stellar radius and R_{Kepler} the *Kepler* response function given in the *Kepler* instrument handbook (Van Cleve & Caldwell 2016). B_* and B_{flare} correspond to the SEDs of the star and the blackbody curve of the flare with T_{flare}. The ratio of these luminosities yields the relative flare luminosity $a_{\text{Kp,flare}}$ as obtained from the *Kepler* LC:

$$a_{\text{Kp,flare}} = \frac{L_{\text{Kp,flare}}}{L_{\text{Kp},*}} \tag{3}$$

Kowalski et al. (2013) measured strong variations in T_{flare} and flare spectra both between events and during the course of individual flares in several dMe stars, limiting the utility of the blackbody approximation. The *Kepler* flare energy $E_{\text{Kp,flare}}$ is more tractable. Instead of bolometric flare luminosity we use the observed flare luminosity in the *Kepler* band $L_{\text{Kp,flare}}$ (eqn. 2) and

obtain

$$
\begin{aligned}
E_{\text{Kp,flare}} &= L_{\text{Kp},*} \int_{t_0}^{t_0 + \Delta t_{\text{flare}}} dt\, a_{\text{Kp,flare}} \\
&= L_{\text{Kp},*} \cdot ED.
\end{aligned} \tag{4}
$$

ED is defined as the area between the LC and the quiescent flux, i.e. the integrated flare flux divided by the median quiescent flux F_0 of the star, integrated over the flare duration (Hunt-Walker et al. 2012):

$$ED = \int dt \frac{F_{flare}(t)}{F_0} \tag{5}$$

Since it is measured relative to the quiescent star, *ED* is a quantity independent of calibration and distance. Note, that by this approximation we miss at least the $\sim 27\,\%$ of the continuum flare flux that resides in the U band relative to the total in UBVR (Hawley & Fisher 1992) and flux from emission lines that lie outside the *Kepler* band (see Kowalski et al. (2013) and references therein). As a consequence, $E_{\text{Kp,flare}}$ should be considered a lower limit to the total released energy of the flare.

2.5.1. Stellar Spectra and L_*

We calculate $L_{\text{Kp},*}$ as defined in eqn. 1 using spectra for FGKM stars computed by Yee et al. (2017). Including a parametrized spectrum mitigates errors arising from deviations from the blackbody assumption which becomes particularly relevant as we move to low-mass stars that have strong absorption features in the *Kepler* band along with an ever lower flux overall, which aggravates the spectral effects on total luminosity. Yee et al. (2017) provide empirical spectra for $3{,}000 - 7{,}000$ K and $0.1 - 16\,R_\odot$ main sequence and giant branch stars with an approximate accuracy of ± 100 K and $10\,\%$, respectively, rendering them approximately as precise as the color-temperature relations in (Pecaut & Mamajek 2013). We use Empirical SpecMatch (SpecMatch-Emp[3]) as a tool to assign template spectra according to the stars' spectral types. The stored spectra have high resolution ($R \approx 60{,}000$) and high signal-to-noise ratios ($S/N \approx 150$). The metallicities are centered around solar values with a standard deviation of ~ 0.2 dex and maximum deviations of ± 0.6 dex. We include the spectral information from derived R_* and T_{eff} to match a spectrum F_{T_{eff},R_*} to each star:

$$L_{\text{Kp},*} = \pi R_*^2 \int d\lambda R_{\text{Kepler}} F_{T_{\text{eff}},R_*}(\lambda) B_*(\lambda) \tag{6}$$

3. Automated Flare Finding

`Appaloosa` is an open-source[4] flare finding and analysis procedure written in Python by Davenport (2016) for *Kepler* LCs. The original version performs two successive steps: First, a model is built for the quiescent stellar brightness. Second, outliers that fulfill a number of detection criteria are analysed and stored in a flare candidate list.

We skip the first step and employ K2SC LCs, designed by Aigrain et al. (2016) for *K2* data, instead, because they already equip us with robust de-trending and variability removal. With K2SC LCs, we can treat the residual flux directly and approximate the quiescent flux by the median of all flux measurements.

We exclusively make use of the second step in the procedure that

---

[3] `https://github.com/samuelyeewl/specmatch-emp`
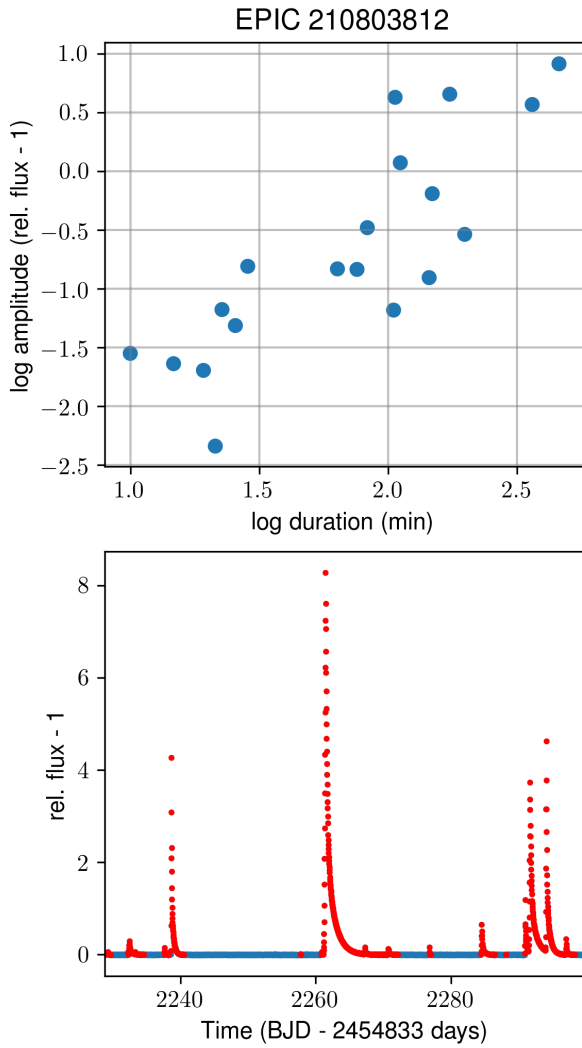[4] `https://github.com/jradavenport/appaloosa`

Fig. 2. Synthetic flare injection for EPIC 210803812 (M45). Left: Distribution of amplitudes and durations of synthetic flare injections. Right: Representation of this distribution in the original K2SC de-trended LC.

checks every flux outlier for three criteria to determine if it is part of a flare: Outliers from the residual LC are treated as flare candidates if they exceed thresholds N_1 and N_2, defined in terms of the LC's variance σ, i.e.

$$\frac{|f_i - \bar{f}|}{\sigma} \geq N_1 \tag{7}$$

$$\frac{|f_i - \bar{f} + w_i|}{\sigma} \geq N_2 \tag{8}$$

where f_i and w_i are the photometric flux and uncertainty at a given time i, $\bar{f}$ and σ are the median value and the statistical uncertainty in a continuous observation period, as introduced by Chang et al. (2015) in their `FINDFlare` algorithm. Eqn. 7 and 8 define the first two criteria; the third criterion (N_3) is the minimum number of consecutive data points that fulfill eqn. 7 and 8. In this work, we require at least 3 consequent data points ($N_3 \geq 3$, i.e. durations of ≥ 1.5 h) for a candidate detection for all outliers that exceed the thresholds $N_1 = 3$ and $N_2 = 4$. For all candidates, the start, peak and end times, and the equivalent duration (ED, eqn. 5) are extracted.
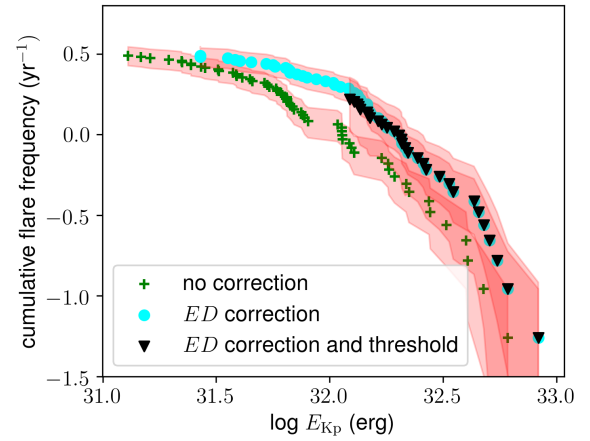

Fig. 3. Flare recovery and correction scheme for M44 ($3, 500 - 3, 749$ K). Green crosses: all detected candidates, no cutoff at low probabilities, no correction for systematic underestimation of ED. Blue dots: all detected candidates, no cutoff at low probabilities, but corrected for systematic underestimation of ED. Black triangles: Final flare distribution after cutting off at the median detection thresholds for the bin and correcting for systematic underestimation of ED. Red shadows: One standard deviation uncertainties.

3.1. Flare Finding Efficiency

A variety of reasons can prevent a flare from being detected in a LC − an event can be lost in the noise, cut off at the end of a continuous observing period, or subjected to filter effects induced by the employed de-trending procedure. Due to the iterative nature of the flare finding procedure and the heterogeneity of LCs it is also not possible to assess the efficiency of the code analytically. Therefore, we address this problem in a cause-neutral, empirical manner: We test `Appaloosa`'s flare recovery efficiency by injecting artificial signatures generated from a semi-analytical flare model derived from the active dMe star GJ 1243 (Davenport et al. 2014).

