

Search for flares and associated CMEs on late-type main-sequence stars in optical SDSS spectra[★]

Florian Koller¹, Martin Leitzinger¹, Manuela Temmer¹, Petra Odert^{1,2}, Paul G. Beck^{1,2}, and Astrid Veronig¹

¹ Institute of Physics/IGAM, University of Graz, Universitätsplatz 5, 8010 Graz, Austria
e-mail: flko@gmx.at

² Instituto de Astrofísica de Canarias, 38200 La Laguna, Tenerife, Spain

Received 22 July 2020 / Accepted 9 November 2020

ABSTRACT

Aims. This work aims to detect and classify stellar flares and potential stellar coronal mass ejection (CME) signatures in optical spectra provided by the Sloan Digital Sky Survey (SDSS) data release 14. The sample is constrained to all F, G, K, and M main-sequence type stars, resulting in more than 630 000 stars. This work makes use of the individual spectral exposures provided by the SDSS.

Methods. An automatic flare search was performed by detecting significant amplitude changes in the $H\alpha$ and $H\beta$ spectral lines after a Gaussian profile was fit to the line core. CMEs were searched for by identifying asymmetries in the Balmer lines caused by the Doppler effect of plasma motions in the line of sight.

Results. We identified 281 flares on late-type stars (spectral types K3 – M9). We identified six possible CME candidates showing excess flux in Balmer line wings. Flare energies in $H\alpha$ were calculated and masses of the CME candidates were estimated. The derived $H\alpha$ flare energies range from 3×10^{28} – 2×10^{33} erg. The $H\alpha$ flare energy increases with earlier types, while the fraction of flaring times increases with later types. Mass estimates for the CME candidates are in the range of 6×10^{16} – 6×10^{18} g, and the highest projected velocities are ~ 300 – 700 km s⁻¹.

Conclusions. The low detection rate of CMEs we obtained agrees with previous studies, suggesting that for late-type main-sequence stars the CME occurrence rate that can be detected with optical spectroscopy is low.

Key words. stars: activity – stars: flare – stars: late-type

1. Introduction

The activity of the present-day Sun is investigated in detail, but an ambiguity remains about how active the Sun was in its earlier life. This varying activity strongly affects the evolution of the planetary atmospheres in the Solar System. Observing solar-like stars with different ages would therefore give us an idea of the Sun in time, meaning the evolution of solar-like stars throughout their lifespan (Sonett et al. 1991; Ribas et al. 2005; Güdel 2007).

While solar flares are frequently observed, they would hardly be detectable using optical disk-integrated measurements (Taylor 1997; Kretzschmar 2011; Pevtsov et al. 2014). Highly energetic stellar flares on late-type stars have been observed several times, with later types (M dwarfs) showing more flaring activity than earlier types (Pettersen 1989; Balona 2015; Van Doorselaere et al. 2017). Highly energetic flares on solar-type stars have also been reported, however (bolometric energy of 10^{33} – 10^{36} erg, up to four orders of magnitude higher than the largest solar flares, see Maehara et al. 2012; Basri 2019). While solar coronal mass ejections (CMEs) have also frequently been detected in the past decades with several CMEs per day during high activity (e.g. Webb & Howard 2012), reported CME detections on other stars have so far been rare. On the Sun, the association between highly energetic flares and CMEs is strong and reaches almost 100% for the strongest flares (Yashiro et al. 2006). If the same association rate of solar CMEs to flare were assumed to hold on stars, the

habitability of exoplanets would be severely affected. It would also lead to a stellar mass-loss rate that would exceed so far observed total mass-loss rates on highly active stars, implying that the strong correlation between energetic flares and CMEs might be more uncertain on these objects (Drake et al. 2013; Odert et al. 2017).

Because direct imaging of stellar CMEs is not feasible with current instrumentation, the detection of CMEs is based on other methods. CMEs that have been reported so far were mainly found using signatures such as Doppler shifts of spectral lines and X-ray absorptions (Moschou et al. 2019).

The Doppler-shift method uses enhancements in the blue or red wings in spectral lines (in particular the Balmer lines), implying that this flux comes from ejected material moving toward or away from the observer along the line of sight. This means that the velocities determined by this method are projected velocities giving a lower estimate. Only the erupting filament or prominence, which builds the dense cores of CMEs, is visible as a Doppler feature. The first CME reported with this method was presented by Houdebine et al. (1990), showing a large blue wing enhancement in $H\gamma$ with a projected maximum velocity up to 5800 km s⁻¹. Guenther & Emerson (1997) reported a CME detection with a bulk velocity of ~ 600 km s⁻¹. Several other blueshifted emissions have been reported, possibly arising from a CME (Gunn et al. 1994; Fuhrmeister & Schmitt 2004; Leitzinger et al. 2011; Vida et al. 2016). The reported velocities of these studies range from ~ 100 km s⁻¹ up to several thousand km s⁻¹; the event studied by Houdebine et al. (1990) was the fastest one. The estimated CME masses from these reports range

[★] Table C.1 is only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/646/A34>

from 10^{15} up to 10^{19} g. [Vida et al. \(2019\)](#) searched for asymmetries in Balmer lines in both wings on near M-dwarfs, finding 478 (9 strong) asymmetries in more than 5500 spectra. They reported projected velocities ranging from 100 to 300 km s^{-1} and estimated masses of 10^{15} – 10^{18} g, concluding that their reported maximum velocities are lower than solar CMEs (with typical velocities of about 500 km s^{-1} , [Gopalswamy et al. 2010](#)) and a somewhat lower event rate, suggesting that CMEs might be suppressed by a strong overlying magnetic field. Observing at shorter wavelengths, [Argiroffi et al. \(2019\)](#) reported a stellar flare event including a possible CME using X-ray measurements, with Doppler-shifted plasma movements during the flare followed by a blueshifted upward motion of $\sim 90 \text{ km s}^{-1}$ of cool plasma. Several studies have also searched for CMEs without yielding any detections ([Leitzinger et al. 2014, 2020](#); [Korhonen et al. 2017](#)). CMEs may produce an absorption feature if the erupting filament lies directly on the line of sight to the observer. [Den & Kornienko \(1993\)](#) and [Ding et al. \(2003\)](#) analyzed an erupting filament on the Sun that was observed as a blue wing absorption in $H\alpha$, while also noting that some parts of the material showed emission instead of absorption. No absorption feature reminiscent of an erupting prominence has so far been reported on stars other than the Sun.

Another method to detect CMEs is based on space-borne measurements of X-ray absorptions. The increase in the absorbing column density is interpreted as a rising CME above an active region. Such X-ray absorption events were reported, for example, by [Favata & Schmitt \(1999\)](#), [Haisch et al. \(1983\)](#), and [Moschou et al. \(2017\)](#). In the extreme UV (EUV) range, we can search for EUV dimmings caused by a CME because plasma is evacuated during the beginning phase. Dimmings associated with CMEs are frequently observed on the Sun (see e.g. [Dissauer et al. 2018](#)). [Ambruster et al. \(1986\)](#) reported evidence for a stellar case of a UV dimming (they interpreted it as obscuring CME matter). Type II bursts in the radio domain are another potential method for identifying CMEs. Type II bursts are a signature of electrons accelerated at shock fronts driven ahead of fast CMEs, but no stellar type II bursts have been detected so far ([Leitzinger et al. 2009](#); [Boiko et al. 2012](#); [Crosley & Osten 2018a,b](#); [Villadsen & Hallinan 2019](#); [Mullan & Paudel 2019](#)). We refer to [Moschou et al. \(2019\)](#) and references therein for a more complete overview on CME reports in the literature.

The searches for flares and CMEs have in common that stellar observations over long time periods are needed in order to detect these events. A promising approach are time-resolved observations of stars in data archives, with which the long observation time necessary for detecting such sporadic stellar activity phenomena can be achieved. The problem of having only snapshots instead of long, consecutive observations of single objects might be compensated by having a large number of objects in an archive.

The aim of this work is to search for flares and CMEs on late-type main-sequence stars (F - M) using optical spectra of the Sloan Digital Sky Survey (SDSS). Most of the stars in the SDSS are observed several times in order to obtain a coadded spectrum with better signal-to-noise ratio (S/N) than the single spectra. However, these single observations can be used to find temporal changes in the Balmer lines that indicate possible activity in the chromosphere. Only a few studies have used time-resolved optical SDSS spectra to infer stellar activity ([Kruse et al. 2010](#); [Hilton et al. 2010](#); [Bell et al. 2012](#)). [Hilton et al. \(2010\)](#) searched for flares on M dwarfs in optical SDSS spectroscopic data. With respect to these previous publications, we extend the dataset, use

a different method, a different sample selection, extend the stellar subtypes to F, G, and K stars, perform searches for line asymmetries and CMEs, and use the latest *Gaia* DR2 data to derive stellar distances, based on which we calculate absolute values of $H\alpha$ flare energies and mass estimates for CME candidates.

2. Data sample

2.1. SDSS spectroscopic data

This work makes use of the optical spectroscopic data of the SDSS data release 14 ([Abolfathi et al. 2018](#)). The spectra were restricted to all stars of type F, G, K, and M. Overall, 630,162 spectra were available for the analysis. Stars that are specifically marked with luminosity class I, II, and III as well as white dwarf stars were excluded. Every spectrum is a combination (called the coadded spectrum) of one or more single spectra. These single spectra were usually observed within one hour to up to several weeks. Only objects with at least three single spectrum observations were used to obtain information on the time evolution. It is possible that one object has several coadded spectra, each comprised of a number of single spectra (this usually happens when an object was observed in different survey programs). The exposure times of the single spectra are in the range of roughly 10–25 min. Each of the single spectra itself is split into a red and a blue part. The spectra are provided in units of spectral flux density [$10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$], and the wavelength axis is given in vacuum wavelengths in the heliocentric frame ([Stoughton et al. 2002](#)).

The SDSS data release 14 includes all the data from previous data releases, covering a wide range of different surveys and programs. Two spectrographs were used: the SDSS spectrograph covering data releases 1–8, including the original SDSS spectroscopic survey as well as the Sloan Extension for Galactic Understanding and Exploration (SEGUE) surveys, and BOSS (Baryon Oscillation Spectroscopic Survey) starting with data release 9. The SDSS observes many targets at once using a plate with predrilled holes for each object in the observed sky patch. The original spectrograph was able to observe 640 objects at the same time with a wavelength range of 3800–9200 Å. The BOSS spectrograph is able to observe 1000 objects at once with a slightly wider wavelength range of 3650 – 10400 Å. Both give a spectral resolution of 1500 at 3800 Å and 2500 at 9000 Å. The split between the red and blue part of the optical data is at about 6000 Å ([Ahn et al. 2012](#)). One example of such a spectrum is displayed in Fig. 1.

Because more stars are observed simultaneously, the runtime of surveys is long, and the datasets are cumulative, the overall summed observation time of the stars in our dataset is about 131 years. All pure exposure times of the files were summed and are listed in Table 1.

Flux errors in the spectra are given as estimated 1σ standard deviations per pixel, assuming a normal distribution for each pixel. The SDSS data pipeline slightly adjusted the spectra to match telluric lines and to correct to the heliocentric frame, reporting an accuracy of the wavelength calibration of 10 km s^{-1} or better ([Stoughton et al. 2002](#)).

This work focuses on the analysis of the $H\alpha$ and $H\beta$ lines because the S/N is better than at higher Balmer lines. With a given approximated spectral resolving power $R = 2000$ within the two Balmer line ranges, we obtain a spectral resolution of at least $\sim 3.3 \text{ \AA}$ around $H\alpha$ and $\sim 2.4 \text{ \AA}$ around $H\beta$ (approximately twice oversampled in both ranges).

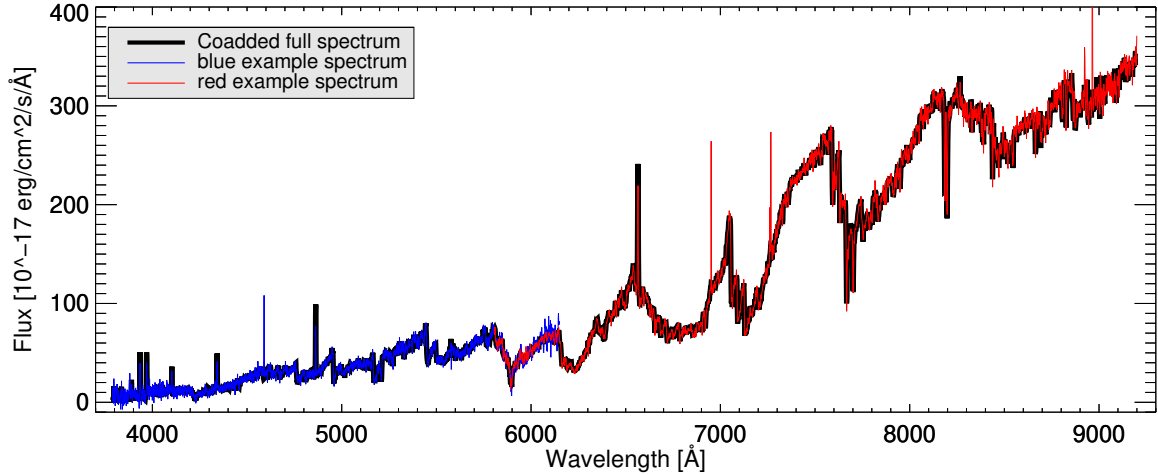


Fig. 1. Full SDSS optical range example plot of a magnetically active M 5e star. The coadded spectrum (black) consists of several spectra that are divided into a red and a blue part, overlapping at ~ 6000 Å. One blue and one red single spectrum is plotted. Cosmic rays in the single spectra are visible as narrow peaks (one in the blue and up to four in the red spectrum range).

Table 1. Summed exposure times of all stars in the dataset.

Stellar spectral type	Summed exposure times per S/N bin [yr]				Sum [yr]
	High S/N	Medium S/N	Low S/N	Lowest S/N	
M	1.39	1.69	2.11	12.56	17.75
K	13.86	4.97	3.67	9.40	31.9
G	4.75	2.26	2.61	6.70	16.32
F	22.55	10.36	12.16	20.10	65.17
Sum	42.55	19.28	20.55	48.76	131.14

Notes. Dataset is divided into stellar spectral types and S/N bins. The cutoff values defining each bin are given in Sect. 2.3.

2.2. Spectral classification

Both spectrographs used specific spectral templates for each stellar subtype, taking the template closest to the spectrum at hand to classify the star. While the original SDSS legacy surveys and the SEGUE had templates for dwarf stars of subtype of M 0–M 9, BOSS only used a template for dwarf stars of subtype M 1 in the same classification region, leading to fewer late-type M stars since SDSS data release 9. It is noted that SDSS spectral classifications often do not explicitly state the luminosity class.

The spectral type of every star that was found by the stellar activity search was confirmed using Skiff (2014) on VizieR, which is a collection of known spectral classifications of different literature sources for each star. One of these sources is West et al. (2011), a catalog of the SDSS M dwarfs up to data release 7. Most of our stellar activity objects were matched with this catalog. West et al. (2011) performed visual inspection of each object to determine the spectral type. The same catalog was also frequently used to obtain distances when an object was not in *Gaia*.

2.3. S/N bins

All spectra were classified into different S/N bins by calculating the S/N within ± 200 Å around the $H\alpha$ line. The reason for the classification based on the surroundings of the $H\alpha$ line is that its continuum flux on late-type stars is higher than for the other Balmer lines. For each object, the mean S/N of all single spectra was calculated. This led to the following classification:

- High- S/N bin: S/N around $H\alpha \geq 20$, resulting in 197 047 stars (31.3%).
- Medium- S/N bin: S/N around $H\alpha$ between <20 and ≥ 15 , resulting in 88 144 stars (14.0%).
- Low- S/N bin: S/N around $H\alpha$ between <15 and ≥ 10 , resulting in 90 119 stars (14.3%).
- Lowest S/N bin: S/N around $H\alpha \leq 10$, resulting in 254 852 stars (40.4%).