The synthetic events are introduced to a LC at different times with varying amplitude and duration while avoiding overlap with real flare signatures (Davenport 2016, see Fig. 2). The contaminated LC passes through the entire flare finding pipeline. A flare is then considered recovered if the flare peak time is contained within the start and end times of any resulting flare event candidates. After relating all successful and failed detections to each other, the recovery rate as a function of ED is returned. Recently, Paudel et al. (2018) used this procedure to determine a minimum energy of artificial flares detected by the algorithm. We went a step further and separated this quantity into detection rates for individual flare energies and introduced improvements to flare energy recovery and the parameter space covered by the injection routine. Details and examples are in Appendix B.

3.1.1. Correction Scheme and Uncertainties

We specified a scheme for the post-detection treatment of individual flare candidates and the associated uncertainties in cumulative FFDs, illustrated in Fig. 3 on the example of M44.

We fitted a power law to an uncorrected FFD with given uncertainties on energy and Poissonian uncertainty in the rate. As a first correction, we then adjusted the recorded flare energies according to the energy recovery ratios obtained from synthetic flare injections. This typically shifted the distribution to higher

energies in the diagram. In a second step, we rejected all flares with recovery probability below 20 %, a number obtained from experimentation and manual vetting of candidates. The recovery probability p then ranged from 20 % to 100 %. We then bolstered the square-root growth of the Poissonian uncertainty by the strongly increasing uncertainty. Induced by decreasing p the factor grew as $p^{-1/2}$ by up to a factor of ~ 2.2.

Besides the uncertainties on the event counts and systematic errors on recovered flare EDs, we estimated the uncertainties for each flare's $E_{\mathrm{Kp,flare}}$ (see overview in Tab. B.1 in Appendix B). Shibayama et al. (2013) report errors on their energy calculation to be around ±60 %, which is consistent with our estimate of ~ 65 %. For most flares, uncertainties mainly stem from the uncertainty on T_{eff}(color) and R_*, or $L_{\mathrm{bol},*}$, which can reach up to 80 % while the uncertainty on ED is typically below 30 % with the systematic error accounted for by the aforementioned ED correction.

All in all, the data set's size and quality present a mixed picture. On one hand, the low time resolution of our LCs limits the investigations to high energy (super-)flares. The targets' characterization in terms of T_{eff} and $L_{\mathrm{bol},*}$ is subject to large uncertainties, not least because of the observed departure of low mass stars from modern stellar models and the lack of standard stars for late-K type stars. These uncertainties propagate all the way through to $E_{\mathrm{Kp,flare}}$. On the other hand, K2SC de-trended LCs nearly restore the original *Kepler* precision and provide good model fits to the systematic variations of individual LCs while preserving astrophysical signal. Our final sample counts 1,638 targets, mostly late-K to mid-M dwarfs, that sum up to over 250 yr of total observation time. Ultimately, our synthetic flare injection procedure allows us to evaluate the quality of each individual LC and additionally validates the detected candidates.

4. Results

We analyse if flaring activity depends on spectral type and age in late-K to mid-M dwarfs. After a short view on the total flare event counts we introduce activity indicators, i.e., the flaring rate (FR), the energy released in flares (FA), the power law fit exponent α and intercept β. These indicators complement each other and can be linked to the underpinning astrophysical models as easy to interpret, direct activity measures. Remaining unresolved effects, such as metallicity and multiplicity of the investigated targets, are discussed in section 5.

4.1. Flare counts

The two clusters yielded a final distribution of 804 flares in total, of which flares in M45 contributed more than three times as many as M44 despite comparable total observation durations. The vast majority of stars and flares were found in the range of late-K to mid-M spectral types. Using these flares, we constructed 5 FFDs per cluster (Fig. 5): four for flares on stars with $T_{\mathrm{eff}} = 3,000 - 4,000$ K divided in 250 K bins, and one for the total sample. The choice of bins reflects the balance between the relatively low spectral type resolution, the number of flares per bin and the comparability of bins. Too few flares were detected on stars hotter than 4,000 K to allow for statistical interpretation.

At high energies we expect artifacts to contaminate the tail of the flare distributions. We inspected the LCs and TargetPixelFiles of all flares above 10^{33} erg using the `interact` function in `lightkurve` (Vinícius et al. 2018), a dedicated Kepler analysis package. We dropped three targets and 27 individual flares, listed in Table A.3 in A.

4.2. Flaring Activity Indicators

The contemporary approach to stellar flaring focuses on empirical studies rather than a robust theoretical description, so that it is unclear which indicator contains the most information about the process' physical underpinning (Benz 2016). There is no single, generally accepted flaring activity indicator per star or group of stars. We consider four parameters that describe flaring activity and allow us to compare to previous work. The flare rate FR and the energy fraction released in flares FA measure how often flares occur and how strong flaring activity is relative to the quiescent stellar flux. The power law exponent α and intercept β can be obtained from fitting a line to the log-log representation of a flare frequency distribution (FFD). α adds differential information to the average values covered by FA und FR. β depends on strongly on α but it can substitute for FR if α is assumed to be universal.

FR Among statistical flaring activity indicators flare rates are the most directly accessible. By treating a sample of stars in a narrow temperature range as a single prototype star we can define a flaring rate FR as the number of flares n_i recovered from all LCs i added up, divided by the sum of these LCs' observation times t_i:

$$FR = \frac{\sum_i n_i}{\sum_i t_i} \tag{9}$$

FR is a measure of average flare frequency. We treat the uncertainty on FR as Poissonian in the total number of flares.

FA The average total energy released in flares per time can be converted into an activity level measure when this energy is related to stellar bolometric luminosity $L_{\mathrm{bol},*}$. We consider, adapting the approach in Lurie et al. (2015), a flaring activity indicator FA that relates the total flare energy $E_{\mathrm{Kp,flare,tot}}$ released in the *Kepler* band relative to the estimated bolometric energy release $t_i \cdot L_{\mathrm{bol},*,i}$, where t_i is the observation time of that star:

$$FA = \frac{1}{N} \sum_i^N FA_i = \frac{1}{N} \sum_i^N \frac{E_{\mathrm{Kp,flare,tot},i}}{t_i \cdot L_{\mathrm{bol},*,i}} \tag{10}$$

N is the number of stars in a certain temperature bin, so that FA becomes an average. The uncertainty on FA is propagated in quadrature from the uncertainty on $E_{\mathrm{Kp,flare},i}$ and $L_{\mathrm{bol},*,i}$. We use bolometric luminosity instead of *Kepler* luminosity because we relate flaring energy to the total energy of a star, but *Kepler* energy contains a different fraction of bolometric energy depending on spectral type. However, we implicitly assume that the fraction of flare energy released in the *Kepler* band does not depend on spectral type, assuming flare production is a universal process.

FFD Using our synthetic flare injection procedure, we have validated a total of 804 flare candidates, 647 in M45, 157 in M44. In Fig. 5, we divide the sample into 250 K bins and fit power laws to cumulative FFDs of stars with similar spectral type as well as to the entire samples in both clusters. We drop all flare candidates with $E_{\mathrm{Kp,flare}}$ below the median detection threshold in each

Table 3. Power law parameters to the FFDs, as in eqn. 11. β_2 indicates a least-square fit with $\alpha \equiv 2$.

cluster	$T_{\min}$	$T_{\max}$	n_*	n_{flares}	α	$\log \beta$	$\log \beta_2$	$\log E_{\min}$
M44	3000	7000	766	155	2.214 ± 0.02	40.14 ± 0.70	$32.99\pm^{0.11}_{0.16}$	32.89
M44	3000	3249	206	22	2.047 ± 0.02	34.50 ± 0.90	$32.91\pm^{0.28}_{1.31}$	33.05
M44	3250	3499	164	55	2.131 ± 0.05	37.50 ± 1.79	$33.14\pm^{0.19}_{0.37}$	32.76
M44	3500	3749	152	88	2.021 ± 0.03	34.00 ± 1.00	$33.28\pm^{0.12}_{0.17}$	32.56
M44	3750	4000	47	13	2.012 ± 0.07	33.39 ± 2.61	$32.97\pm^{0.20}_{0.38}$	32.60
M45	3000	7000	737	596	2.159 ± 0.01	39.03 ± 0.53	$33.76\pm^{0.05}_{0.06}$	32.93
M45	3000	3249	224	94	2.051 ± 0.01	34.89 ± 0.65	$33.18\pm^{0.14}_{0.22}$	32.68
M45	3250	3499	195	262	2.137 ± 0.02	38.37 ± 0.92	$33.79\pm^{0.08}_{0.10}$	32.77
M45	3500	3749	130	179	2.369 ± 0.03	46.13 ± 1.21	$33.84\pm^{0.10}_{0.13}$	32.88
M45	3750	4000	47	53	2.149 ± 0.06	39.07 ± 2.20	$34.04\pm^{0.12}_{0.18}$	32.99
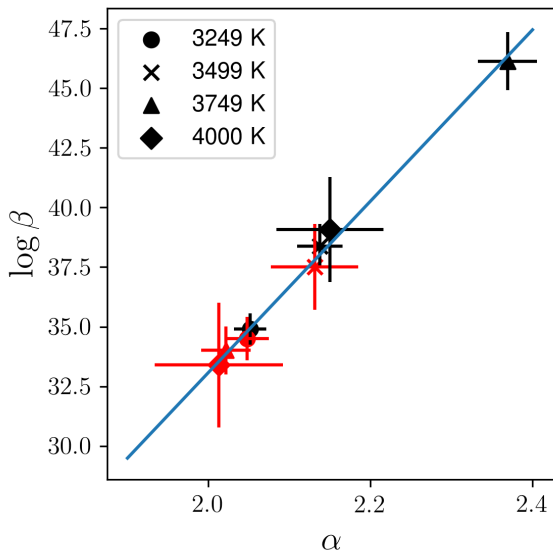
Fig. 4. Power law parameters to the FFDs, as in eqn. 11. See also Table 3.

bin. Consequently, the FFDs contain the contribution of $> 50\%$ of all targets with detected flare candidates in the entire considered energy range. Thus the cut accounts for the consequence of superimposing FFDs derived from LCs of different quality.