Figure 2 shows the number of stars according to the SDSS spectral subtype classification for each S/N bin. As was mentioned in Sect. 2.2, this classification is dependent on the spectral templates used by SDSS. Since BOSS changed the spectral templates, some subtypes are either over- or underrepresented in the data. Figure 2 shows that the later spectral subtypes are mostly prevalent with poor S/N , as expected because their luminosities are low.

2.4. Cosmic rays and flags

Cosmic rays are visible in the data, usually compromising two neighboring data points. Every spectrum found by our algorithms was analyzed by eye to rule out that cosmic rays result in false positives in the stellar activity search. The SDSS provides flags for each data point, which provides information on occurring problems. These flags are resourceful in noting large cosmic rays as well as spectra with general problems, while on the other hand, stellar activity is often also flagged (due to sudden changes from one spectrum to another, as well as large differences to the SDSS template spectrum). We therefore considered flags and

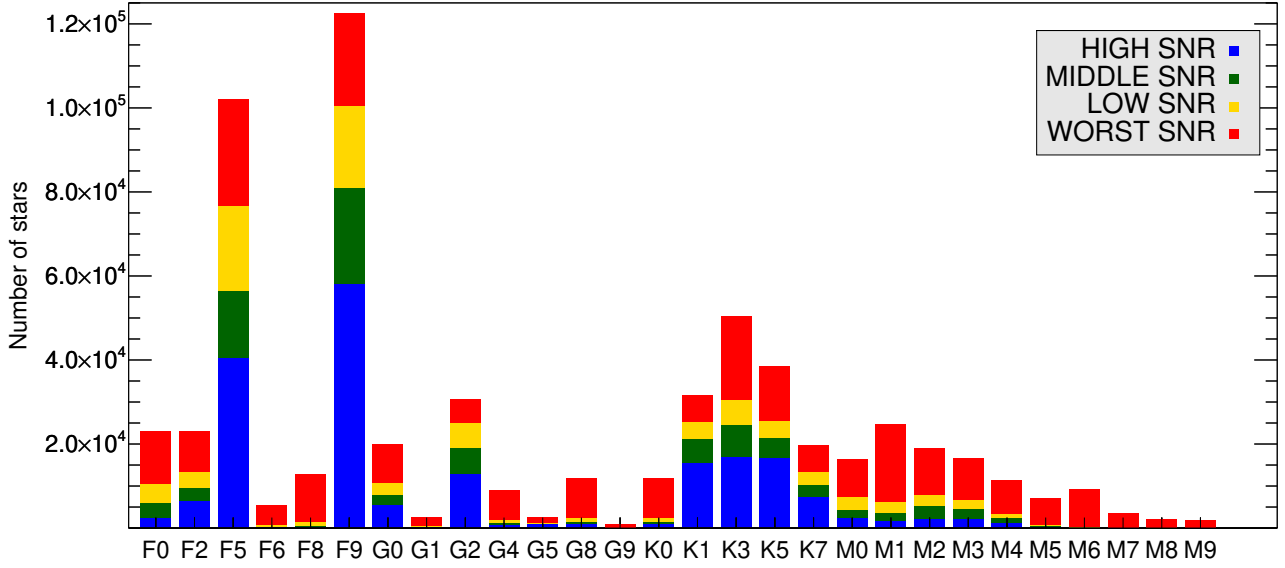


Fig. 2. Number of stars per stellar subtypes in each S/N bin. S/N bins are stacked in this visualization. Spectral classification according to the SDSS.

ruled out cosmic rays after the first search for stellar activity because the SDSS pipeline does not differentiate between cosmic rays and stellar activity. More background information on the pipelines, calibration, and flags of the spectroscopic data can be found in [Stoughton et al. \(2002\)](#).

2.5. *Gaia* DR2 data

We used the second *Gaia* data release¹. *Gaia* DR2 gives information on position and apparent G magnitudes for approximately 1.7 billion sources, parallaxes of 1.3 billion sources, and apparent red photometer (RP) and blue photometer (BP) magnitudes for 1.4 billion sources ([Gaia Collaboration 2018](#)). Only sources brighter than 21 magnitudes are considered by *Gaia*.

To match each of the 630 162 spectroscopic sources from SDSS with their *Gaia* DR2 counterparts, a VizieR search within the *Gaia* DR2 catalog was conducted for the position of every single SDSS source. The search radius was 4 arcseconds. *Gaia* G magnitude information is available for approximately 591 000 stars in our sample (~94%), while BP-RP color information is available for ~90% of the stars in the sample. Both *Gaia* G magnitudes and parallaxes are available for approximately 480 000 stars (~76%).

The distances were calculated using the simple relation of inverting the *Gaia* DR2 parallax (in units of arcseconds) to derive the distance (in units of parsecs). There are more sophisticated methods to derive the correct distance from *Gaia* DR2 data (see [Bailer-Jones et al. 2018](#)), but the resulting difference in distance values between our work and [Bailer-Jones et al. \(2018\)](#) lies within the parallax errors provided by *Gaia* DR2. While the cross-matching of our objects with the *Gaia* DR2 catalog is valid for the majority of the stars, a few special cases remained that were treated separately (e.g., the possibility of two *Gaia* sources within 4 arcseconds, see Appendix B for examples of such cases).

2.6. Data preparation

The methods we used to search for flares and CMEs are based on fitting algorithms of Balmer lines. To correctly perform these

Table 2. Line and continuum regions for Balmer lines.

Line	Continuum 1 (Å)	Line region (Å)	Continuum 2 (Å)
$H\alpha$	6500–6550	6557.61–6571.61	6575–6625
$H\beta$	4810–4850	4855.72–4870.00	4880–4900
$H\gamma$	4270–4320	4331.69–4350.00	4360–4410

Notes. All values adopted from [Hilton et al. \(2010\)](#).

fits, the spectroscopic lines were normalized. Spectral ranges for lines and continua were adopted from [Hilton et al. \(2010\)](#) and are listed in Table 2. For each Balmer line, the mean values of the two associated continuum regions were used to calculate a linear normalization line. The measured flux values were divided by this line, leading to the normalized line profile. Problems using the normalization occurred at spectra with low flux close to zero. Noise occasionally led to nonphysical values below zero, which jeopardized the normalization efforts. Such spectra were not further considered in the analysis.

3. Methods

3.1. Flare detection methods

To search for stellar flares in the spectra, we used methods based on identifying variations in the Balmer lines. The following subsections describe our approaches in more detail.

3.1.1. Amplitude variability search: high Threshold Method (HTM) and Low Threshold Method (LTM)

Using the normalized coadded spectra and the normalized single spectra, the $H\alpha$ and the $H\beta$ Balmer lines were fit by a simple Gaussian curve,

$$I = 1 + I_0 \times \exp\left(-\frac{(\lambda - \lambda_0)^2}{\sigma}\right). \quad (1)$$

Here, I_0 denotes the line amplitude, λ_0 the wavelength of the line center, and σ the width of the Gaussian profile. These three

¹ <http://cdsweb.u-strasbg.fr/gaia>

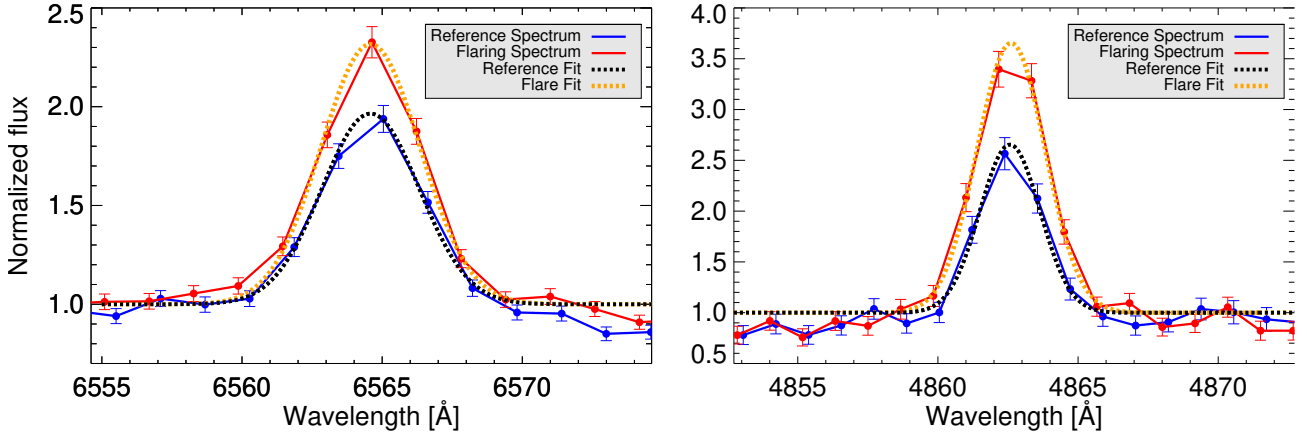


Fig. 3. Example of a flaring $H\alpha$ line (left panel) and flaring $H\beta$ line (right panel) detected on a M 5e star. Two single spectra are shown, one as a reference and one as an example of a flare. The data points with error bars together with the Gaussian fits are shown.

parameters were fit, and reasonable initial start values ($I_0 = 0.0$, $\lambda_0 =$ Balmer line core value, $\sigma = 4 \text{ \AA}$) were chosen. The initial amplitude value was set to zero to account for absorption lines (negative I_0) and emission lines (positive I_0). All values around $\pm 10 \text{ \AA}$ of the core wavelength of each Balmer line were used in the fit.

Activity in $H\alpha$ was defined as a significant change of the fitted Gaussian line amplitude I_0 from one time instant to another. The amplitude change was required to exceed the error on the amplitude I_0 as derived from the fits. This amplitude error was determined by Gaussian error propagation and includes the amplitude errors of both Gaussians that are compared to each other, as each fit gives an error value. Two thresholds were used: amplitude changes greater than twice (high-threshold method, henceforth abbreviated as HTM) or 1.5 times (low-threshold method, henceforth abbreviated as LTM) of the amplitude error. To rule out nonphysical line fits, the resulting shift of the line core in the x -axis was required to be less than 2 \AA from one time instant to another.

Because of the low S/N, spectral resolution, and the possibility of outliers in the data, a significant change of the $H\alpha$ line alone is not sufficient to declare the spectrum as flaring. The same algorithm was therefore used on $H\beta$ as well. Because the noise around $H\beta$ is higher than around $H\alpha$, this second condition further decreases the number of eligible events compared to using $H\alpha$ alone. A star was defined to be in flaring state when the fitted $H\alpha$ and $H\beta$ lines both show amplitude variations above the thresholds.

The reasoning behind this method is that the cores of Balmer lines follow Gaussian curves arising from Doppler broadening. The fit is robust against outliers (e.g., cosmic rays) and therefore represents the overall line shape more faithfully, while trends in the line wings and outside of the Balmer lines are ignored. Especially late-type M-stars show strong contamination of molecular bands near $H\alpha$, for example. In addition to that, it is unlikely that both $H\alpha$ and $H\beta$ simultaneously change significantly above the amplitude error without flaring activity.

On the downside, the line-core fitting algorithm overestimates its fit significance for spectra with very low S/N by ignoring the noise around the Balmer lines. This made the visual post-inspection of each candidate event indispensable (see Sect. 3.1.3). Large outliers and faulty or missing data points could also lead to erroneous fits. Because of the spectral resolution of the SDSS spectra, the Balmer line cores (especially for

M stars) usually consist of fewer than ten data points. Cosmic rays usually affect between 2 and 4 pixels of the camera on the spectrograph. Depending on the position and size of the cosmic ray hit, false data points may have similarities to emission lines. An example of a successful flare detection with this method is given in Fig. 3.

We applied this search strategy, and the HTM resulted in 538 flare candidates on M stars (183, 90, 73, and 192 flares for high, medium, low, and lowest S/N, respectively) and 47 candidates on K stars (19, 17, 6, and 5 flares for high, medium, low, and lowest S/N, respectively). The LTM detected an additional 420 M star flare candidates (9, 10, 35, and 366 flares for high, medium, low, and lowest S/N, respectively) and 132 on K stars (18, 22, 15, and 77 flares for high, medium, low, and lowest S/N, respectively). After manual inspection, all of the 54 detections on G stars and 169 detections on F stars were considered to be either nonflaring or had an incorrect spectral classification. The main reasons for this misidentification are either high noise, observational or data reduction errors that affect the whole spectrum, cosmic ray hits in both Balmer lines, the star being a close binary affecting the Balmer lines, or the star being a different type (e.g., RR Lyrae).

3.1.2. $H\alpha$ emission line method (ELM)

High noise and variations in the wings could lead to missing flares in the line-core fitting search algorithm. Because stars showing Balmer lines in emission are magnetically active and because the probability of detecting flares is therefore much higher, a search for stars showing the $H\alpha$ line in emission was started. We again used the Gaussian fitting algorithm (Eq. (1)), but this time, all stars were identified with an emission line amplitude I_0 higher than 1.5 at least. The resulting list of stars were then manually searched for changes exceeding the typical errors from one time instant to another in both $H\alpha$ and $H\beta$.

The $H\alpha$ emission line search algorithm for M stars found 177 in the high, 146 in the medium, 203 in the low, and 5049 stars in the lowest S/N bin. These numbers include several flaring stars that were previously detected with the HTM and LTM algorithms. For K stars, the search yielded 167 candidates (147 from the lowest S/N bin alone), but almost all of these detections are due to erroneous data values. There were no reasonable findings of strong Balmer emission in G and F stars except for detections arising from erroneous data. Detections in the lowest S/N bin were not further considered because the high

number of false emission line detections in this list rendered a manual inspection unfeasible. After manual inspection, this method yielded six additional flares that were not found with the other methods.

3.1.3. Visual inspection of the flare candidate spectra

After we obtained the numerous detections by the two algorithms, each identified candidate was verified manually to determine whether the flare was valid considering the following aspects: first we sorted out spectra with obvious data problems (e.g. incorrect flux levels due to data reduction errors or severe cosmic-ray hits). We also checked that the change in the line amplitude is significantly above the surrounding noise in both $H\alpha$ and $H\beta$. Considering the relative flare line amplitude and the error bars of the individual data points, we did not consider any amplitude changes that were too small (similar height as the errors or smaller). In addition to that, we checked for visible activity in other Balmer lines (especially $H\gamma$). Stars with no visible activity (enhanced data points in the same observation) in other lines despite having sufficiently good S/N were sorted out. Data point flags for each $H\alpha$, $H\beta$, and $H\gamma$ line were controlled. While flares are usually automatically flagged by the SDSS, several other aspects also result in flagged data points. Individual flags were controlled and spectra with many flags were noted. Finally, we cross-checked with flares identified in Hilton et al. (2010) to determine whether a flare was mistakenly rejected based on one of these aspects. These events were then reevaluated (henceforth denoted Reeval.), which happened in four cases.