The occurrence rate of flares as a function of their energies, also termed flare frequency distribution (FFD; Lacy et al. 1976), can be written as

$$N(E)\mathrm{d}E = \beta E^{-\alpha}\mathrm{d}E. \tag{11}$$

In the log-log representation the above relation reads

$$\log N(E)\mathrm{d}E = [\log \beta - \alpha \log E]\mathrm{d}E. \tag{12}$$

In the corresponding cumulative distribution, a widely used representation in literature (see Audard et al. 2000; Hawley et al. 2014; Paudel et al. 2018, and references therein), the exponent becomes $\hat{\alpha} = \alpha - 1$. We fit the parameters given in eqn. 12 to the cumulative FFDs in Fig. 5. We used Orthogonal Distance Regression (ODR) to fit the line because it takes into account both uncertainties on flare rates and energies (see Hogg et al. 2010). The method yielded the best fit parameters $\hat{\alpha}$ and $\log\beta$, whose uncertainties we estimated using the delete-1 *jackknife* algorithm (Quenouille 1956).

Table 3 summarizes the resulting power law fit parameters for all FFDs. Overall, a power law slope $\alpha \approx 2.0 - 2.5$ is similar in most temperature bins and across both clusters, except for $T_{\mathrm{eff}} > 3,500\,\mathrm{K}$ in M44, where the slopes are steeper but also suffer from low flare numbers. In the total sample and, marginally, in some of the 250 K bins, we notice a departure from the single power law at high energies.

The values for β are typically dominated by α, as we can see in Fig. 4:

$$\log \beta = (33.1 \pm 1.3)\alpha - (33.6 \pm 2.9) \tag{13}$$

If we attribute at least some of the deviation from a single power law to pixel saturation, we expect α to be lower. If we then assume α to be universal, we can perform a fit with $\alpha \equiv 2$ and optimize only for the intercept $\log\beta_2$, Fig. 6 shows that β_2 clearly depends on age and less on mass. Overall, β_2 is roughly consistent with the trends in FR.

4.3. Aging of FA and FR

Both FA and FR decline with age (Fig. 7). In the time range from roughly ZAMS (125 Myr) to 630 Myr flaring activity decreases faster with higher T_{eff}. Stars in the 3,000-3,250 K bin follow the Skumanich $t^{-1/2}$ law most closely for both indicators, hotter stars' activity declines faster. Above the highest detection threshold we can compare activity levels for fixed ages at the cost of higher uncertainties, as shown in Fig. 8, or alternatively in the three bottom panels in Fig. 9. In this representation FR at young ages is similar for all T_{eff}. For FA, there is an initial spread where hotter stars have lower activity levels that also decline faster with age resulting in an increasing spread. Note that for FA the two lowest T_{eff} bins have very similar activity levels at young ages and have not diverged after 630 Myr.

Our values of FR and FA confirm previous work that suggests an age dependence of flaring activity either from kinematic ages (Kowalski 2009; Hilton 2011), other OCs like intermediate-age M37 compared to young clusters and the solar neighborhood (Chang et al. 2015), or indirectly by showing that rotation, which is an age indicator itself, predicts flaring activity levels (Doorsselaere et al. 2017). Clarke et al. (2018) searched *Kepler* time-domain photometry of 33 equal-mass wide binaries for flares. They analysed the relative luminosity emitted in flares on individual stars and found that it was similar within each system: Stars that are alike in mass, metallicity and age exhibit similar magnetic activity levels. As our sample clusters all have similar, close to solar-like metallicities (see Tab. 1) and all stars in a cluster share approximately the same age, comparing stars with
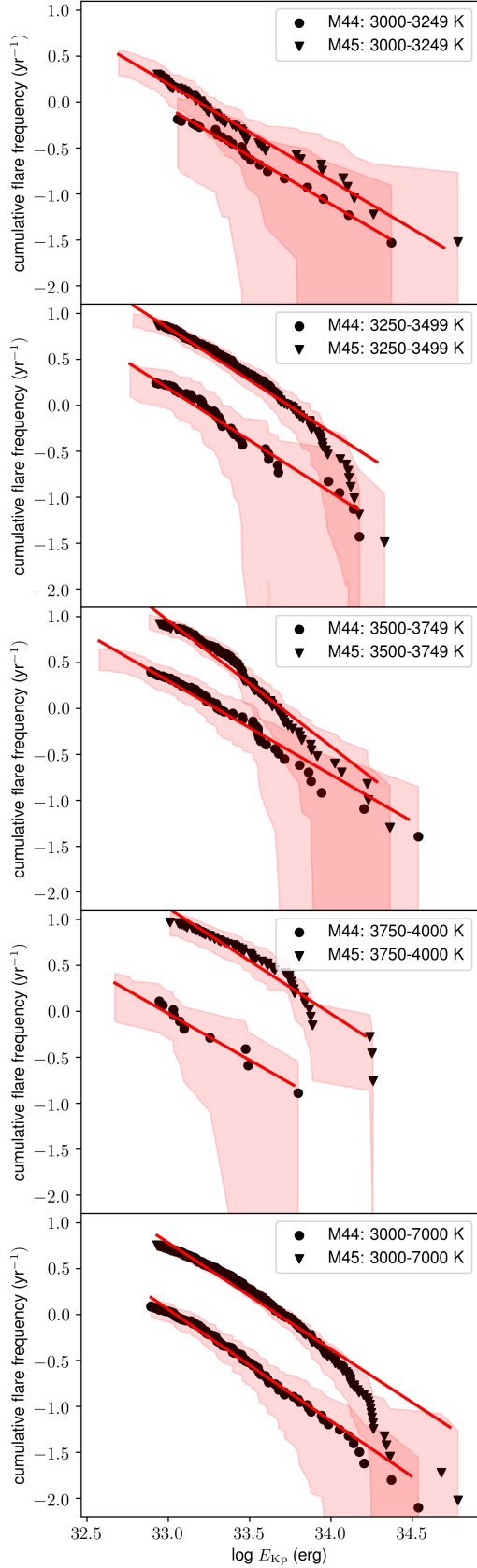
Fig. 5. Cumulative FFDs. Red lines: ODR fits. Red shadows: weighted Poissonian uncertainties on the frequency. Blue markers: flares vetted by eye. The low energy thresholds for are given in Table 3.
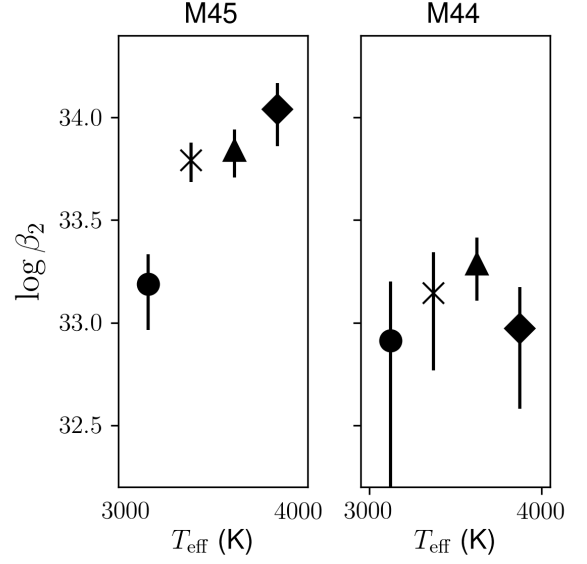

Fig. 6. Power law parameters to the FFDs, as in eqn. 11 with $\alpha \equiv 2$. See also Table 3.
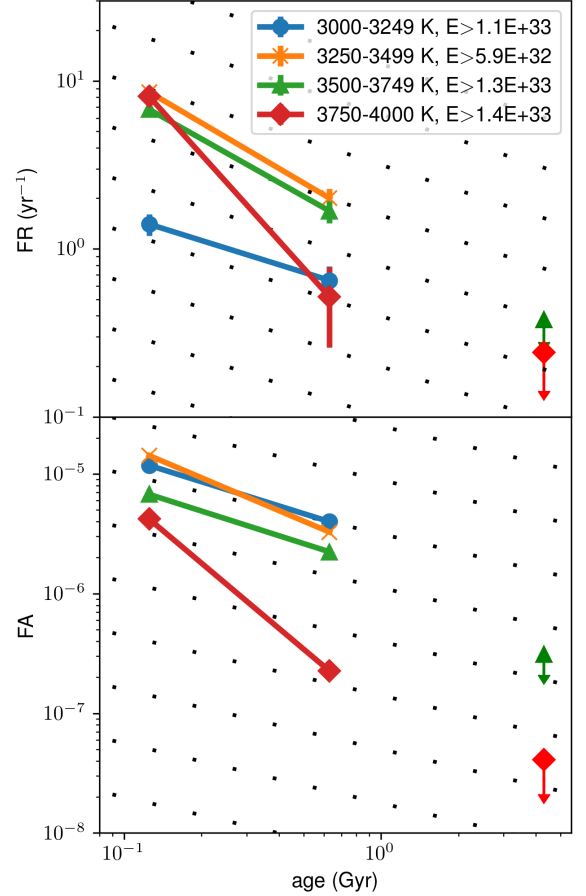