Based on these aspects, each spectrum was categorized into one of three groups “Yes”, “Probable” or “No” to indicate whether the spectrum should be considered or discarded from the flare list. The stars in the category “Yes” were considered to show distinct flares. The spectra in the “Probable” category show clear activity in (at least) $H\alpha$ and $H\beta$, but more caution in handling them is appropriate. They either show no visible activity in $H\gamma$ (due to high noise) or several or all of the data points in one emission line were flagged, indicating a problem with this part of the spectrum. In these cases, all Balmer lines in the blue spectrum as well as the Ca II H and K lines were checked. Only events that clearly resembled usual flaring behavior were kept (e.g., simultaneous flux increase in all Balmer lines instead of flux changes at different observation times; Ca II K line peaking later than the Balmer lines; see also Kowalski et al. 2013). Four “Probable” flares were also reported in Hilton et al. (2010). Spectra with unusual properties (e.g. shifted Balmer lines) were checked regarding their spectral type classification. Eight stars (mis-) classified as M dwarfs by SDSS were identified as either cataclysmic variables (CV) or close/eclipsing white dwarf - M star binaries and were removed from the flare event list. CVs can be identified by the shift in their emission lines within a short period of time and the strong periodic changes in the strength of the lines.

Based on all these aspects, the final list of 281 flare candidates was created. This is the basis for further in-depth analysis.

3.2. CME detection methods

To search for signatures of stellar CMEs, we used methods based on the idea of identifying time variations in the Balmer line wings. This indicates motions of CME material with a detectable line-of-sight velocity component. The following subsections describe our approaches in more detail.

3.2.1. CME-train search

To detect enhanced emission or absorption in the blue or red wing of the emission or absorption line under study, two wavelength windows were defined covering the far red and blue wings of $H\alpha$ and $H\beta$ (corresponding to Doppler-velocity ranges of ~ 350 – 1500 km s⁻¹). A search was conducted to determine whether several successive data points (a “train”) changed significantly from one time instant to another. All trains that exceeded a predefined length were reported. Train sizes of 5 and 3 Å were tested (equivalent to velocity ranges of ~ 230 and ~ 140 km s⁻¹ in $H\alpha$ and ~ 310 and ~ 180 km s⁻¹ in $H\beta$). The finding of one train in a spectrum triggered a detection that was subsequently cross-checked manually.

The idea behind this method was to find wide CME-related emissions like those found by Houdebine et al. (1990). The algorithm was successfully tested on an CME-related blue wing feature in the data from V374 Peg presented in Vida et al. (2016). We note that this algorithm has two weaknesses: slow or small changes over several consecutive spectra are not recognized, and one single data point in the whole enhanced train with larger error or with an erroneous value may prevent a detection.

This algorithm yielded 375 detections, which only consisted of either unusually large and wide cosmic-ray hits or spectra showing incorrect flux levels at one observation time due to faulty data reduction. A short discussion of the reasons for nondetections of CMEs is given in Sect. 5.8

3.2.2. Manual search for flare events and emission lines

All spectra found by the flare detection algorithms were inspected manually to search for Balmer line asymmetries or line wing enhancements or absorptions as an indicator of a CME. Every spectrum showing an asymmetry in either $H\alpha$ or $H\beta$ was noted (even when it was one data point). Spectra with clear asymmetries in both lines were recognized as possible CME candidates and analyzed in more detail. In these cases, the $H\gamma$ line and additional Balmer lines were included in the analysis, while all data point flags and error bars were controlled. Figure 4 shows an example of a low S/N star where an asymmetry was suspected. We found six possible CME candidates with this method (Sect. 4.3.3).

4. Analysis and results

4.1. Flare-search results

Using these methods, we detected no flares for dwarf stars of types G or F, 272 events were found for M dwarfs, 6 on K dwarfs, and 3 on M-type stars that are known T Tauri stars. These 3 stars are labeled primarily by their luminosity class (T Tauri) from now on to differentiate them from the M dwarfs. This gives us a final list of 281 flare candidates.

Table 3 shows all flares per spectral subtype listed for each S/N bin. One of the K stars is most likely a giant or subgiant star; this detail as well as the ambiguity of the other K-type stars is discussed in Appendix B. Table 4 lists the final number of flares that were found by each method, showing that HTM proved to be the most successful flare-search method in this work.

4.2. Flare parameters

Efforts were made to constrain flare parameters, in particular the flare occurrence, flare energies, flare luminosities, and flare

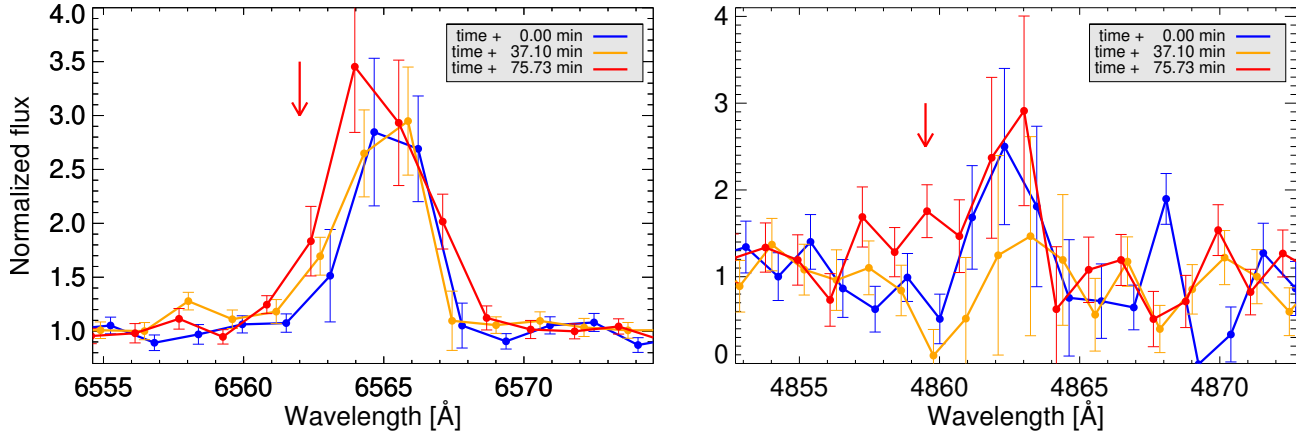


Fig. 4. Example of possible blue-wing line asymmetries (indicated by arrows) and a possible flare in $H\alpha$ (left panel) and $H\beta$ (right panel) on an M 4 star. This object has a low S/N and shows too little activity to be categorized as a CME or flare candidate.

Table 3. Spectral type of flaring stars per S/N bin.

S/N bin	K3	K5	K7	M0	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	TT	Total
High	1	2	3	1	0	4	6	35	29	19	5	0	0	2	107
Medium	0	0	0	0	1	0	3	3	15	13	3	1	0	1	40
Low	0	0	0	0	0	1	2	4	2	16	14	1	1	0	41
Lowest	0	0	0	0	0	0	2	1	5	30	29	9	17	0	93
Total	1	2	3	1	1	5	13	43	51	78	51	11	18	3	281

Table 4. Number of detected flaring stars per method.

S/N bin	HTM	LTM	ELM	Reeval.	Other	Total
High	99	5	2	0	1	107
Medium	38	1	0	1	0	40
Low	32	4	3	2	0	41
Lowest	67	24	1	1	0	93
Total	236	34	6	4	1	281

Notes. HTM and LTM: High- and low-threshold method, ELM: emission line method, and Reeval.: reevaluated spectra. The one flare in the “Other” category was found on a K star by inspecting $H\alpha$ activity alone.

evolution for each identified flare. We briefly describe the preceding analysis and object parameters necessary to yield these results. Regarding the stellar distances, 259 (92.2%) of the flaring stars have associated *Gaia* DR2 parallaxes (see Sect. 2.5). Stars with missing parallaxes were checked for other sources within VizieR, and if possible, a value for the distance or the parallax was added. The spectral classification cross check (see Sect. 2.2) served as an indicator to determine whether the SDSS spectral classification was valid. We found 253 of the 281 flaring stars within the *Skiff* (2014) catalog. Missing stars were again checked for other sources. If not stated otherwise, the spectral types from the literature are used in the following subsections.

It is necessary to define a reference spectrum for each star for comparison with the flaring spectra and to gather the flare parameters. A quiet flag” was introduced to mark reference spectra that were in a nonflaring state. Spectra were considered to be quiet when at least two of them at a completely different observation time than the flaring spectra show negligible differences. Spectra

were considered to be “probably quiet” when they appeared before or after a consecutive flaring series without any further flux changes. Of the 281 events, 59 show quiet spectra and 64 show probably quiet spectra.

4.2.1. Flaring occurrence, positions in the color–magnitude diagram, and sample purity

To estimate the flare occurrence rates for each stellar subtype, the flaring times of all events were summed for each subtype and divided by the overall exposure times of all available stars of the corresponding subtype. The flaring percentage of the later M-type stars is highly dependent on whether a possible flare was detected despite the poor S/N. Not enough flares were found on K-type stars for meaningful statistics.

Figure 5 shows the results of the flaring occurrence separately for the spectral subtypes and S/N classes. The high and medium S/N bins are more reliable than the other bins. No flare on an M 8 or M 9 star was detected in the high S/N class, which is again a consequence of their low flux. Overall, the flaring occurrence appears to be higher for later type stars.

It is important to note that the overall exposure times per subtype were based on the original SDSS classification, while the flaring stars were classified using several sources. The large flaring fraction reaches high values (up to 55% for type M 7) because very few late-type stars lie in these S/N bins (only about ten stars of type M 7 stars are in “High”). Here, the mix of different spectral classifications may compromise the absolute values, but the overall trend is valid.

Using the data given by *Gaia* DR2, we can plot the position of the flaring stars in the color–magnitude diagram (CMD). The basic idea behind this is to test whether the stars in question are in fact late-type dwarf stars, resulting in a high sample purity regarding the luminosity class.

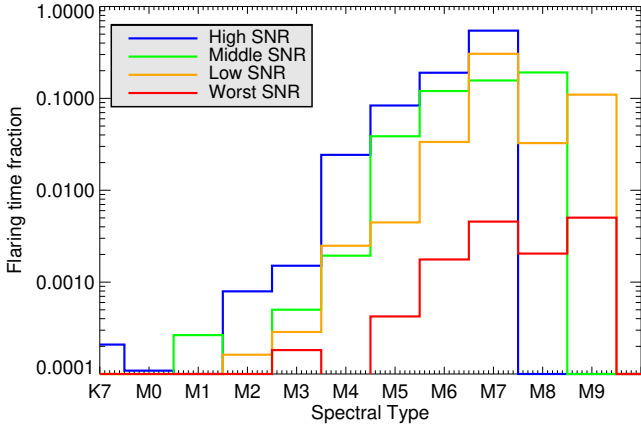


Fig. 5. Flaring occurrence per stellar subtype in percent (fraction of observed flaring time compared to overall exposure time in the dataset for each spectral subtype). The S/N bins are color-coded.

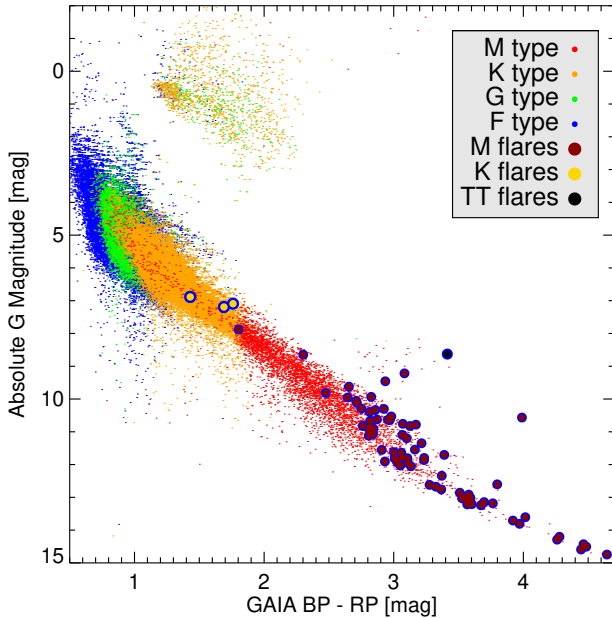


Fig. 6. CMD of the SDSS sample with errors <1 mag in both axes for the original sample. The flaring subsample has errors of <2 mag in absolute magnitude and <1 mag in the color axis. The main constraining factor is the availability of error values for the color.

Figure 6 shows the flaring stars in the CMD based on their absolute G magnitude derived from *Gaia* DR2 and their *Gaia* BP - RP color. The respective errors for the BP - RP color as well as the absolute magnitude were restricted to be smaller than 1 mag. This restriction reduces the whole SDSS sample to approximately 121 000 stars ($\sim 19\%$ of all available stars) and the number of flaring stars to 76 (30% of all flaring stars) in the plot. This reduced number of stars in this plot is mostly due to large errors or missing values in the BP - RP value from *Gaia*. We find that most of the displayed stars lie well within the main sequence. Some stars are above the M star main-sequence region, which is an indicator for pre-main-sequence stars (T Tauri stars), one of which was confirmed to be one. None of the flaring stars presented in this plot shows signs of being an evolved star on the RGB.

Figure 7 shows all M and K stars in the dataset with their *Gaia* G magnitudes over their distances. All flaring stars

are displayed as well, and all flaring M stars lie well within 1000 parsec. The G magnitude of a solar-like star was plotted as a function of the distance to give a distinction line between stars with subsolar luminosity (such as M and K dwarfs) and stars with higher than solar luminosity (like red giants). One instance of a flaring K star above this line is a clear indication that this star is either a subgiant or a giant star (other possibilities include T Tauri stellar type). The function of a well-known M 5.5 dwarf (Proxima Centauri) is also displayed. The function of an M0 type star (GJ 270) is displayed to approximately show the threshold between K and M stars. The error of the distance is color coded as described in the figure caption. The diameter of the Milky Way is given as an upper limit for unrealistic distance values. The plot shows that we obtained the object information out of two different datasets (SDSS and *Gaia* DR2). The lower threshold in magnitude is most obvious because *Gaia* DR2 restricts its measurements of stars brighter than 21 mag in G (*Gaia* Collaboration 2018). Figures 6 and 7 show the observed stellar magnitudes without a correction for extinction, which would affect the true magnitude values of the most distant stars. Figure 7 further indicates that the displayed sample of flaring M stars only consists of M dwarfs. It also works as a tool for detecting giant stars in the flaring K star subsample. For a short discussion of the sample purity, see Sect. 5.4.

4.2.2. Flare energies, flare luminosities, and detection threshold

We calculated the flare energy E_{flare} and peak luminosity L_{flare} in $H\alpha$. Other chromospheric lines were not considered because their S/N is lower than in $H\alpha$.

As the data values in the spectra are given as spectral flux density I_{λ} in units of $[10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}]$, we calculated the fluxes F ($[\text{erg s}^{-1} \text{ cm}^{-2}]$) for each spectrum (abbreviated by “spec”) by

$$F_{\text{spec}} = \int_{\lambda_1}^{\lambda_2} I_{\lambda}(\text{\AA}) d\text{\AA}, \quad (2)$$

where λ_1 and λ_2 determine the wavelength range for the integration. The window around $H\alpha$ is 6555–6575 \AA . Prior to the integration, linear interpolation was used. This was necessary as each spectrum usually has a different wavelength axis, which often results in a different number of data points in the integration range. The excess flaring flux was obtained by subtracting the reference, $F = F_{\text{flare}} - F_{\text{ref}}$. The reference (ref.) spectrum was either a spectrum declared as quiet, the coadded spectrum, or the spectrum showing the least amount of activity. Spectra showing cosmic rays in the integration range as well as flares with high noise in the reference were excluded. The total flare flux (sum of all flaring fluxes of an event) as well as the peak flaring flux were evaluated. The total radiant energy per area H ($[\text{erg cm}^{-2}]$) was obtained by multiplying the flaring excess flux of each spectrum with the exposure time and summing,

$$H = \sum_i^N F_i * t_{\text{exp}, i}. \quad (3)$$

Here, i denotes each single spectrum, and N is the total number of consecutive flaring spectra. The total flare energy (in units of [erg]) is calculated by

$$E_{\text{flare}} = H * 4\pi d^2, \quad (4)$$

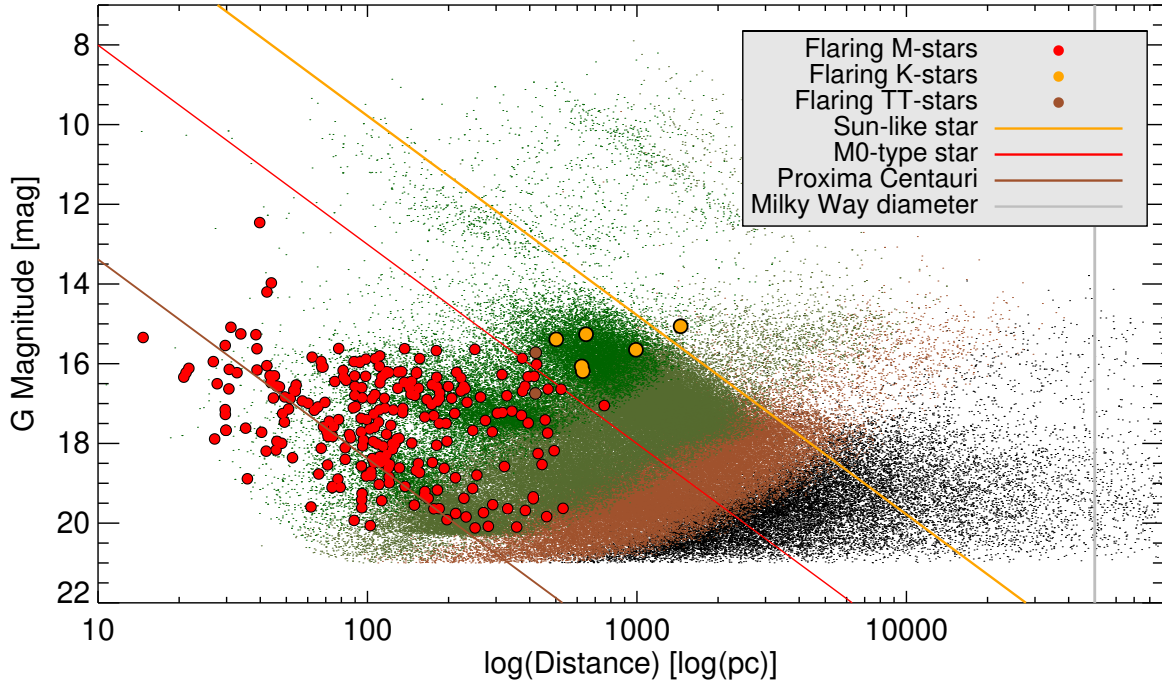


Fig. 7. M- and K-type stars (using the SDSS spectral type classification) with their respective distances and G magnitudes from *Gaia* DR2. Flaring stars are shown within the plot. Example for a dwarf star (Proxima Centauri), an M0-type star (GJ 270), and a Sun-like star are given. The diameter of the Milky Way is given as an upper constraint on the distance. Errors in the distance are color-coded. Green shows error smaller than 5%. Olive shows an error between 5 and 20%. Brown shows an error between 20 and 50%. Black shows an error higher than 50%.

and the instantaneous flare luminosity (in units of $[\text{erg s}^{-1}]$) is given as

$$L_{\text{flare}} = F * 4\pi d^2, \quad (5)$$

with d being the distance to the star. E_{flare} (as well as H) gives a total value for the flare energy, while L_{flare} gives a luminosity value for a single flaring spectrum.

The integration errors for the energy and luminosity were determined as follows:

$$\Delta F_{\text{spec}} = \Delta \lambda * \sqrt{\sum_j^M (\text{error}_{(j)})^2}, \quad (6)$$

with $\Delta \lambda$ being the mean wavelength spacing of the data points, M the number of all data points in the range (without interpolation), and $\text{error}_{(j)}$ the error bar associated with data point j . Using Gaussian error propagation,

$$\Delta F = \sqrt{\Delta F_{\text{ref}}^2 + \Delta F_{\text{flare}}^2} \quad (7)$$

and

$$\Delta H = \sqrt{\sum_i^N (\Delta F_i * t_{\text{exp},i})^2}, \quad (8)$$

the final result of the error in flare energy was determined by

$$\Delta E_{\text{flare}} = |\Delta H 4\pi d^2| + |\Delta d H * 8\pi d|, \quad (9)$$

and for the luminosity by

$$\Delta L_{\text{flare}} = |\Delta F 4\pi d^2| + |\Delta d F 8\pi d|, \quad (10)$$

with Δd being the associated error determined by the *Gaia* DR2 parallax. If no distance error value was available, the second term in Eqs. (9) and (10) was neglected.

Figure 8 shows the result of the $H\alpha$ flare energy over the spectral subtype. The flare energy arising from excess flux in $H\alpha$ ranges over five orders of magnitude from $\sim 3 \times 10^{28}$ erg up to 2×10^{33} erg. The plot shows a clear trend in the flare energy, with late-type stars showing less energetic flares than earlier stars. S/N bins are color-coded for each flare. The highest calculated flare energies are observed on K stars and earlier M stars. This is probably a selection effect: because K stars are brighter, they need higher energy releases for flares to be detected. Another explanation for this trend is the simple correlation with distance: there are more late-M stars in the solar neighborhood, making their flares (even at small energies) easier to detect, while there are no late-M dwarf stars in the dataset with large distances to the Solar System. Their flux is not high enough for spectroscopic observation in the SDSS. K stars, on the other hand, show higher absolute magnitudes and are easier to observe at larger distances.

Figure 9 shows the result of the $H\alpha$ flare energy over the distance. Stars that have quiet reference spectra are marked in blue. The energy values of these flares are more reliable because the reference is not part of the flare, which would compromise the calculation. It was not clear whether the reference spectrum was quiet or still in a flaring state for all flares marked in red. The two groups clearly agree, indicating that the obtained values marked in red are also fairly reliable.

We again found the same range for the flare energy arising from excess flux in $H\alpha$ with the typical spread of about one order of magnitude in energy. The more distant a star, the more energetic the flare has to be to be detected with our methods. This is also another validation of the significance of the flares that we found. This plot mostly depends on the reliability of the *Gaia* parallax information. Calculated error bars (using Eq. (9) for the

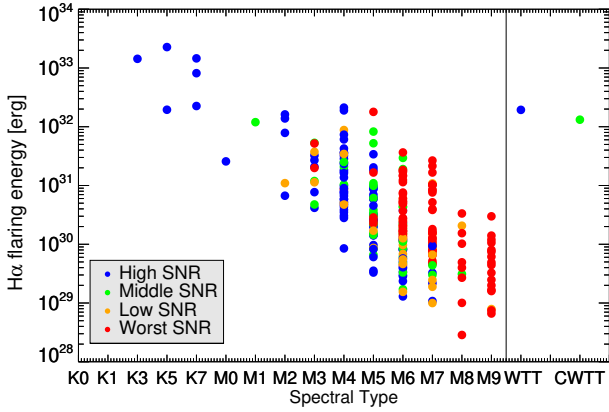


Fig. 8. Flare energy as a function of stellar subtype. T Tauri stars are displayed separately. The S/N bin for each star is given.

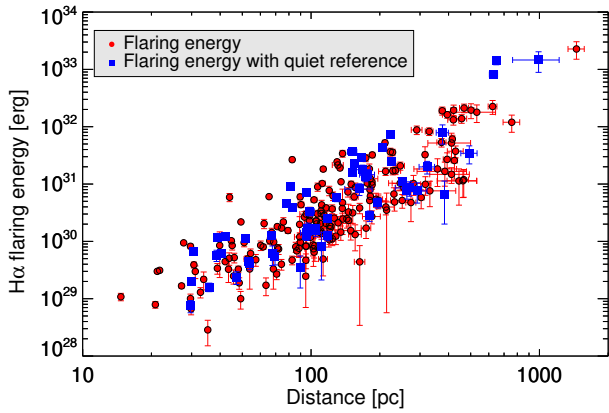


Fig. 9. $H\alpha$ flare energy as function of distance from the *Gaia* DR2 catalog. Stars with quiet reference spectra are marked in blue. Stars where the reference spectra is not unambiguously in a quiet state are marked in red.

energy and *Gaia* parallax errors for the distance) are displayed. The apparent upper limit of energetic flares (showing that there appear to be no energetic flares at a close distance) of nearby stars may be a result of a selection effect. This is discussed in Sect. 5.5.

To answer the question which minimum flare threshold is still detectable in this dataset, we considered the relation between the maximum calculated $H\alpha$ flare luminosity (in units of [erg/s]) and the distance of each flare star. A simple quadratic equation was used,

$$L_{\text{fit}}(d) = 10^c \times d^2. \quad (11)$$

Here, L_{fit} denotes the fitted function of flare luminosity as a function of the distance d in parsec, and c is a free parameter determining the detection threshold. The coefficient 10^c is given in units of [erg s⁻¹ pc⁻²]. For the mean flare luminosity as a function of distance, a value of $c = 23.1344$ was found. The lower threshold was estimated with a value of $c \sim 22.3$. This means that a star at 100 pc distance needs a measured $H\alpha$ flare luminosity of at least $10^{22.3} \times 100^2 = 10^{26.3}$ erg s⁻¹ to be detected in this dataset (see Fig. 10).

4.2.3. Flare evolution

The $H\alpha$, $H\beta$, and $H\gamma$ Balmer lines were manually inspected. To estimate the evolution of the flare, every event was categorized as

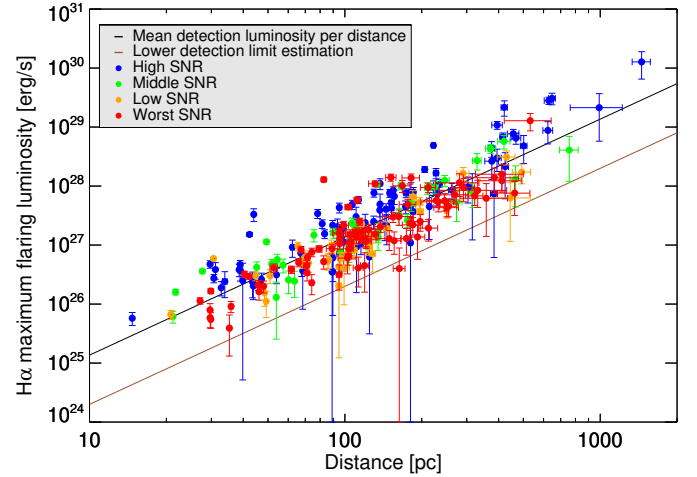


Fig. 10. Flare peak luminosity as a function of stellar distance. S/N bin of each star and the respective error bars are color-coded. The mean detection luminosity at a given distance as well as an estimate of the lower detection limit is displayed by the black and brown lines.

Table 5. Flare evolution estimates of all 281 flares.

	Rise	Decay	Full	Activity at diff. time	Other	Total
#	82	67	55	67	10	281

follows: “Rise” (Balmer lines grow over consecutive time series), “Decay” (Balmer lines decay over consecutive time series), “Full” (entire flare evolution including peak), “Activity at different time” (a single flaring spectrum at a separated observation date, preventing an evolution estimate), and “Other” (unable to classify the evolution of the flare). Irregular cases, that is, classes in which the flare decreases and rises again immediately afterward, fall into the “Other” category. At each instant, the whole available spectrum (especially other Balmer lines) was checked to gain further information.

The results of the manual flare evolution identification is summarized in Table 5. These results were cross-checked with the overlapping flares from Hilton et al. (2010), giving the same results. More flares are declared to be in a rising phase than in decay. This finding is discussed in more detail in Sect. 5.6.

4.3. CME search and results

4.3.1. Overview of CME candidates

We present six possible CME candidates, all detected on M stars, that show enhancements in Balmer line wings (see Sect. 4.3.3). They were found by noting wing asymmetries in all stars that were manually inspected (mostly flare candidate stars). Five candidates show a red wing, and one shows a blue-wing enhancement. We discuss the interpretation of red-wing enhancements in Sect. 5.7. An overview of these candidates is given in Table 6, including general information on the objects.

4.3.2. Mass and velocity estimations

Under the assumption that the line wing enhancements we found are due to CME material having a velocity component along the line of sight, we can estimate the velocity of the outflowing material. The maximum speed was estimated as the point at which

Table 6. CME candidates.

No	SDSS ID	RA	Dec	Type	Flare	S/N bin	Lines with wing enhancement
1	J052700.12+010136.8	81.75054	1.02691	M 3.2e* ^[1] known WTT ^[2]	No	High	Hα , Hβ , Hγ
2	J085556.66+174942.0	133.98612	17.82835	M 4e ^[3]	Yes	High	Hα , Hβ , Hγ
3	J110743.30+075624.9	166.93047	7.94027	M 4e ^[3]	No	High	Hα , Hβ , Hγ , Hδ , Hϵ , H8
4	J042139.64+264913.8	65.41520	26.82050	M 6e* ^[4]	Yes	Medium	Hα , Hβ , Hγ , Hδ
5	J053709.91-011050.3	84.29134	-1.18067	M 5e* ^[5] known WTT ^[6]	Yes	high	Hα , Hβ , Hγ , Hδ , Hϵ , H8 , Ca I H&K
6	J052839.66-000322.5	82.16528	-0.05627	M 4e ^[3]	Yes	high	Hα , Hβ , Hγ , Hδ

Notes. Listed properties including SDSS ID, right ascension and declination, spectral subtype, if an associated flare is visible, S/N bin, and the lines showing the enhancement. Significant enhancements in line wings (above all error bars) are marked in bold font. *The “e” for emission line star was added to these stars based on the visual inspection.

References. Spectral type sources: ^[1] Suárez et al. (2017), ^[2] McGehee (2006), ^[3] West et al. (2011), ^[4] Luhman et al. (2017), ^[5] Briceño et al. (2019), ^[6] Koenig et al. (2015).

Table 7. CME candidates with associated values.

No	Enhanced Wing	Distance [pc]	Mean H α S/N	Max. velocity [km s ⁻¹]	Bulk velocity [km s ⁻¹]	Est.mass H α [g]	Error %	Est.mass H γ [g]	Error %
1	Red	379.3	28	594	494	3.0E+17	81	X	X
2	Red	234.2	22	708	275	6.4E+17	22	X	X
3	Red	70.5	28	667	300	6.1E+16	51	1.3E+17	24
4	Red	166.3	17	703	358	8.5E+17	11	X	X
5	red	–	21	708	340	–	–	–	–
6	Blue	467.8	24	-365	-240	6.1E+18	19	X	X

Notes. Listed properties include indication of the direction of the Doppler shift, the distance in parsec based on *Gaia* DR2, their maximum and bulk velocity, and the estimated mass (including rough error estimates in percent) using the flux power in H α and H γ . A cross marks cases where the mass estimation was not possible for specific reasons (see Sect. 4.3.3).

the enhancement merges with the continuum, while the bulk velocity was determined by fitting a Gaussian to the enhanced feature, determining its central wavelength. If a Gaussian fit to the enhanced line wing feature was not possible, the bulk velocity was estimated using the mean wavelength value of the enhancement.