Fig. 7. Age dependent activity indicators. Color and line shapes distinguish different T_{eff} bins. Detection thresholds are derived from synthetic flare injections and averaged in each bin. For M67, available upper limits are given. The highest energy we detect is $\sim 2 \cdot 10^{34}$ erg. Top: flare rate *FR*; bottom: Flaring activity *FA* (see definitions in section 4.2). Grey dotted lines indicate the Skumanich age$^{-1/2}$ law for comparison.
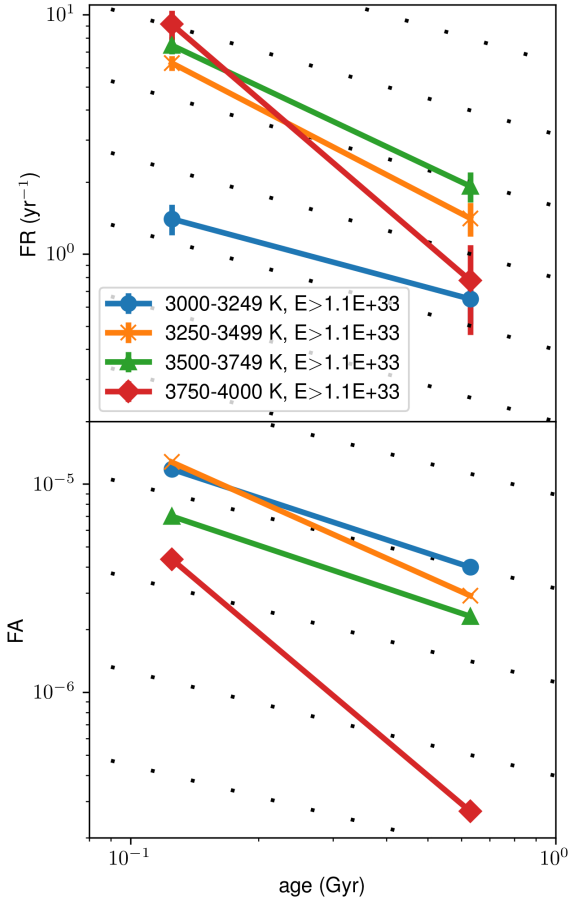
Fig. 8. Same as Fig. 7 but with equal detection thresholds for all T_{eff} bins.

similar masses (T_{eff}) between clusters should isolate the aging effect on flaring activity within the given uncertainties.

Recent work by Wright & Drake (2016) has shown that the saturation of X-ray emission L_X with increasing Rossby numbers Ro (rotation period divided by convective turnover time) is not due to the lack of a tachocline by finding fully convective M dwarfs in the non-saturated regime of the $L_X(Ro)$ relation. Since flaring activity is tightly correlated with X-ray emission (Neupert 1968; Crosby et al. 1993; Hannah et al. 2011), our results can be interpreted in the context of the conclusions from Wright & Drake (2016): All stars follow a hotter-stars-deplete-faster rule for both FA and FR in Fig. 8. M dwarfs, however, retain high activity levels for several hundred Myr, as Shkolnik et al. (2014) found in X-ray, NUV, and FUV observations. In our results, the lower flaring rates for the young and probably fully convective stars as well as the apparent saturation in FA hint at some regime change, possibly caused by the full convection boundary, and/or the transition from PMS to ZAMS to MS.

4.4. Flaring Activity Indicators as a Function of Mass

We can confirm that M dwarfs flare more often than K dwarfs, as Walkowicz et al. (2011) found in *Kepler* Quarter 1 data (see also Candelaresi et al. (2014) and Doorsselaere et al. (2017)). As an illustration: we find that $FR_{>4000\,\mathrm{K}}/FR_{<4000\,\mathrm{K}}$ is approximately 0.05 and 0.03 in M45 and M44, respectively. Our flaring rates for M44 are also consistent with work from Davenport

et al. (2012), who found that the time spent flaring increases for mid-M dwarfs compared to early-M type stars. This is not true, however for young M45 stars (Fig. 9, third panel from the top).

FR, FA, and FA per flare Flaring activity, quantified by FA and FR, is a function of stellar mass and age. A combination of these indicators is FA per flare, i.e. a measure of average individual flare energy. It is noteworthy that while both clusters show different levels of activity in the former two indicators, FA per flare is very similar for the cooler half of the sample and only begins to diverge for $T_{\mathrm{eff}} > 3,500\,\mathrm{K}$. Flares on mid-M type stars are the most violent in our sample. They release the highest energy fraction in fewest events (Fig.9) irrespective of age. While for even cooler, late-M to L type dwarfs, it is believed that the ionization fraction decreases and atmospheric density increases leading to higher resistivity and decoupling of magnetic fields from the atmosphere (relaxation of the frozen-in condition), such that the random walk of surface magnetic loop footprints no longer twists energy into the field topology (Mohanty et al. 2002; Rodriguez-Barrera et al. 2015), our sample of stars is not yet severely affected. We interpret that, in late-K to mid-M dwarfs, cooler atmospheres can store more energy in their surface magnetic field before a reconnection event is triggered, possibly because it produced larger starspots (Candelaresi et al. 2014), and/or is less complex (Garraffo et al. 2018) such that instabilities in magnetic topology are more difficult to cause.

The main caveat is the uncertainty on T_{eff}. Around $0.3\,M_{\odot}$ ($\sim M3.5$, $T_{\mathrm{eff}} \approx 3,250\,\mathrm{K}$; Hansen & Kawaler 1994; Delfosse et al. 2000) stars are thought transition to be fully convective, possibly altering their magnetic topology such that it impacts flare production (see, e.g. Reiners & Basri 2009). A $2,750 - 3,000\,\mathrm{K}$ bin we lack here would more certainly reside below the full convection boundary. We expect our T_{eff} estimates to be a good relative scale, but we shall be cautious about interpretations that involve the absolute parameters of our targets.

5. Discussion

5.1. Time Evolution of the Power Law Slopes

Shakhovskaya (1989) concluded in her seminal work that flaring activity is probably more a function of age than of spectral type and that this effect originates in a change in surface magnetic field as the stars spin down over time. She found values for $\alpha \sim 1.6$ in M44 and $\sim 1.8 - 2.0$ in M45 red dwarfs. More recently, Paudel et al. (2018) used kinematic ages of 10 UCDs to discover a weak trend towards shallower α in stellar flaring activity for older ages.

An evolution towards a shallower slope along with an overall decline in FA can be read as follows: Although the overall magnetic energy of a star dissipates over time, the surface magnetic field topology evolves such that longer buildup periods of magnetic stress can occur, caused by slower surface convection and/or rotation. This allows the production of more strong flares relative to weak ones[5]. A description of the concrete physical model of this interpretation is a worthwhile project of its own and should certainly involve, besides basic atmospheric parameters like density and temperature with their respective effects on ionization fractions, considerations of mass loss via flare associated CMEs (Drake et al. 2013; Alvarado-Gómez et al. 2018) and

---

[5] The argument works vice versa as well but with shorter buildup periods of magnetic stress that can no longer occur. This allows the production of fewer lower energy flares relative to stronger ones.
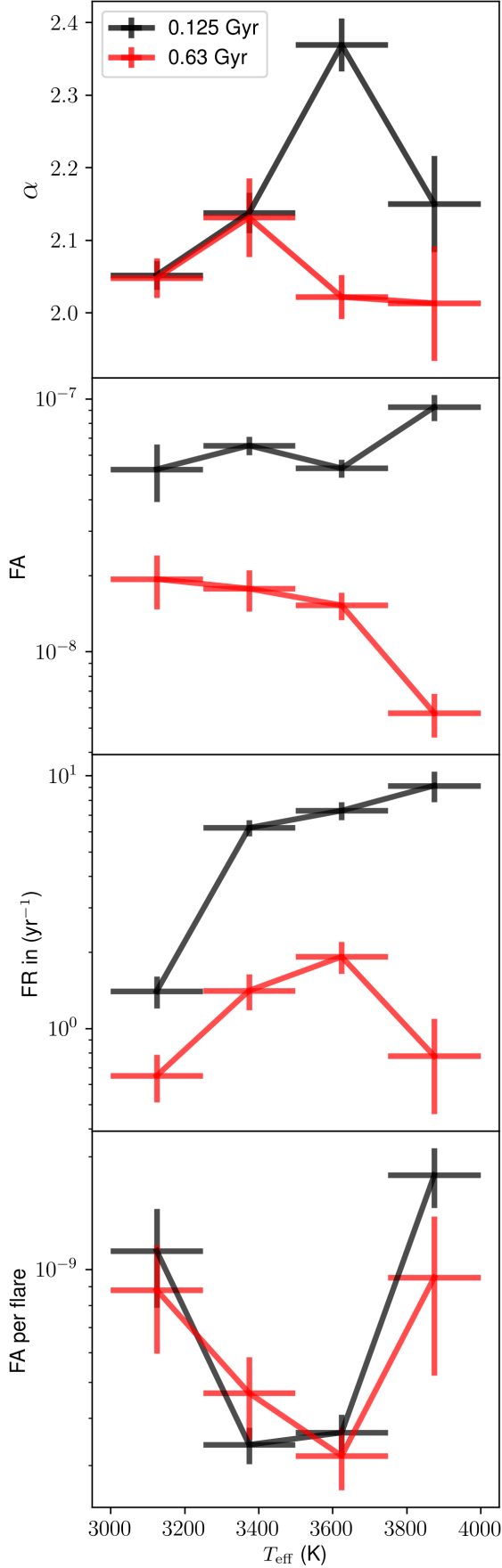
Fig. 9. Threshold for FA and FR $3.4 \cdot 10^{32}$ erg.

field complexity levels on observed bimodal spin-down modes, which are a topic of current research and theoretical studies by themselves (Barnes 2010; Newton et al. 2016; Garraffo et al. 2018). However, we find values of $\alpha \approx 2.0 - 2.5$ with no apparent age trend.