To estimate the hydrogen mass, the following relation from Houdebine et al. (1990) was used:

$$M_{\text{CME}} \geq \frac{4\pi d^2 F_{\text{em}} (N_{\text{tot}}/N_j) m_{\text{H}} \eta_{\text{OD}}}{h\nu_{j-i} A_{j-i}}. \quad (12)$$

Here, d denotes the distance of the star, F_{em} the integrated excess flux from the emission feature, N_{tot}/N_j the ratio between the total number of hydrogen atoms to the hydrogen atoms being at excitation level j , m_{H} the mass of the hydrogen atom, η_{OD} the opacity damping factor (giving a value to indicate how much flux escapes the plasma due to optical thickness), h the Planck constant, ν_{j-i} the frequency difference between excitation level j and i , and A_{j-i} the Einstein coefficient for spontaneous decay from excitation level j to i .

While the CME candidates are most prevalent in the H α line (transition from $j = 3$ to $i = 2$), no value is available in the literature to estimate the ratio N_{tot}/N_3 . However, the transition for H γ ($j = 5$, $i = 2$) is estimated to be $N_{\text{tot}}/N_5 = 2 \times 10^9$ by Houdebine et al. (1990) using nonlocal thermal equilibrium (NLTE) modeling for the active M star AD Leo. η_{OD} was set to 2, indicating that 50% of the radiation escapes the plasma (Leitzinger et al. 2014), while the Einstein coefficient is $A_{5-2} = 2.53 \times 10^6$ (Wiese & Fuhr 2009).

As the SDSS spectra cover the whole visible spectrum, the H γ line can be used, but poses difficulties because the S/N is low, especially for faint late-type stars. We can therefore approximate the real value of the CME flux in H γ by scaling the flux from H α with the Balmer decrement BD ($F_{\text{em},\gamma} = F_{\text{em},\alpha}/\text{BD}$), with BD set to 3 (adopted from solar and stellar flares, Butler et al. 1988). When possible, we used both H α and H γ measurements for CME mass estimations for each candidate.

For error estimations, the same method as in the flare flux error analysis was used (see Eq. (6)). For all CME excess flux measurements, the flux as well as the errors of the coadded spectrum was used as reference. The coadded spectrum has smaller uncertainties because the S/N is higher.

Table 7 shows calculated values such as maximum velocity, bulk velocity, and mass estimations based on the CME assumptions. The derived maximum velocities range from 365 to 700 km s⁻¹ and the bulk velocities from 240 to 500 km s⁻¹ (absolute values). The derived CME masses range from 6×10^{16} to 6×10^{18} g.

4.3.3. Detailed analysis of the six CME candidates showing enhanced Balmer line wings

Candidate 1: SDSS J052700.12+010136.8 (Figs. 12 and 11). The star with subtype M 3.2 (SDSS classification M 4) was found with our ELM algorithm. McGehee (2006) also reported it as a weak-line T-Tauri star (WTTs). The SDSS spectrum shows five consecutive spectra with exposure times of 15 min for spectra 1 and 5, and 25 min for spectra 2, 3, and 4. Figure 11 shows the red wing of the H α , H β , and H γ lines of

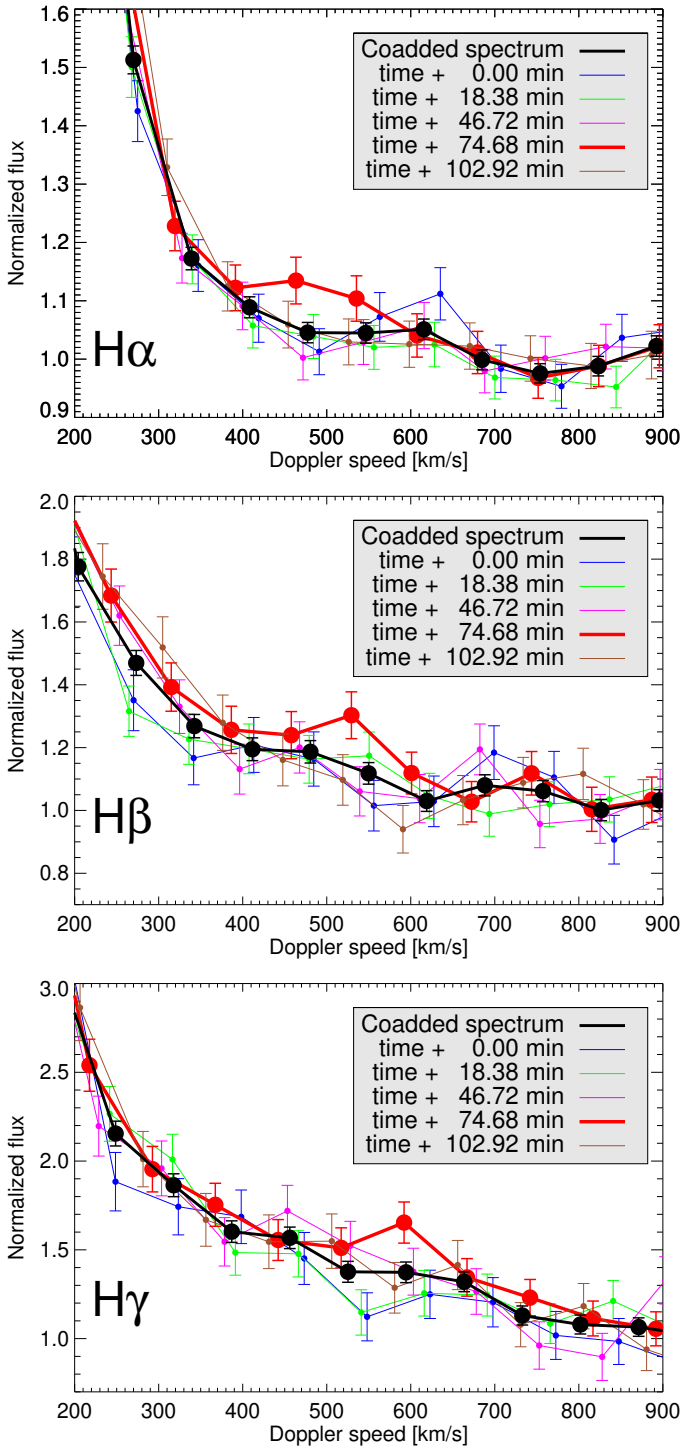


Fig. 11. Red wing of CME candidate 1 in $H\alpha$ (upper panel), $H\beta$ (middle panel), and $H\gamma$ (lower panel) for five time steps. The spectral line at the time step showing the enhancement is marked in red.

these spectra². Spectrum 4 shows a significant increase in flux on two neighboring data points in the red wing of $H\alpha$ as well as one (not significantly enhanced) higher data point on $H\beta$ and one significantly enhanced data point in $H\gamma$. We recall that the

² We note that the enhancement in the spectrum at about -1000 km s^{-1} is due to a molecular band that is not considered in the linear normalization procedure and thus can be ignored. The same holds for Figs. 14, 17, 19, and 21.

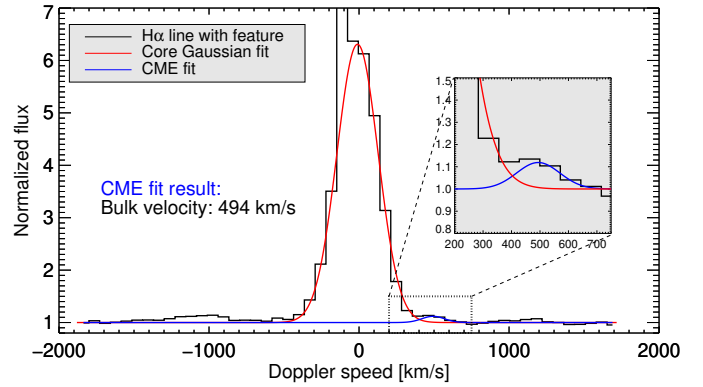


Fig. 12. $H\alpha$ spectral line of CME candidate 1. The line core (red line) as well as the possible CME feature (blue line) is fitted, with the bulk velocity resulting from the Gaussian CME fit annotated. A cosmic-ray hit affects the spectrum near the $H\alpha$ line peak. The inset shows a zoom into the spectral range with the enhanced wing.

wavelength axis is oversampled with respect to the spectral resolution, making single enhanced data points hard to interpret. We note, however, that these enhancements are significantly higher in two of the Balmer lines and occur at a similar Doppler shift. The star exhibits a higher S/N than other SDSS spectra. It shows strong chromospheric emission lines that peak at approximately six times the flux of the surrounding continuum. The maximum Doppler velocity (manually read out from the spectrum) is approximately 600, 630, and 640 km s^{-1} for $H\alpha$, $H\beta$, and $H\gamma$, respectively. A bulk velocity of $\sim 500 \text{ km s}^{-1}$ was derived using a Gaussian fit in the CME signature in $H\alpha$. Figure 12 shows the fitted Gaussian profiles to the Balmer line core and the enhanced feature. The mass estimated from integrated flux measurements in $H\alpha$ using Eq. (12) is approximately $3 \times 10^{17} \text{ g}$. The value is just an order of magnitude estimate and should be treated carefully because the method uses linear interpolation over the two data points to integrate over the enhanced range. The mass estimate based on the $H\gamma$ line is too uncertain because only one data point in the spectrum is enhanced.

While the quality of the plate is categorized as “bad” by SDSS, no data point flags in the wings of the emission lines indicate problematic data (only the $H\alpha$ line core is flagged in several spectra).

Candidate 2: SDSS J085556.66+174942.0 (Figs. 13 and 14). The star with subtype M 4e (SDSS classification M 4) was found with the HTM algorithm while searching for flares. The flare is part of the M-star flare in the high S/N bin. The coadded spectrum consists of two sets, each containing two consecutive spectra. The latter two spectra (which are observed two days after the first set) show higher Balmer emission lines that are interpreted as flaring activity. Figure 13 shows the spectra near the $H\alpha$, $H\beta$, and $H\gamma$ lines. Spectrum 3 shows a clear and significantly enhanced broad red wing in $H\alpha$ over seven neighboring data points. Two data points are enhanced in $H\beta$ and one data point in $H\gamma$, but none of them significantly. The noise in other chromospheric lines is too high to detect wing enhancements while the flare is still visible.

The maximum Doppler velocity derived is 710, 300, and 310 km s^{-1} for $H\alpha$, $H\beta$, and $H\gamma$, respectively. A bulk velocity of $\sim 300 \text{ km s}^{-1}$ is deduced, while the Gaussian fit to the CME signature in $H\alpha$ (see Fig. 14) did not work properly (because the normalization was not ideal on this wing). The mass estimated from integrated flux measurement in $H\alpha$ is $6.4 \times 10^{17} \text{ g}$.

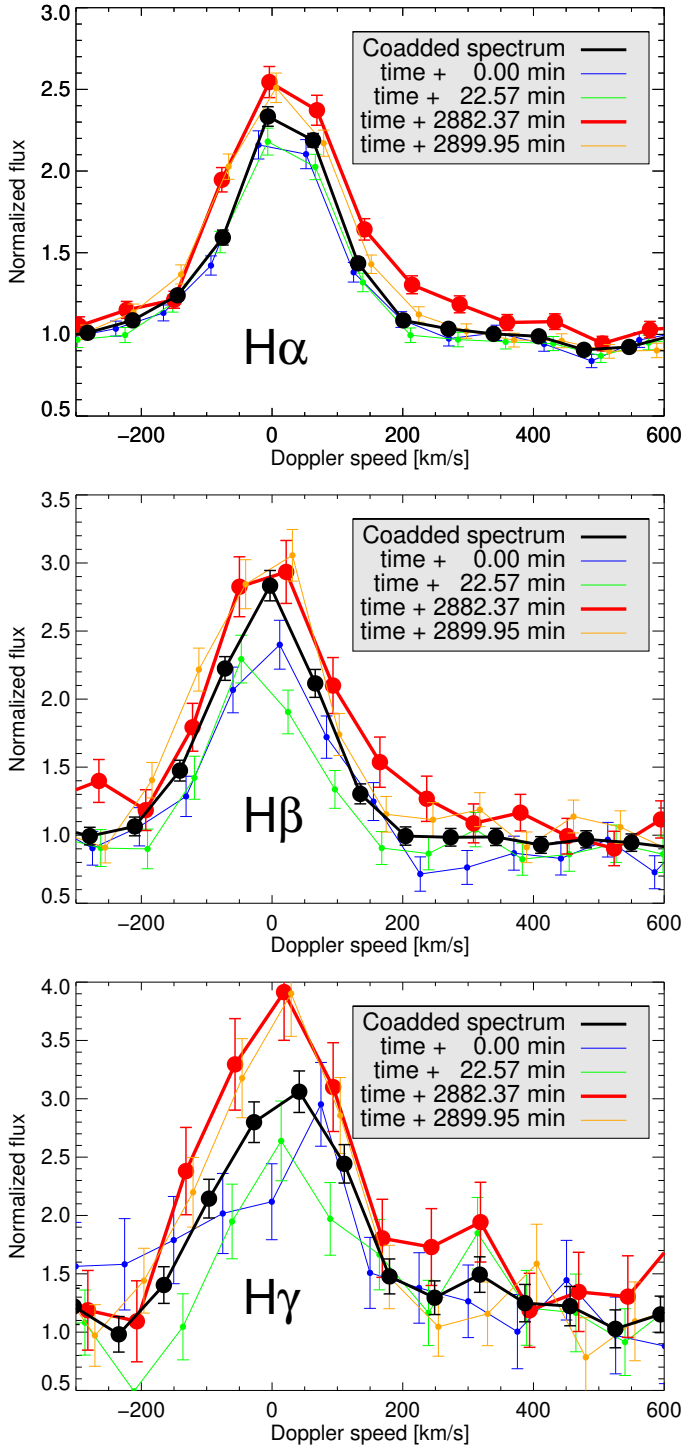


Fig. 13. CME candidate 2 in $H\alpha$ (upper panel), $H\beta$ (middle panel), and $H\gamma$ (lower panel). The spectral line at the time step showing the enhancement is marked in red.

The plate quality is categorized as “good” by SDSS. Four data point flags are present in the flaring Balmer emission line, which is not unusual for flaring activity.

Candidate 3: *SDSS J110743.30+075624.9* (Fig. 15). The star with subtype M 4e (SDSS classification M 4) was found with the ELM algorithm. The coadded spectrum consists of two sets, each containing two consecutive spectra, with the latter set being observed four days after the first set. The exposure times

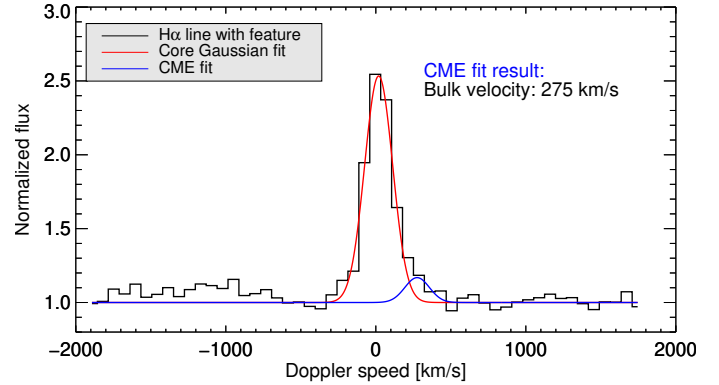


Fig. 14. $H\alpha$ spectral line of CME candidate 2. The line core (red line) as well as the possible CME feature (blue line) is fit; the bulk velocity of the Gaussian CME fit is annotated.

are 20 min for the first two spectra, and 12 and 10 min for the latter two. Figure 15 shows the red wing of the $H\alpha$, $H\beta$, and $H\gamma$ lines of these spectra. Spectrum 4 shows both an enhanced broad red wing in $H\alpha$ (between five and eight neighboring data points) and an enhanced blue wing (two neighboring data points). The enhanced data points in the red wing on their own are not significantly higher, that is, not above the error bars, but it is remarkable that there eight neighboring points are higher than all the other spectra in this range. The red $H\beta$ wing shows two enhanced data points (one significantly higher), and the red $H\gamma$ wing shows a clear enhancement of six higher points; three points are significantly enhanced. While it not significant considering the error bars, a red wing enhancement can also be seen in $H\delta$, the overlapping $He/Ca II H$ line, and the $H8$ line (but none is visible in Ca II K). Because there appears to be an enhancement in all of these lines, the assumption that this is a CME becomes more probable, even though most of the data points are not significantly enhanced.

The maximum Doppler velocity is 670, 300, and 650 km s^{-1} for $H\alpha$, $H\beta$, and $H\gamma$, respectively. A bulk velocity of $\sim 300 \text{ km s}^{-1}$ is deduced. Fitting a Gaussian to the CME signature in $H\alpha$ failed on the normalized flux of the spectrum with the enhanced wing. The mass estimated from the integrated flux measurement in $H\alpha$ and $H\gamma$ is 6×10^{16} and 1.3×10^{17} g, respectively. The integration over data points is possible in this case because of several neighboring data points in $H\alpha$ and $H\gamma$ (using linear interpolation).

The plate quality is categorized as marginal by the SDSS. No important data point flags affect this analysis, indicating no problem with the data.

Candidate 4: *SDSS J042139.64+264913.8* (Figs. 16 and 17). The star with subtype M 6 (SDSS classification M 8) was found with the HTM algorithm when we searched for flares. The flare is part of the M-star flare in the medium S/N bin. The coadded spectrum consists of five single spectra observed within four hours. Figure 16 shows the spectra near the $H\alpha$, $H\beta$, and $H\gamma$ lines. While the first three spectra show quiet Balmer emission lines (down to $H\gamma$), the fourth spectrum shows flaring activity with an enhanced peak and a broad red wing that is visible in the Balmer lines down to $H\delta$. The fifth and last spectrum is observed one hour after the last spectrum, only showing an enhanced emission line core (which is higher in $H\alpha$, of similar height in $H\beta$, and somewhat lower than the previous one in $H\gamma$, but it is not visible in $H\delta$ at all). The red-wing enhancement shows seven significantly enhanced data points in $H\alpha$, eight

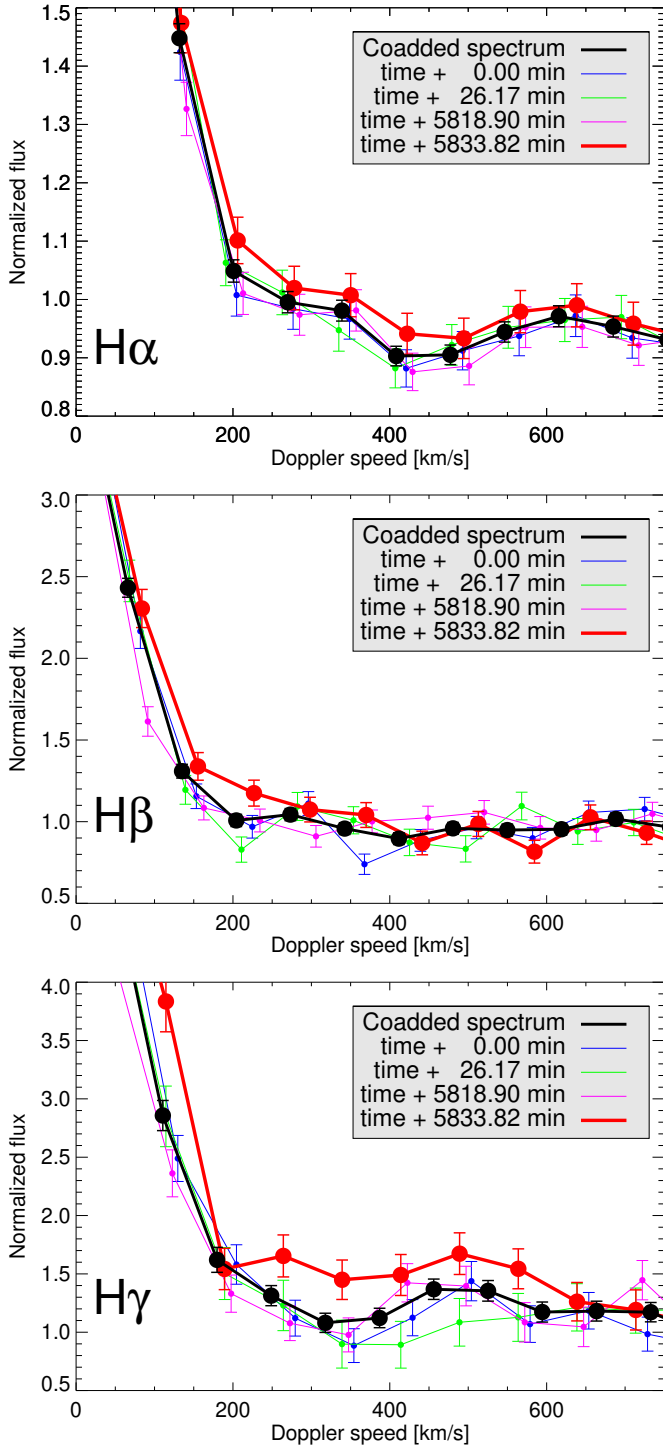


Fig. 15. Red wing of CME candidate 3 in $H\alpha$ (upper panel), $H\beta$ (middle panel), and $H\gamma$ (lower panel). The spectral line at the time step showing the enhancement is marked in red.

higher data points (three that are significantly enhanced) in $H\beta$, and two higher but not significantly enhanced points in $H\gamma$. An enhancement is also visible in $H\delta$, although it is not significantly higher. A cosmic-ray hit disturbs the blue wing of $H\beta$ in this spectrum.

The maximum Doppler velocity we derived is 700, 700, and 240 km s⁻¹ for $H\alpha$, $H\beta$, and $H\gamma$, respectively. A bulk velocity of ~ 360 km s⁻¹ was derived using a Gaussian fit to the CME signature in $H\alpha$. Figure 17 shows the fitted Gaussian profiles to the

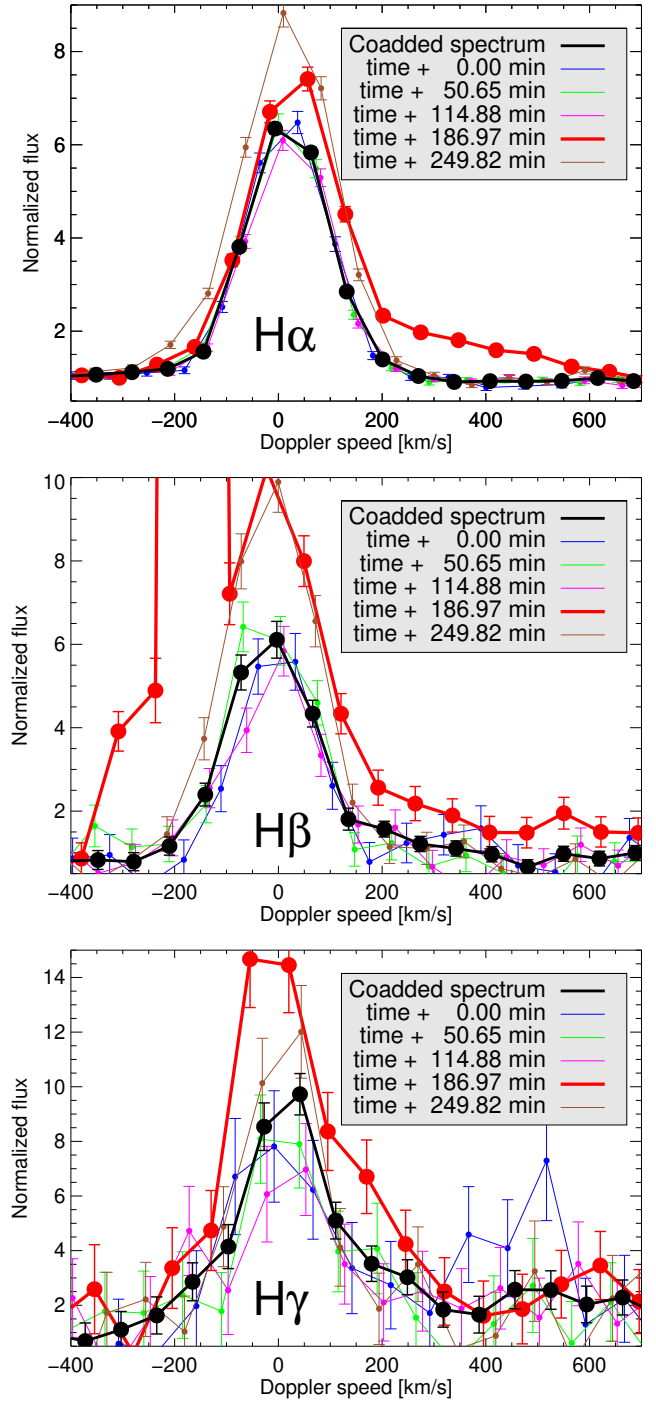


Fig. 16. CME candidate 4 in $H\alpha$ (upper panel), $H\beta$ (middle panel), and $H\gamma$ (lower panel). The spectral line at the time step showing the enhancement is marked in red.

Balmer line core and the enhanced feature. The mass estimated from integrated flux measurements in $H\alpha$ is 8.5×10^{17} g. While the enhanced feature in $H\alpha$ is sufficiently broad to be safely integrated, the feature only shows one clearly enhanced data point in $H\gamma$, thus the mass estimation was omitted in $H\gamma$.

The plate quality is categorized as “bad” by the SDSS. The flaring $H\alpha$ and $H\beta$ lines are mostly flagged, as is the whole emission feature in the $H\alpha$ red wing. Because these values deviate strongly from the quiet spectra, the flagging of these values is not unexpected.

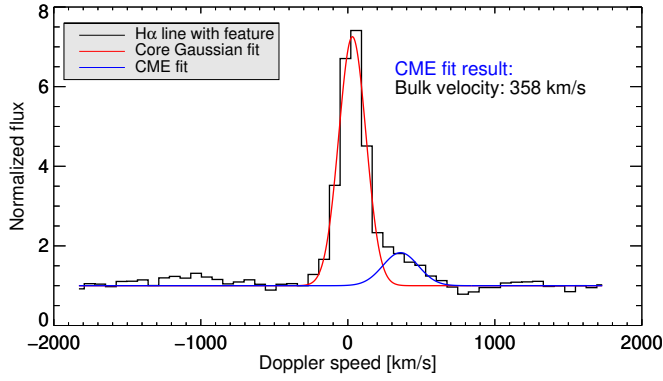


Fig. 17. $H\alpha$ spectral line of CME candidate 4. The line core (red line) as well as the possible CME feature (blue line) is fit; the bulk velocity of the Gaussian CME fit is annotated.

Candidate 5: *SDSS J053709.91-011050.3* (Figs. 18 and 19). Candidate star 5 *SDSS J053709.91-011050.3* is reported as a WTTS with SDSS classification M 5 and was found with the HTM algorithm when we searched for flares. The flare is part of the M-star flare in the high S/N bin. The coadded spectrum consists of five single spectra, the first with an exposure time of 15 min, and the other four spectra build a consecutive time series one day after the first spectrum, also with 15 min exposure time each. Figure 18 shows the spectra near the $H\alpha$, $H\beta$, and $H\gamma$ lines. The last spectrum shows a large flare in the chromospheric lines, which includes a broad flux enhancement in the red wing. The flare and flux excess is visible in all available Balmer lines down to $H\zeta$. The enhancement is also visible in the Ca II K line, but without the accompanying flare. The wing enhancement is significantly higher than the quiet spectrum in at least six neighboring data points in $H\alpha$, $H\beta$, and $H\gamma$ and is still significantly higher with at least one data point in $H\delta$, $H\epsilon$, and Ca II K.

The maximum Doppler velocity we derived is 710, 760, and 590 km s^{-1} for $H\alpha$, $H\beta$, and $H\gamma$, respectively. A bulk velocity of $\sim 340 \text{ km s}^{-1}$ was derived using a Gaussian fit to the CME signature in $H\alpha$. Figure 19 shows the fitted Gaussian profiles to the Balmer line core and the enhanced feature. There is no mass estimate because no distance is available for the star.

The plate quality is categorized as “bad” by SDSS. Most data points of the $H\alpha$ emission line are flagged in all spectra, and the red wing excess is also flagged, which is expected because of the high difference of these values from the quiet spectrum.

Candidate 6: *SDSS J052839.66-000322.5* (Figs. 20 and 21). Candidate star 6, *SDSS J052839.66-000322.5*, other designation “Kiso A-0903 163”, is reported as an M 4e type (SDSS classification M 4) and was found with the HTM algorithm when we searched for flares. The flare belongs to the M-star flare in the high S/N bin. The coadded spectrum consists of five consecutive single spectra, with exposure times of 15 min for spectra 1 and 5, and 25 min for spectra 2, 3, and 4. Figure 20 shows the spectra near the $H\alpha$, $H\beta$, and $H\gamma$ lines. The spectra show a decaying flare with a maximum in the first spectrum. The first three spectra also show an enhanced blue wing in $H\alpha$, which also decays over time. The entire blue wing (four datapoints) of the first two spectra in $H\alpha$ is significantly enhanced compared to the last spectrum, which shows the least activity. The blue wing is also enhanced in $H\beta$, but only two data points in the blue wing in each spectrum 1 and spectrum 2 are significantly higher than the quiescent spectrum. While an enhancement is still visible in $H\gamma$, no data point

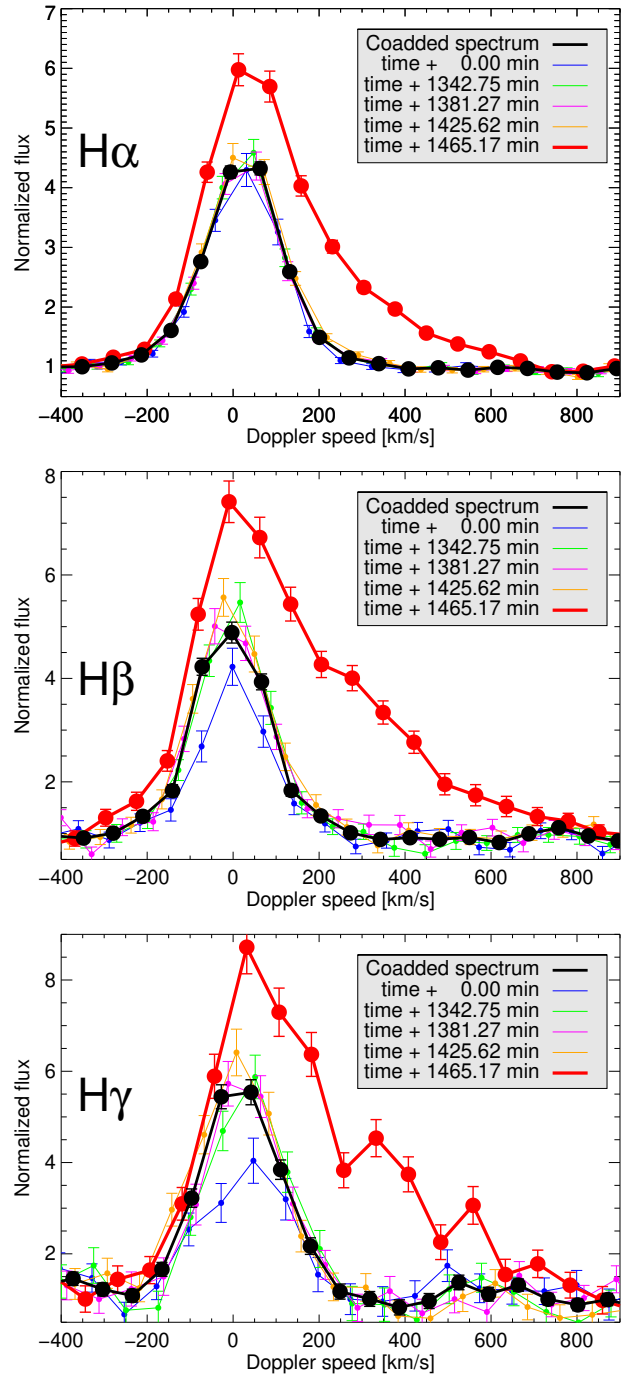


Fig. 18. CME candidate 5 (*SDSS J053709.91-011050.3*) in $H\alpha$ (upper panel), $H\beta$ (middle panel), and $H\gamma$ (lower panel). The spectral line at the time step showing the enhancement is marked in red.

is significantly above the error bars of the quiet spectrum. The feature in $H\gamma$ appears to be at higher velocities than in the other Balmer lines. The enhancement also appears in $H\delta$, and two of the four enhanced data points are significantly higher.