5.2. Flares as Coronal Heating Mechanisms in late-K to mid-M Dwarfs

Since the outset of statistical flare studies, the energy distributions of flares are most frequently described by power laws (Lacy et al. 1976). The function's exponent, i.e. the slope α in the log-log representation, is highly relevant for our understanding of stellar coronae. Flares have been proposed as coronal heating mechanisms for M dwarfs early by Doyle & Butler (1985), who found X-ray luminosity to be correlated with the average UV energy released in flares. If $\alpha \geq 2$, the total energy released in flares can be arbitrarily high (Güdel et al. 2002). On the Sun, flare energies can be as low as 10^{26} erg in microflares (Shimizu 1995)[6]. Given a sufficiently steep distribution, one can imagine very low energy flares to be the main coronal heating mechanism. For $\alpha < 2$, the highest energy flares contribute most to the total flare energy release and flares would be excluded from the discussion because high energy flares occur too rarely to be responsible for the quiescent coronal temperature.

Numbers found in literature all revolve around the value of 2: Güdel et al. (2003) observed AD Leo (dMe3.5) in EUV and soft X-ray and found $\alpha = 2 - 2.5$, Davenport et al. (2012) constrained α to be around 2 in M0-M6 dwarfs in red-optical and NIR in the *SDSS* and *2MASS* photometric surveys. Lurie et al. (2015) determined $\alpha \approx 2$ for two dMe5 dwarfs, Gizis (2017) found $\alpha = 1.8 \pm 0.2$ for for an M8 dwarf, and Paudel et al. (2018) studied time domain photometry of 10 L dwarfs in *K2* data and concluded that $\alpha < 2$ for the vast majority of investigated LCs.

As we find $\alpha \gtrsim 2$ for all temperatures and ages in our results, it seems possible that flares present the main coronal heating mechanism. However, given the proximity to the critical value and the unconclusive results in other work, the question has to remain open, at least for stars other than the Sun. For the latter, Hannah et al. (2011) conclude that microflares are not sufficient to heat the steady-state corona.

5.3. Departures from the Power Law Relation in FFDs

Flare production is a self-similar process with respect to the released energy because we find the FFDs to follow a power law in a broad energy range across different spectral types and stellar ages. A break in or a deviation from this power law would reflect a change in the underlying physics.

A single power law is not always a good fit to our data, as one can see from the full sample FFDs for both clusters in the bottom row in Fig. 5. Such deviations have already been noticed in single target FFDs, e.g. by Hawley et al. (2014) and Davenport (2016). Paudel et al. (2018) find a broken power law and a power law with an exponential cutoff at the highest energies a better fit to the FFDs of two L dwarfs in *K2* LCs. Mullan & Paudel (2018) argue that there could be more than one flare generating regime, i.e. that flares with small energies are fundamentally different from larger flares regarding e.g. the size of the reconnecting magnetic loop. Gershberg (2005) notes that both the instrument's sensitivity/saturation and a maximum energy a star of

---

[6] We do not consider nanoflares here, because they are produced by a different mechanism than classical flares (Benz 2016).

a certain type can release, may contribute to the deviations they observe for solar type stars.

Assuming that such a maximum energy $E_{\mathrm{max,flare}}$ exists for late K to mid-M dwarfs, we can argue that the departure from a single power law is the superposition of several stars' FFDs with different $E_{\mathrm{max,flare}}$. However, in contrast to our FFDs, the deviations discovered in previous work are associated with individual targets' FFDs. Within works that use, as expected, the same analysis methods for all targets in their samples (Paudel et al. 2018; Hawley et al. 2014; Davenport 2016) some but far from all single target FFDs show this type of departure. An undetected multiplicity, which would create the required superposition of FFDs, offers one explanation to this conspicuity.

However, to test this hypothesis or other physical interpretations, instrumental effects have to be entirely removed first, as Paudel et al. (2018) also admit. In our case, the deviations can also stem from outliers that were not rejected by `Appaloosa` and/or the de-trending procedure applied to the LCs before passing it on to the detection pipeline. While the algorithm removed many low-probability flares on the low-energy side, some outliers were kept and resulted in a deviation from the fitted power law, which can likely in part be attributed to the heterogeneous detection thresholds determined by the synthetic flare injection procedure. Furthermore, from by-eye inspection of LCs we expect some of the higher energy flare candidates to be artifacts, as `Appaloosa` does not perform well on LCs with strong peak-like systematics. We removed a handful of LCs in obvious cases, but several high-energy events may have underestimated $E_{\mathrm{Kp,flare}}$ due to pixel saturation.

As a result, high-energy outliers affect the slope for the entire sample, which would be shallower if we excluded them.

5.4. Metallicity

We expect metallicity to be a relevant parameter for flare activity studies, because it directly affects the atmosphere within which flaring takes place, but our sample can probably be treated as if metallicity was controled for.

On one hand, Gray et al. (2006) found that metal rich stars ([M/H] > −0.2) have a bimodal chromospheric activity distribution while lower metallicity stars show a single peaked spread. Karoff et al. (2018) suggested an effect of metallicity on stellar differential rotation and the undelying dynamo[7]. They point out that increasing the metallicity increases the opacity, which in turn will increase the temperature gradient. Then the criterion for convection will be satisfied deeper in the star (Schwarzschild 1906). Deeper convection zones lead to longer convective turnover times near the base of the outer convection zones (Brun et al. 2017). Stronger differential rotation, a key parameter in the classical $\alpha\Omega$-dynamo, is the consequence (Bessolaz & Brun 2011).

On the other hand, magnetic activity in wide binaries, that naturally have the same metallicities, is similar within the pairs, as Clarke et al. (2018) show for 33 equal-mass binaries found in *Kepler* photometry. Thus, the limitation of our OC sample to solar like metallicity clusters dashes joy with pain: We cannot offhand extrapolate the determined age-activity-mass-relation to higher or lower metallicities, but we can mostly exclude that the different activity levels observed in the clusters are an effect of metallicity rather than age: In our sample, M45 has close to solar metallicity, while M44 is even more metal-rich, see also Tab. 1,

---

[7] Here, Karoff et al. (2018) only discuss the classical, tachocline-dependent dynamo paradigm.

Netopil et al. 2016). Unfortunately, the evidence found by Karoff et al. (2018) does not allow to precisely quantify the effect for M44 because they only compare two solar-type stars. From Gray et al. (2006) we can follow that we can neglect the differences in [Fe/H] because all of our studied clusters clearly fall into the high-metallicity regime.

5.5. Multiplicity

Multiplicity is ubiquitous among solar-type and lower mass stars affecting $\sim 50\%$ of all systems (Duquennoy & Mayor 1991; Fischer & Marcy 1992) and causes the misattribution of flares from secondaries to the primary. Bouvier et al. (2001) observe a binary frequency of about $\sim 50\%$, independent of age of the cluster by comparing M44, M45 and a star forming 2 Myr old cluster IC 348 for G and K type stars. The frequency is lower for M dwarfs ($42 \pm 9\%$; Fischer & Marcy 1992) due to the decreasing mass range for companions, and keeps declining towards lower masses, as Boudreault et al. (2012) found for the $0.07 < M_\odot < 0.45$ range for M44, but does not fall below $\sim 17\%$ within uncertainties for any mass bin they studied. For M67, Geller et al. (2015) estimated the binary frequency to be as high as $57 \pm 4\%$. Thus, the binary fraction considerably increases the number of individual stars in all clusters for low mass stars, similarly for M44 and M45, and even more notably in M67, where higher mass stars prevail. The effects on both the total activity level and the flare energy distributions are twofold:

– The true total flaring rate is smaller, as the true number of stars and hence the cumulated observation time is significantly higher than the number of targets.
– Energies of flares in unresolved binaries are underestimated, regardless of whether they occur on the larger or smaller star, because their *EDs* are measured relative to the quiescent luminosity of the whole system, but are multiplied only by the luminosity of the large companion that dominates the color indices, while it should be the sum of both (see eqn. 4).

These considerations imply

– a reattribution of flaring activity levels to cooler stars within the $3,000 − 4,000\,\mathrm{K}$ range. This could mean, for instance, that the single flare detected for the $T_{\mathrm{eff}} = 3780\,\mathrm{K}$ (M0.5−M1; Pecaut & Mamajek 2013, classified as M67 member (M) by Gonzalez 2016) star may stem from an even cooler dwarf with probability $\sim 29\%$.
– an additional systematic offset to the location of the full convection boundary as may be marked by flares, which would reside at higher assigned spectral types because of the misattribution.

We note that the dependence of magnetic activity on the presence of close companions may also affect flaring activity measurements, as discussed by Gao et al. (2016). A quantitative comparison of flaring activity in single and binary members is beyond the scope of this paper and requests information about both confirmed binary and single stars.