The maximum Doppler velocity we derived is -365 , -320 , and -450 km s^{-1} for $H\alpha$, $H\beta$, and $H\gamma$, respectively. A bulk velocity of about -240 km s^{-1} was derived using a Gaussian fit to the CME signature in $H\alpha$. Figure 21 shows the fitted Gaussian profiles to the Balmer line core and the enhanced feature. The mass estimated from integrated flux measurements in $H\alpha$ using

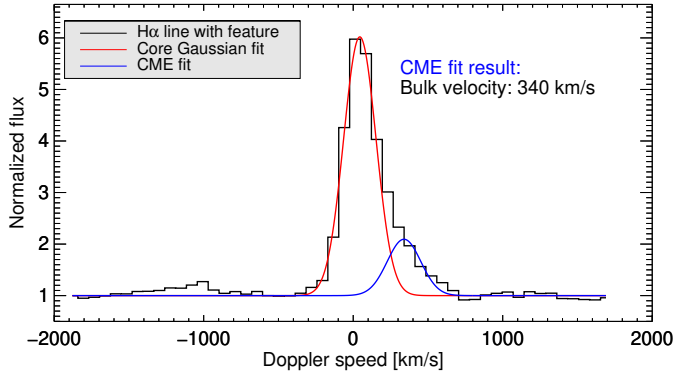


Fig. 19. $H\alpha$ spectral line of CME candidate 5. The line core (red line) as well as the possible CME feature (blue line) is fit, and the bulk velocity of the Gaussian CME fit is annotated.

Eq. (12) is 6.1×10^{18} g. While the enhanced feature in $H\alpha$ is sufficiently broad to be safely integrated, the feature appears weaker in $H\gamma$, thus the $H\gamma$ mass estimation was omitted.

The plate quality is categorized as “bad” by SDSS. The $H\alpha$ line is flagged in every spectrum with the flag “BADSKYCHI”. The blue-wing enhancements in the first three spectra are flagged with the flag “COMBINEREJ”, which is a common flag for unusual enhancements. It states that these points are rejected for the coadded spectrum.

5. Discussion

5.1. Difference to Hilton et al. (2010)

Hilton et al. (2010) performed a search for flares on M stars within an older data release of the spectroscopic SDSS data, finding 63 flaring events. While there is obvious overlap with our study, the differences in method and findings are substantial. Data release 14 that we used here covers additional and different objects compared to DR6, which was used by Hilton et al. (2010). Hilton et al. (2010) preselected their sample using photometric SDSS flags and color selection to obtain M-star spectral types, while our selection of stars is solely based on the original SDSS spectral classification with restriction to F, G, K, and M dwarf or subgiant stars. Hilton et al. (2010) made high and medium S/N samples and did not consider low S/N objects. Their quality cuts were based on the continuum level compared to the standard deviations. The emission line strength was measured by Hilton et al. (2010) by defining a “flare line index” (FLI), where the mean value within the line region minus the continuum flux was divided by the continuum standard deviation. It measures the strength of the line and the significance compared to the surrounding noise. Their flare criteria were either based on difference in FLI values between maximum and minimum exposures in $H\alpha$ and $H\beta$ or on threshold FLI values in strong emission lines. The latter method used a null distribution of inactive stars to determine the quiet emission line value and declared a flare threshold for each subtype, each $H\alpha$ and $H\beta$ line, and each S/N sample based on false-discovery rate analysis. This method used by Hilton et al. (2010) is quite different from the methods presented here, where Gaussian fits were used to solve the low S/N problem in the data. We also required temporal changes in the emission to be mandatory to declare a flaring state.

Of the 63 flares reported by Hilton et al. (2010), 41 (65%) are part of the flares we found. Of the missing 22 flares, 8 were found by Hilton et al. (2010) through their strong emission line criteria

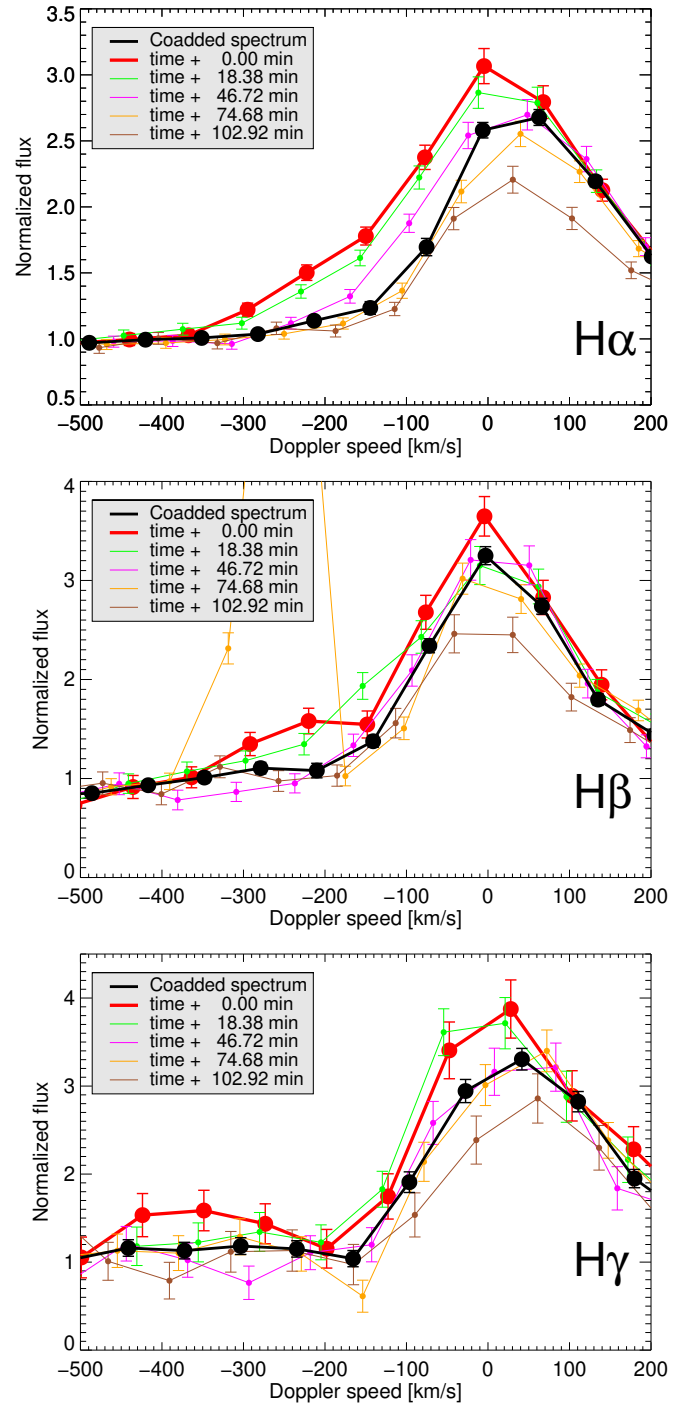


Fig. 20. Blue wing of CME candidate 6 (SDSS J052839.66-000322.5) in $H\alpha$ (upper panel), $H\beta$ (middle panel), and $H\gamma$ (lower panel). The spectral line at the time step showing the enhancement is marked in red.

only, which explains that we did not detect them here because they showed too little or no amplitude changes to declare them as flaring based on our categorizations, even after visual inspection. Of the remaining 14 flares, one was found by the LTM and declared as too small for flaring condition. Two were found by the ELM search but were declared as nonflaring after visual inspection. One was found by the ELM search in the lowest S/N bin. Because of too many false detections, the objects detected by ELM in the lowest S/N bin were not analyzed further, therefore this flare was overlooked. Ten remaining flares were not detected

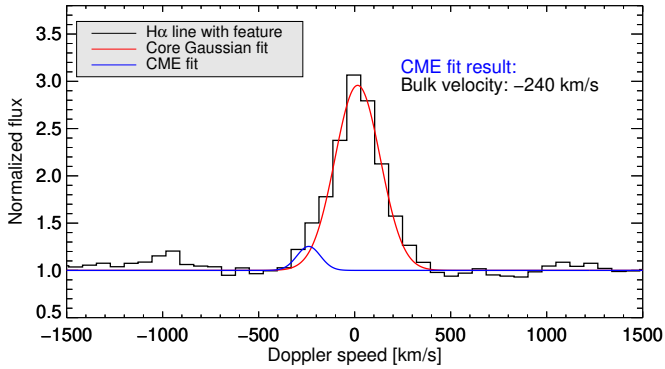


Fig. 21. $H\alpha$ spectral line of CME candidate 6. The line core (red line) as well as the possible CME feature (blue line) is fit, and the bulk velocity of the Gaussian CME fit is annotated.

by our algorithms. To understand why they were not detected, these ten objects are discussed in detail in Appendix A. The nondetections are caused by too weak changes in one Balmer line, missing data points, or by cosmic-ray hits in the reference spectrum that is used for the direct comparison.

5.2. Constraints, statistics, and significance of detections

The limited data quality allowed only a few flare detections compared to the size of the sample. The short and more or less random observational windows further complicated the search, and only allowed findings through the size of the dataset. K and M stars (which have a higher probability to show flares) make up 24.1% and 17.7% of the stars in the original dataset. Only 9849 stars or 1.6% of the original sample are M stars with high S/N, which is the subsample that showed the most flares (101 flares, which is 1% of all high S/N M stars) and all CME wing enhancement candidates. Of all M stars in the dataset, 70% have a very low S/N of < 10 near the $H\alpha$ line. These stars in the lowest S/N bin were still considered in the search. The significance of our flare findings in all S/N bins are reinforced by the (partly overlapping) flares reported by Hilton et al. (2010) in the same dataset.

5.3. Emission line stars in the flaring sample subset

The sample of 281 flaring stars is clearly dominated by emission line stars, that is, by stars that show Balmer lines in emission even without a flare. The SDSS spectral classification does not explicitly state whether a star is an emission line star. Taking all stars where we found a source for the spectral type that explicitly state the emission line status (249 stars or 88.6% of all flaring stars), we find that 90% of these stars are declared to be emission line stars. After visual inspection, we find that no star in our resulting flaring list shows the $H\alpha$ line in absorption, even considering all nonflaring spectra. We find 14 stars (5%) that showed neither emission nor absorption Balmer lines in quiescence, but significant Balmer emission lines during the flare. Ten of these 14 objects are part of the lowest S/N categorization.

Overall, these findings are expected: Balmer lines in emission are the indicator for magnetic activity, and our methods focused on fitting these lines (although both HTM and LTM should work on absorption lines as well). The low S/N of our data allowed us to only find flares on late-type stars with low luminosity and no flares on earlier types (F, G) that usually show Balmer lines in absorption.

5.4. Sample purity validation

Validating the purity of the luminosity type classification is an important step toward a robust flare detection statistic. Multi-color photometry, for example, can only determine the effective temperature, but not the luminosity class of an object and therefore cannot distinguish between close, faint dwarfs and distant, luminous red giant stars. Mathur et al. (2016) performed asteroseismic analysis of 45400 stars, which were classified as dwarfs in the *Kepler* input catalog based on multicolor photometric classification. Mathur and collaborators found that $\sim 2\%$ of the sample were cool giants.

In light of this finding and because most spectral subtypes determined by the SDSS are given without luminosity class, it is sensible to confirm whether the detected flaring stars are indeed dwarf stars. Based on our analysis in Sect. 4.2.1, we find that only one of our 281 flaring stars appears to be a giant or subgiant K star.

5.5. Flare energy and flaring occurrence

The calculated flare energies cover a wide range from 3×10^{28} erg to 2×10^{33} erg. With a dataset this large, some findings of extraordinary strong flares (beyond 10^{32} erg in $H\alpha$) are expected (compare with stellar $H\alpha$ flare energies of approximately 10^{29} – 10^{30} erg, Hawley et al. 2003; Leitzinger et al. 2014). Recent spectroscopy analysis of the active M dwarf AD Leo by Muheki et al. (2020) derived similar results, with the strongest $H\alpha$ flare emitting an energy of 2.12×10^{32} erg.

The correlation between the energy or luminosity of the flares and the S/N of the objects is unclear from the whole dataset. This is a result of the mix of different spectral subtypes. Considering one specific subtype (e.g. M 6 in Fig. 8), we find that lower flare energies are found at higher S/N values (equivalent to stars that are closer to us), which is expected.

The lower threshold of flare energy compared to the distance (see Fig. 9) was expected because only large flares are observable on distant stars. On the other hand, there is a visible upper threshold showing no high-energy flares within the surroundings of the Solar System. Brasseur et al. (2019) found similar relations between flare energies and distances as an effect of observational sensitivity. The most probable explanations are selection effects in the data. With increasing distance, the volume and thus the number of stars and the probability of detecting high-energy flares obviously increases. We verified the distances per subtype and found that in the given dataset only few earlier subtypes (K stars and M 0 or M 1 stars) are close to the Sun (within 200 pc). This leads to missing detections of high-energy flares on nearby stars because energetic events occur mainly on earlier types. Our findings that earlier types show higher maximum flare energies (see Fig. 8) is consistent with previous reports in the literature (e.g. Davenport 2016, Fig. 5 or Balona 2015, Fig. 10).

The flaring occurrence rates shown in Fig. 5 are compatible with the result in Kowalski et al. (2009), showing a rising flaring occurrence rate until spectral type M 6. Unlike our result, which used a spectroscopic analysis of Balmer lines, Kowalski et al. (2009) used photometric light curves, giving the fraction of flaring epochs. West et al. (2011) determined the magnetic activity fraction for SDSS M dwarfs (based on “quiet” chromospheric emission) and reported the same relation. This further confirms our flaring occurrence results because flares are often related to high magnetic activity.

5.6. Flare evolution

The higher number of rising flares compared to decaying flares given in Table 5 is most probably an effect of the data and the search method we used. Our algorithm is most sensitive to strong changes in consecutive spectra. Decaying spectra may show too little changes to be detected due to the more gradual evolution. A second factor might be that the short rising period of flares (expected to last less than 10 min, see, e.g., $H\alpha$ flare evolution curves in Kowalski et al. 2013) may be shorter than the exposure times (up to 25 min) of the single spectra. Most of our rising flare cases appear to show two consecutive rising spectra (resulting in 20–50 min), which is in contrast to the expected short timescale of the rise phase of flares. We could explain this as follows: the first flaring spectrum could show a mix of quiet and flare rise flux, including the peak. The integrated flux might therefore be lower than the subsequent spectrum, where the flux already decreases, but the integrated flux level is higher than the first spectrum. This would give the impression that the first and second flaring spectrum both increase and thus show the flare rise, although the flare is already in the decaying phase in the second spectrum. This would intrinsically favor rise phases over decay phases.