6. Summary and Conclusions

Using `K2SC` de-trended *Kepler/K2* LCs we investigated the flaring activity of two solar metallicity OCs, the ZAMS cluster M45 and the intermediate age M44, a total of more than 250 years of cumulative observation time at 30 min cadence from $1,638$ targets, mostly late-K and early- to mid-M dwarfs.

Pan-STARRS and *2MASS* multiband photometry yielded $T_{\rm eff}$ and radii of individual stars, using solar metallicity standards (Pickles 1998), computed by Covey et al. (2007), and color-temperature relations from Pecaut & Mamajek (2013). From these we derived quiescent luminosities and the *Kepler* band energies of flares detected by the flare finding and analysis pipeline `Appaloosa` (Davenport 2016).

We improved `Appaloosa`'s performance using GP Regression de-trended and variablity cleared LCs from Aigrain et al. (2016) compared to raw data. We introduced a synthetic flare injection and recovery routine to characterize long cadence K2SC de-trended LCs of heterogeneous quality. Over 9 million artificial flare signatures were injected and recovered. They yielded both *ED*-dependent detection thresholds and correction factors to the sampling induced systematic energy underestimation.

We found 760 flare candidates with $E_{\rm Kp,flare}$ ranging from $5 \cdot 10^{31}$ erg to 10^{34} erg in two clusters, of which 647 belong to M45 and 201 to M44. From the flare rates (*FR*) and energy released relative to bolometric luminosity (*FA*) as a function of $T_{\rm eff}$ and age we conclude:

– Both *FA* and *FR* substantially decline with age for late-K to mid-M dwarfs and follow a hotter-stars-deplete-faster rule. Cooler stars retain their elevated activity for several hundred Myr.

– *FA* per flare is higher for $T_{\rm eff} < 3,500$ K possibly due to decreasing surface magnetic field complexity, that leads to relatively higher energies stored in magnetic fields, which in turn are less likely to get destabilized by surface field motions. An analysis of both cooler and hotter stars will show if this trend substantiates and could be attributed to the transition to fully convective interiors.

Assuming that our FFDs can be described by single power laws for all flare energies, it remains unclear if flares can be the main coronal heating mechanism because the power law exponents α are around or above the critical value of 2. A single power law is not always a good fit to the FFDs, although it well describes the distributions in most individual temperature bins. This may be caused by saturation effects at the FFDs' high-energy ends. Alternatively, a high energy flaring limit that varies among stars could produce the broken power law by superimposing their FFDs.

We confirm that we can treat flaring activity measurements as age tracers for late-K to mid-M dwarfs. Our findings back previous suggestive evidence that a flaring-temperature-age relation exists. Although we do not find an age dependence in the FFD power law exponent α, we observe a distinct age dependence for various other activity measures that could prove to be useful age-dating techniques for low-mass stars, complementing existing methods. The lack of any trend in α suggests the universality of the flare production process across a broad range of masses, presumably below the full convection boundary.

Besides M45 and M44, *K2* has endowed us with months of continuous monitoring of several OCs, spanning a range of ages from PMS to solar age. We will expand the analysis described in this paper to these targets in an upcoming project, and explore the effects of multiplicity and metallicity on the gears of a stellar flaring "clock".

References

Aigrain, S., Parviainen, H., & Pope, B. J. S. 2016, MNRAS, 459, 2408
Alvarado-Gómez, J. D., Drake, J. J., Cohen, O., Moschou, S. P., & Garraffo, C. 2018, ApJ, 862, 93
Audard, M., Gudel, M., Drake, J. J., & Kashyap, V. L. 2000, ApJ, 541, 396
Auvergne, M., Bodin, P., Boisnard, L., et al. 2009, A&A, 506, 411
Barnes, S. A. 2003, ApJ, 586, 464
Barnes, S. A. 2010, ApJ, 722, 222
Barnes, S. A., Weingrill, J., Fritzewski, D., Strassmeier, K. G., & Platais, I. 2016, ApJ, 823, 16
Bell, C. P. M., Naylor, T., Mayne, N. J., Jeffries, R. D., & Littlefair, S. P. 2012, MNRAS, 424, 3178
Benedict, G. F., Henry, T. J., Franz, O. G., et al. 2016, AJ, 152, 141
Benz, A. O. 2016, Living Reviews in Solar Physics, 14, 2
Benz, A. O. & Güdel, M. 2010, ARA&A, 48, 241
Bessolaz, N. & Brun, A. S. 2011, ApJ, 728, 115
Boudreault, S., Lodieu, N., Deacon, N. R., & Hambly, N. C. 2012, MNRAS, 426, 3419
Bouvier, J., Duchêne, G., Mermilliod, J.-C., & Simon, T. 2001, A&A, 375, 989
Bouy, H., Bertin, E., Sarro, L. M., et al. 2015, A&A, 577, A148
Brun, A. S., Strugarek, A., Varela, J., et al. 2017, ApJ, 836, 192
Candelaresi, S., Hillier, A., Maehara, H., Brandenburg, A., & Shibata, K. 2014, ApJ, 792, 67
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, ArXiv e-prints [arXiv:1612.05560]
Chang, S.-W., Byun, Y.-I., & Hartman, J. D. 2015, ApJ, 814, 35
Clarke, R. W., Davenport, J. R. A., Covey, K. R., & Baranec, C. 2018, ApJ, 853, 59
Covey, K. R., Ivezić, Ž., Schlegel, D., et al. 2007, AJ, 134, 2398
Crosby, N. B., Aschwanden, M. J., & Dennis, B. R. 1993, Sol. Phys., 143, 275
Dahn, C. C., Harris, H. C., Subasavage, J. P., et al. 2017, AJ, 154, 147
Davenport, J. R. A. 2016, ApJ, 829, 23
Davenport, J. R. A., Becker, A. C., Kowalski, A. F., et al. 2012, ApJ, 748, 58
Davenport, J. R. A., Hawley, S. L., Hebb, L., et al. 2014, ApJ, 797, 122
Delfosse, X., Forveille, T., Ségransan, D., et al. 2000, A&A, 364, 217
Dias, W. S., Monteiro, H., Caetano, T. C., & Oliveira, A. F. 2012, A&A, 539, A125
Dieterich, S. B., Henry, T. J., Jao, W.-C., et al. 2014, AJ, 147, 94
Doorsselaere, T. V., Shariati, H., & Debosscher, J. 2017, ApJS, 232, 26
Douglas, S. T. 2016, ApJ, 822, 47
Douglas, S. T., Agüeros, M. A., Covey, K. R., et al. 2014, ApJ, 795, 161
Douglas, S. T., Agüeros, M. A., Covey, K. R., & Kraus, A. 2017, ApJ, 842, 83
Doyle, J. G. & Butler, C. J. 1985, Nature, 313, 378
Drake, J. J., Cohen, O., Yashiro, S., & Gopalswamy, N. 2013, ApJ, 764, 170
Duquennoy, A. & Mayor, M. 1991, A&A, 248, 485
Eker, Z., Soydugan, F., Soydugan, E., et al. 2015, AJ, 149, 131
Esselstein, R., Aigrain, S., Vanderburg, A., et al. 2018, ApJ, 859, 167
Filippazzo, J. C., Rice, E. L., Faherty, J., et al. 2015, ApJ, 810, 158
Finkbeiner, D. P., Schlafly, E. F., Schlegel, D. J., et al. 2016, ApJ, 822, 66
Fischer, D. A. & Marcy, G. W. 1992, ApJ, 396, 178
Gao, Q., Xin, Y., Liu, J.-F., Zhang, X.-B., & Gao, S. 2016, ApJS, 224, 37
Garraffo, C., Drake, J. J., Dotter, A., et al. 2018, ApJ, 862, 90
Geller, A. M., Latham, D. W., & Mathieu, R. D. 2015, AJ, 150, 97
Gershberg, R. E. 2005, Solar-Type Activity in Main-Sequence Stars (Springer Berlin Heidelberg)
Gizis, J. E. 2017, ApJ, 845, 33
Gonzalez, G. 2016, MNRAS, 459, 1060
Gray, R. O., Corbally, C. J., Garrison, R. F., et al. 2006, AJ, 132, 161

Güdel, M., Audard, M., Kashyap, V., Drake, J. J., & Guinan, E. F. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 277, Stellar Coronae in the Chandra and XMM-NEWTON Era, ed. F. Favata & J. J. Drake, 491

Güdel, M., Audard, M., Kashyap, V. L., Drake, J. J., & Guinan, E. F. 2003, ApJ, 582, 423

Hannah, I. G., Hudson, H. S., Battaglia, M., et al. 2011, Space Sci. Rev., 159, 263

Hansen, C. J. & Kawaler, S. D. 1994, Stellar Interiors. Physical Principles, Structure, and Evolution. (Springer Berlin Heidelberg New York)

Hawley, S. L., Davenport, J. R. A., Kowalski, A. F., et al. 2014, ApJ, 797, 121

Hawley, S. L. & Fisher, G. H. 1992, ApJS, 78, 565

Hilton, E. J. 2011, PhD thesis, University of Washington

Hilton, E. J., West, A. A., Hawley, S. L., & Kowalski, A. F. 2010, AJ, 140, 1402

Hogg, D. W., Bovy, J., & Lang, D. 2010, ArXiv e-prints [arXiv:1008.4686]

Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398

Hunt-Walker, N. M., Hilton, E. J., Kowalski, A. F., Hawley, S. L., & Matthews, J. M. 2012, PASP, 124, 545

Jackson, R. J., Deliyannis, C. P., & Jeffries, R. D. 2018, MNRAS, 476, 3245

Jenkins, J. M. 2010, ApJ, 713, L120

Kaltcheva, N., Golev, V., & Paunzen, E. 2017, PASP, 129, 114201

Karoff, C., Metcalfe, T. S., Santos, Â. R. G., et al. 2018, ApJ, 852, 46

Kawaler, S. D. 1988, ApJ, 333, 236

Kirkpatrick, D. J., Cushing, M. C., Gelino, C. R., et al. 2011, ApJS, 197, 19

Kirkpatrick, J. D., Reid, I. N., Liebert, J., et al. 1999, ApJ, 519, 802

Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, ApJ, 713, L79

Kowalski, A. F. 2009, AJ, 138, 633

Kowalski, A. F., Hawley, S. L., Wisniewski, J. P., et al. 2013, ApJS, 207, 15

Kraus, A. L. & Hillenbrand, L. A. 2007, AJ, 134, 2340

Kretzschmar, M. 2011, A&A, 530, A84

Lacy, C. H., Moffett, T. J., & Evans, D. S. 1976, ApJS, 30, 85

Leggett, S. K., Morley, C. V., Marley, M. S., & Saumon, D. 2015, ApJ, 799, 37

Lépine, S., Hilton, E. J., Mann, A. W., et al. 2013, AJ, 145, 102

Lodieu, N., Deacon, N. R., & Hambly, N. C. 2012, MNRAS, 422, 1495

Lorenzo-Oliveira, D., Porto de Mello, G. F., & Schiavon, R. P. 2016, A&A, 594, L3

Luger, R., Agol, E., Kruse, E., et al. 2016, AJ, 152, 100

Lurie, J. C., Davenport, J. R. A., Hawley, S. L., et al. 2015, ApJ, 800, 95

Maehara, H., Shibayama, T., Notsu, S., et al. 2012, Nature, 485, 478

McQuillan, A., Mazeh, T., & Aigrain, S. 2014, VizieR Online Data Catalog, 221

Meibom, S., Barnes, S. A., Platais, I., et al. 2015, Nature, 517, 589

Mestel, L. 1984, in Cool Stars, Stellar Systems, and the Sun, ed. S. L. Baliunas & L. Hartmann, Vol. 193 (Berlin, Heidelberg: Springer Berlin Heidelberg), 49–59

Mohanty, S., Basri, G., Shu, F., Allard, F., & Chabrier, G. 2002, ApJ, 571, 469

Montet, B. T., Tovar, G., & Foreman-Mackey, D. 2017, ApJ, 851, 116

Mullan, D. J. & Paudel, R. R. 2018, ApJ, 854, 14

Nardiello, D., Libralato, M., Bedin, L. R., et al. 2016, MNRAS, 455, 2337

Netopil, M., Paunzen, E., Heiter, U., & Soubiran, C. 2016, A&A, 585, A150

Neupert, W. M. 1968, ApJ, 153, L59

Newton, E. R., Irwin, J., Charbonneau, D., et al. 2016, ApJ, 821, 93

Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K., & Vaughan, A. H. 1984, ApJ, 279, 763

Pace, G. 2013, A&A, 551, L8

Parker, E. N. 1979, Cosmical Magnetic Fields: Their Origin and Their Activity (Oxford, Clarendon Press; New York, Oxford University Press)

Patel, R. I., Metchev, S. A., & Heinze, A. 2014, ApJS, 212, 10

Patten, B. M. & Simon, T. 1996, ApJS, 106, 489

Paudel, R. R., Gizis, J. E., Mullan, D. J., et al. 2018, ApJ, 858, 55

Pecaut, M. J. & Mamajek, E. E. 2013, ApJS, 208, 9

Pickles, A. J. 1998, PASP, 110, 863

Piskunov, A. E., Kharchenko, N. V., Schilbach, E., et al. 2011, A&A, 525, A122

Priest, E. & Forbes, T. 2002, A&A Rev., 10, 313

Queloz, D., Allain, S., Mermilliod, J.-C., Bouvier, J., & Mayor, M. 1998, A&A, 335, 183

Quenouille, M. H. 1956, Biometrika, 43, 353

Radick, R. R., Thompson, D. T., Lockwood, G. W., Duncan, D. K., & Baggett, W. E. 1987, ApJ, 321, 459

Rebull, L. M., Stauffer, J. R., Bouvier, J., et al. 2016, AJ, 152, 113

Reiners, A. & Basri, G. 2009, A&A, 496, 787

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Rodriguez-Barrera, M. I., Helling, C., Stark, C. R., & Rice, A. M. 2015, MNRAS, 454, 3977

Sarro, L. M., Bouy, H., Berihuete, A., et al. 2014, A&A, 563, A45

Schaefer, B. E., King, J. R., & Deliyannis, C. P. 2000, ApJ, 529, 1026

Schatzman, E. 1962, Annales d'Astrophysique, 25, 18

Schneider, A. C., Cushing, M. C., Kirkpatrick, J. D., et al. 2015, ApJ, 804, 92

Schwarzschild, K. 1906, Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, 195, 41

Shakhovskaya, N. 1989, Sol. Phys., 121, 375

Shibayama, T., Maehara, H., Notsu, S., et al. 2013, ApJS, 209, 5

Shimizu, T. 1995, Publications of the ASJ, 47, 251

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Soderblom, D. R., Duncan, D. K., & Johnson, D. R. H. 1991, ApJ, 375, 722

Van Cleve, J. E. & Caldwell, D. A. 2016, Kepler Instrument Handbook (KSCI-19033-002)

Van Cleve, J. E., Howell, S. B., Smith, J. C., et al. 2016, PASP, 128, 075002

Van Saders, J. L., Ceillier, T., Metcalfe, T. S., et al. 2016, Nature, 529, 181

Vidotto, A. A., Gregory, S. G., Jardine, M., et al. 2014, MNRAS, 441, 2361

Vinícius, Z., Barentsen, G., Hedges, C., Gully-Santiago, M., & Cody, A. M. 2018

Walkowicz, L. M., Basri, G., Batalha, N., et al. 2011, AJ, 141, 50

West, A. A., Morgan, D. P., Bochanski, J. J., et al. 2011, AJ, 141, 97

Wright, N. J. & Drake, J. J. 2016, Nature, 535, 526

Wright, N. J., Drake, J. J., Mamajek, E. E., & Henry, G. W. 2011, ApJ, 743, 48

Yadav, R. K. S., Bedin, L. R., Piotto, G., et al. 2008, A&A, 484, 609

Yang, H., Liu, J., Gao, Q., et al. 2017, ApJ, 849, 36

Yang, H., Liu, J., Qiao, E., et al. 2018, ApJ, 859, 87

Yee, S. W., Petigura, E. A., & von Braun, K. 2017, ApJ, 836, 77

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Appendix A: Excluded Data

Cosmic Ray Contamination Cosmic ray (CR) events occur randomly, possibly adding an offset to the flaring rate of a star because they may be mimicking flare signatures. Unlike the original *Kepler* mission, *K2* does not identify and remove cosmic ray signals in optimal aperture due to the reduced calibration and photometric analysis pipelines. CRs in pixels collateral to optimal aperture, however, are flagged and can be used to extract information about the data quality. In Campaigns 0-11 a CR is flagged if the counts in a pixel exceed the 4σ limit compared to a rolling band median of five cadences. In our sample, CR flags make up $\sim 4 - 7\,\%$ of all data points depending on the campaign. Note that from Campaign 11 on the threshold was increased to 7σ to minimize false detections triggered by K2's roll motion[8].
A CR hitting the optimal aperture or spilling signal into it from collateral pixels and signals stemming from flares are possibly ambiguous in long cadence data. Both produce a sudden increase in flux. A higher temporal resolution mitigates the problem: True flares can be identified by the characteristic shape of the flare signal with its impulsive rise and decline followed by a more gradual decay phase. For flares with durations shorter than $\sim 1.5\,$h, or 3 data points, *K2* long cadence data do not resolve any of these features and we do not include any candidates with less than 3 consecutive data points in our analysis (i.e. $N_3 \geq 3$, see section 3). However, a flare that occurs simultaneously with a CR impact in a sufficiently close spot on the CCD will erroneously increase the detected flare energy, amplitude and/or duration.
A CR induced flaring rate offset can vary, particularly when the spacecraft drifts and changes its attitude towards the Sun until it rotates again for a new campaign. Different campaigns and the relative position of the target on the CCD can play a role: The spacecraft may shield certain parts of the detector from CR impact better than others.
In this work, we acknowledge the possible CR contamination in our data but treat them as a statistical offset that affects all light curves in a similar manner: We interpret the similar amounts of CR flags we found in all OC samples (see Tab. A.2) to indicate comparability of the OCs. The weak trend may even reflect true flaring activity differences in the sample.

Thruster Firings and Field-Wide Systematics While we keep CR flagged data points, we remove those labeled as coinciding with the spacecraft's thruster firings as well as their preceding and succeeding measurements that are often flagged as CRs. This procedure, however, excludes $11 - 18\,\%$ of all data points from the individual LCs (see Tab. A.2).
Even if the de-trending procedure removes most systematic effects from the LC, some may remain undetected. To identify false positive detections, we compare all LCs and their respective flare candidates to each other in order to find field-wide systematics. We assume that flares occur randomly and reject as precautionary measure all flare detections that are detected in more LCs simultaneously than this assumption would allow.

Saturated targets According to Van Cleve & Caldwell (2016), if the flux in a pixel exceeds the full well depth of 10093 DN by a factor < 10, the detector response may still be linear if one considers the total flux in a sufficiently large aperture. We

---

excluded 24 targets from our analysis, where the detected flare candidates exceed this threshold, listed in Tab. A.1.

Table A.1. Excluded data. Extremely saturated targets.

cluster	EPIC ID
M45	210877423
	210879932
	210966700
	210978650
	210997197
	211014186
	211020371
	211033155
	211053737
	211063235
	211066337
	211067634
	211072441
	211086025
	211087059
	211093684
	211096368
	211101761
	211115638
	211120842
	211125210
	211132233
M44	211934056
	212034371

Table A.2. Excluded data. We removed both possible and confirmed thruster firings from the data, CR flags have not been removed except for those adjacent to or overlapping with thruster firing flags. C: campaign. % thruster firings: percentage of data points flagged as (possible) thruster firings. $\overline{\%CR}$: average contamination of LCs with CR flags. Field-wide: regions excluded due to simultaneous flares (total number of data points they comprise).

cluster	C	% thruster firings	$\overline{\%CR}$	field-wide
M45	4	3.85	6.8 ± 2.75	50 (251)
M44	5	4.53	5.2 ± 3.24	7 (28)
M67	5	4.53	4.0 ± 1.60	1 (6)

Appendix B: Synthetic Flare Injection and Recovery

Improvements to the Routine We carried out several enhancements to the original artificial flare injection subroutine to improve the validity of the returned recovery rates. The original code injected 100 synthetic flares into a LC at once, a number too small to characterize the whole LC but too large to create mostly single and not superimposed events. We introduced an artificial flare grid in the amplitude-duration space and conducted more injections with fewer flares instead. We assumed that a LC's properties are uniform throughout the campaign during which it was recorded. The convergence of the recovery probability distribution was a reasonable expectation because the flare finding procedure is deterministic, i.e. the same fake flare infested LC yields the same recovered flares. Note that artificial flare injection cannot take care of false positive signals. We addressed the issue by manually vetting high-energy events, and comparing multiple LCs to identify field-wide systematics.

Table A.3. Flare candidates detected in targets with given EPIC ID in the time interval $[t_0, t_f]$ and rejected after manual inspection of LCs and TargetPixelFiles. If t_0, t_f are not given, the entire target was dropped.

EPIC	t_0	t_f	note
210966700	2258.671619	2258.732914	flagged cadences
211038389	2251.643200	2251.867949	entire aperture flashes
211038389	2252.051835	2252.215290	no flares in LC
211060530			bright CCD row
211066337	2279.471032	2279.532327	no clear brightening
211066337	2236.993468	2237.034332	no clear brightening
211073598			detector edge
211089323			saturated
211091848	2243.245708	2243.286571	no clear brightening
211096368	2274.996528	2275.078254	bright CCD row
211114329	2271.094166	2271.135030	no clear brightening
211120842	2243.654339	2243.695202	dropout CCD row
211909748	2317.909660	2317.950524	flagged cadences
211936906	2336.012180	2336.093907	asteroid transit
211955036	2338.463959	2338.484391	saturated
211955036	2338.647844	2338.770434	saturated
211972627			broken pixel
211975426	2331.435440	2331.476304	asteroid transit
212034371	2316.070632	2316.111496	no clear brightening
212034371	2338.361678	2338.422973	contamination
212034371	2343.857791	2343.898654	no clear brightening

We could directly infer the recovery probability for the injected flare *ED*s. Observed flares' *ED*s, however, are systematically underestimated in data with low time sampling (30 min) compared to 1 min cadence LCs, as Yang et al. (2018) found for *Kepler* data. Since we only used long cadence LCs in this work, we compared the injected with the recovered *ED* for every artificial flare. From the ratio, i.e., the recovered share of the injected energy, we derived a correction factor that allowed us to ascribe both a more realistic flare energy and the corresponding recovery rate to each candidate (see Fig. B.1).

In summary, we obtained two *ED*-dependent corrections to the *ED*s and the recovered flare rates for each flare: an *ED* correction factor and a correction to the Poissonian count uncertainty, respectively. We corrected for the systematic energy underestimation on a LC by LC basis (Fig. B.1). Each LC obtained a unique flare recovery relation from the series of synthetic flare injections as a function of injected *ED*. This detection rate of synthetic events revealed the detection threshold for flares and provided a means to account for flares detected within the transition region by increasing the uncertainty on the detected number of flares accordingly (Fig. B.2).

For M dwarfs, Hawley et al. (2014) found clear evidence for an exponential distribution of flare amplitudes as a function of duration. We enabled users of `Appaloosa` to generate artificial flares generously covering this empirical parameter space (Fig. 2). The main advantage is a better coverage of the low-energy/low-detectability end of the expected flare distribution which is also most densely populated.

Realistic Synthetic Flare Injections   Caution is recommended regarding the universality of injected flare shapes. Although the self-similarity of flares with different durations and amplitudes for a large parameter range is striking, the semi-analytical flare model used to generate synthetic events is derived from the flare distribution of a single star, GJ 1243 (Davenport et al. 2014), a bright dMe flare star in the northern sky (Lépine et al. 2013). Fur-

thermore, Davenport et al. (2014) found a small portion (1.3 %) of flares on that star that could only poorly be fitted by this single star model. The injection procedure is therefore tailored to characterize a LC's quality with respect to the most common flares on GJ 1243. It can be extended to treat superpositions of classical flares, i.e. multiflare events, that can amount to 15 % of all events (Davenport et al. 2014) and are occasionally created, but complex shapes cannot be represented by the injections, although they can be recovered by the detection routine we employ here. We acknowledge that there is room for improvement in this respect.

Appendix B.0.1: Light Curve Quality from Synthetic Flare Injections

Relying on LCs alone for the detection of flare candidates requires detailed quality assessment of each LC. We therefore injected and recovered 20 randomly distributed flares into every LC and repeated the process 300 times each using `Appaloosa`. Of these injections $\sim$ 5 %, were dropped, because they were juxtaposed with more energetic events yielding more complex signatures. Consequently, $\sim$ 5,700 synthetic flares injected into each LC − a total of more than 9 million events to cover all LCs − are used to retrieve information about their quality with respect to thresholds of detection, and systematic errors on energy recovery. For each cluster, this results in individual recovery rates based on the properties of the de-trended LCs, e.g. noise levels, time resolution, or variability time scales and magnitudes. More properties may be implicitly covered, because the procedure does not make any a priori assumptions about what affects flare signature detectability and to what degree. Overall, flare energies are systematically underestimated with significant uncertainties that become smaller with increasing energy (see Fig. B.1). Recovery rates typically exceed 80 % as soon as some LC specific threshold for detection is exceeded, e.g. the time resolution limit. However, 100 % recovery of both events and their
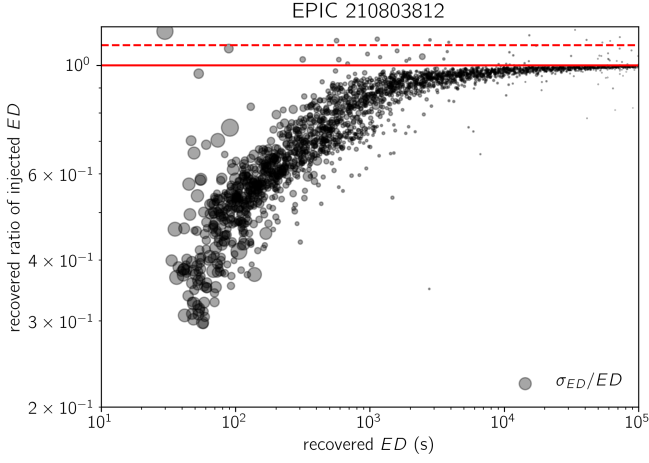
Fig. B.1. Example of recovered energy ratios from synthetic flare injections. Ratio of the originally injected *ED* recovered by `Appaloosa` in the K2SC de-trended LC light curve for EPIC 210803812. The circle size represents the calculated relative uncertainty σ_{ED}/ED. The red lines are placed at 100 % (filled) and 110 % (dashed) energy recovery, given for orientation.
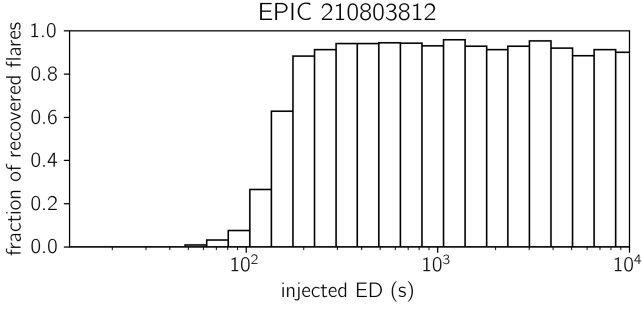

Fig. B.2. Recovery rate of classic flare events as a function of *ED* for EPIC 210803812.

Table B.1. Average uncertainties propagating to flare energies (and their standard deviations). LCs: number of light curves in final analysis. Flares: number of flare candidates used for final analysis.

	LCs	flares	$\overline{\sigma_{ED}}(\sigma)$	$\overline{\sigma_{L_{bol}}}(\sigma)$	$\overline{\sigma_{E_{Kp,flare}}}(\sigma)$
M45	723	547	0.29(0.52)	0.55(0.27)	0.66(0.52)
M44	769	157	0.26(0.24)	0.53(0.27)	0.63(0.24)
M67	258	0	-(-)	-(-)	-(-)

energies is never achieved due to the cutoff of continuous observation periods (Fig. B.2).