Some M star spectra show highly active Balmer lines with significant changes, sometimes showing several possible flares observed days apart from each other. The more prominent and easier to categorize flaring phase was chosen to represent the evolution of each set of spectra. The possibility of several flares in one set of spectra is thus not accounted for.

5.7. CME candidates

Of the six CME candidates showing Balmer line wing excesses presented in Sect. 4.3, five show a red and one a blue wing enhancement. CME candidates 2 (Fig. 13), 4 (Fig. 16), and 5 (Fig. 18) show an enhanced red wing during a flare. If this is a CME feature, it is counterintuitive to what is expected from the standard model of a CME onset because the flare together with the ejected material is directed toward the observer and therefore should be blueshifted. There are possible explanations for this feature: (a) a backward-directed CME related to a flare occurring close to the limb; (b) the visible feature is cool plasma material that falls back onto the surface of the star, albeit with very high velocities; (c) chromospheric condensation (Kowalski & Allred 2018), or (d) a broad flare line profile superimposed with blue wing absorption from a CME, causing an apparent red-wing asymmetry. For CME candidate 2 (Fig. 13) in particular, we stress that the last spectrum no longer shows the red-wing enhancement, while the line core is still distinctly enhanced compared to the quiet spectra. A possible explanation in this case (other than downfalling material) is that the CME material becomes ionized on its way, making it invisible in Balmer lines in the last spectrum.

The CME candidate 6 (Fig. 20) shows a strong and broad wing enhancement during the flare similar to candidates 2, 4, and 5. The excess flux is blueshifted, however, which is in agreement with the interpretation as a CME.

The CME candidate 1 (Fig. 11) has significant red-wing enhancements in $H\alpha$, $H\beta$, and $H\gamma$, but only with either one or two enhanced data points each, making the detection more doubtful. CME candidate 3 (Fig. 15), on the other hand, shows several enhanced data points in the red wing of each Balmer line, rendering it a good example of a possible CME in this dataset. Most of the enhanced data points for candidate 3 are not

significantly higher than the other spectra, but the large number of neighboring and consistently higher data points and the appearance of the enhancement in eight chromospheric lines gives reason to take this event into consideration. Candidates 1 and 3 show no accompanying flare, which is in agreement with the theory of a backward-directed CME.

5.8. CME properties, line asymmetries, rates, and nondetections

The derived CME masses range from 6×10^{16} g to 6×10^{18} g. This lies within the expected range (10^{14} – 10^{19} g, see Table 1 in Moschou et al. 2019). Odert et al. (2020) estimated the S/Ns that are required to detect a certain CME mass depending on stellar spectral type. Their estimate for emission features (see Fig. 1 in Odert et al. 2020) shows that for stars of type M 5.5, at least an S/N of ~ 20 is required to detect CME masses above 6×10^{15} g in $H\alpha$. For stars of type M 2, the detectable CME masses with an S/N of ~ 20 rise to $\sim 5 \times 10^{16}$ g in $H\alpha$. The S/N near $H\alpha$ of the identified CME candidates ranges between 17 and 28, and thus their derived masses lie well above the estimated minimum detectable value calculated by Odert et al. (2020).

The number of CME candidates is surprisingly low considering the minimum S/N – CME mass detection calculations by Odert et al. (2020), the assumed high association rate between high-energy flares and CMEs (based on solar relations), and the large number of high-activity late-type M stars. Still, this finding is consistent with other studies from literature. Leitzinger et al. (2020) used a similar approach by using optical spectral archive data and searched for CMEs on dF, dG and dK stars, finding no CME-like features and only little flaring activity on 425 stars with more than 3700 h of observing time. Muheki et al. (2020) analyzed $H\alpha$ line asymmetries in high-resolution spectroscopy of the active M star AD Leo with a total monitoring time of about 240 h. All determined Doppler-shift velocities of the study (blueshifts up to 260 km s^{-1} and redshifts up to 190 km s^{-1}) lie well below the escape velocity of the star.

Explanations for our sparse findings based on the dataset quality may be a) low spectral resolution or b) flares giving wing enhancements that may outshine a possible CME-like feature. The long exposure times are another explanation for nondetection because the long integration times may diminish a short flux enhancement. Moreover, the minimum S/N – CME mass detection calculations by Odert et al. (2020) are estimates for the maximum possible flux, further limiting the chances of detecting a CME, especially at low S/N.

A physical explanation for sparse CME findings may be that CMEs are suppressed by the strong overlying magnetic fields of the star, as was suggested by numerical simulations of stellar CMEs on M dwarfs by Alvarado-Gómez et al. (2018). This would result in a significant reduction of velocity, suppressing weaker CMEs while only strong ejections could break free (Vida et al. 2019). Nonetheless, no clear and detached emission features were found. Another explanation lies in a possible strong ionization of the prominence material under the intensive EUV emission of active M stars (Howard 2015).

Alternative explanations for asymmetries arise by using the connection to flares (Canfield et al. 1990). Balmer line asymmetries are generally associated with chromospheric material moving either downward or upward as a reaction to the impulsive energy input by flare-accelerated electrons causing chromospheric evaporation. The associated velocities to this mechanism on the Sun are usually in the range of tens of km s^{-1} in chromospheric spectral lines, that is, far below the estimated

velocities of our CME candidates (Li et al. 2019). Cho et al. (2016) reported strong preflare $H\alpha$ blueshifts on the Sun, with a wide velocity range of -130 to 38 km s^{-1} , while the bulk velocity was low. The broadening was spatially correlated to a rising filament; however, there was no reported CME associated with this event. Using observations of a dMe 4.5 star, Gunn et al. (1994) reported a blue-wing emission feature similar to our CME candidates as a high-velocity chromospheric evaporation (maximum projected velocity of 600 km s^{-1}) during a flare, while the possibility of an associated CME was not discussed. Notably, their emission feature is also visible in the Ca II H & K lines, similar to our CME candidate 5, and both show similar velocities and asymmetry shapes, but they have opposite directions.

Redshifted $H\alpha$ emission may arise from coronal rain (cooling plasma flows along post-flare loops) or from chromospheric condensation during the impulsive flare phase (Fuhrmeister et al. 2018; Vida et al. 2019). Redshifts attributed to coronal rain in the solar case show velocities up to 120 km s^{-1} with average values of roughly $60\text{--}70 \text{ km s}^{-1}$ (Antolin et al. 2012; Lacatus et al. 2017; Fuhrmeister et al. 2018). Chromospheric condensation arising from downward-moving material due to nonthermal electron beams is observed during the impulsive phases of solar flares with reported velocities of about $40\text{--}100 \text{ km s}^{-1}$ (Ichimoto & Kurokawa 1984). More recent solar $H\alpha$ red-wing asymmetries arising from chromospheric condensation reported by Asai et al. (2012) are at about 50 km s^{-1} . These redshifted velocities are significantly below the derived values of our CME candidates 1–5 with bulk velocities of $\sim 300 \text{ km s}^{-1}$ and maximum velocities up to 700 km s^{-1} . This might be an indicator that processes with higher velocities (such as CMEs) occur, although the measured bulk velocities are still below the escape velocities of the stars (which are about $500\text{--}600 \text{ km s}^{-1}$). We also recall that all CME velocity estimates are projected velocities, which decreases the true value by an unknown factor.

Regarding the temporal evolution of our CME candidates, the low amount of spectra and long exposure times prohibit a detailed analysis. Still, CME candidates 2, 4, and 5 appear to show a red asymmetry during the first flaring spectrum, which might indicate chromospheric condensation. The velocities nevertheless significantly exceed chromospheric velocity values known from the Sun. Recent efforts in numerical simulations of the coronal response during magnetically suppressed CME events on M stars suggest that weakly suppressed CMEs might display a temporal evolution of Doppler shifts from red to blue with velocities up to 200 km s^{-1} , while fully suppressed CMEs might show redshifted emission from in-falling material with speeds of $< -50 \text{ km s}^{-1}$ (Alvarado-Gómez et al. 2019). However, we note that Alvarado-Gómez et al. (2019) primarily focused on hot plasma and X-ray emission, in contrast to the cool plasma that produces Balmer lines. Alternatively, Kuridze et al. (2015) showed that asymmetries in $H\alpha$ lines may also arise from shifts in the wavelength of the maximum opacity, in contrast to plasma motions.

This work did not employ an upper occurrence rate estimation on stellar CMEs due to the large fraction of spectra that are probably too low in S/N to perform a reasonable CME search. The overall observation time that is usable for CME analysis was therefore greatly reduced.

6. Conclusions and outlook

We searched optical SDSS spectra of 630 162 stars (consisting mostly of F, G, K, and M dwarfs) in order to find flares and CMEs using both automatic algorithms and visual inspections.

We found no flares on F and G type stars, 6 on K stars, 272 on M dwarfs, and 3 on M stars that are known to be WTT stars. We report 6 possible CME-like enhancements in the wings of Balmer lines and derive maximum velocities of up to 700 km s^{-1} and CME masses between $6 \times 10^{16} \text{ g}$ and $6 \times 10^{18} \text{ g}$.

With stellar CMEs playing a crucial role in the evolution of exoplanetary atmospheres and with their effect on stellar mass and angular momentum loss, it is important to constrain CME occurrences and parameters (Lammer et al. 2007; Airapetian et al. 2020). Like Hilton et al. (2010), we demonstrated that searching through large archival data proves a viable method to search for stellar flares in optical spectra, even with limited S/N. A search for associated CMEs is reasonable considering minimum S/N constraints for CME masses (Odert et al. 2020) and the high association rate between energetic flares and CMEs on the Sun (Yashiro & Gopalswamy 2009). While other works have implemented similar search approaches (e.g., Vida et al. 2019; Leitzinger et al. 2020), numerous data archives with spectral data have not yet been analyzed within the scope of stellar CME search.

Still, considerable information is lost in similar data archives because of the short observational time windows (and thus only few single spectra) for each star, compared to long and continuous observations of one single star. To solve the problem of blocking a telescope for long observations, a flare-alert system like the one presented by Hansmeier et al. (2017) might be a promising concept. When continuous photometric observations with small telescopes are used, a detected flare can be reported in real time to larger observatories so that they can conduct spectroscopic follow-up observations. While this work focused on main-sequence stars, we may shift the focus to include pre-main-sequence stars in order to search for CMEs because these objects are known for their high activity.

We focused on CME-like features that are visible at all Balmer lines, while the possibility of CME material being optically thick in one Balmer line only might give reason to reconsider this approach. This needs to be further analyzed.

Acknowledgements. F.K., M.L., and P.O. gratefully acknowledge the Austrian Science Fund (FWF): P30949-N36 for supporting this project. M.L., P.O., and A.M.V. gratefully acknowledge support by the Austrian Space Applications Programme of the Austrian Research Promotion Agency FFG (ASAP-14 865972, BMVIT). Funding for the Sloan Digital Sky Survey (SDSS) has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the US Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>. The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, Princeton University, the United States Naval Observatory, and the University of Washington. This research has made use of the VizieR catalogue access tool, CDS, Strasbourg, France (DOI : 10.26093/cds/vizieer). The original description of the VizieR service was published in Ochsenbein et al. (2000). This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

References

- Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, *ApJS*, 235, 42
 Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, *ApJS*, 203, 21
 Airapetian, V. S., Barnes, R., Cohen, O., et al. 2020, *Int. J. Astrobiol.*, 19, 136

- Alvarado-Gómez, J. D., Drake, J. J., Cohen, O., Moschou, S. P., & Garraffo, C. 2018, *ApJ*, **862**, 93
- Alvarado-Gómez, J. D., Drake, J. J., Moschou, S. P., et al. 2019, *ApJ*, **884**, L13
- Ambruster, C. W., Pettersen, B. R., Hawley, S., Coleman, L. A., & Sandmann, W. H. 1986, *ESA SP*, **263**, 137
- Antolin, P., Vissers, G., & Rouppe van der Voort, L. 2012, *Sol. Phys.*, **280**, 457
- Argiroffi, C., Reale, F., Drake, J. J., et al. 2019, *Nat. Astron.*, **3**, 742
- Asai, A., Ichimoto, K., Kita, R., Kurokawa, H., & Shibata, K. 2012, *PASJ*, **64**, 20
- Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, *AJ*, **156**, 58
- Balona, L. A. 2015, *MNRAS*, **447**, 2714
- Basri, G. 2019, *The Sun as a Guide to Stellar Physics*, eds. O. Engvold, J.-C. Vial, & A. Skumanich (Amsterdam: Elsevier), 363
- Becker, A. C., Bochanski, J. J., Hawley, S. L., et al. 2011, *ApJ*, **731**, 17
- Bell, K. J., Hilton, E. J., Davenport, J. R. A., et al. 2012, *PASP*, **124**, 14
- Boiko, A. I., Konvalenko, A. A., Koliadin, V. L., & Melnik, V. N. 2012, *Adv. Astron. Space Phys.*, **2**, 121
- Brasseur, C. E., Osten, R. A., & Fleming, S. W. 2019, *ApJ*, **883**, 88
- Briceño, C., Calvet, N., Hernández, J., et al. 2019, *AJ*, **157**, 85
- Butler, C. J., Rodono, M., & Foing, B. H. 1988, *A&A*, **206**, L1
- Canfield, R. C., Penn, M. J., Wulser, J.-P., & Kiplinger, A. L. 1990, *ApJ*, **363**, 318
- Cho, K., Lee, J., Chae, J., et al. 2016, *Sol. Phys.*, **291**, 2391
- Cook, N. J., Pinfield, D. J., Marocco, F., et al. 2016, *MNRAS*, **457**, 2192
- Crosley, M. K., & Osten, R. A. 2018a, *ApJ*, **856**, 39
- Crosley, M. K., & Osten, R. A. 2018b, *ApJ*, **862**, 113
- Davenport, J. R. A. 2016, *ApJ*, **829**, 23
- Den, O. E., & Kornienko, G. I. 1993, *Astron. Rep.*, **37**, 76
- Ding, M. D., Chen, Q. R., Li, J. P., & Chen, P. F. 2003, *ApJ*, **598**, 683
- Dissauer, K., Veronig, A. M., Temmer, M., Podladchikova, T., & Vanninathan, K. 2018, *ApJ*, **855**, 137
- Drake, J. J., Cohen, O., Yashiro, S., & Gopalswamy, N. 2013, *ApJ*, **764**, 170
- Drake, A. J., Graham, M. J., Djorgovski, S. G., et al. 2014, *ApJS*, **213**, 9
- Favata, F., & Schmitt, J. H. M. M. 1999, *A&A*, **350**, 900
- Fuhrmeister, B., & Schmitt, J. H. M. M. 2004, *A&A*, **420**, 1079
- Fuhrmeister, B., Czesla, S., Schmitt, J. H. M. M., et al. 2018, *A&A*, **615**, A14
- Gaia Collaboration (Brown, A. G. A., et al.) 2018, *A&A*, **616**, A1
- Gopalswamy, N., Akiyama, S., Yashiro, S., & Mäkelä, P. 2010, *Astrophys. Space Sci. Proc.*, **19**, 289
- Güdel, M. 2007, *Liv. Rev. Sol. Phys.*, **4**, 3
- Guenther, E. W., & Emerson, J. P. 1997, *A&A*, **321**, 803
- Guieu, S., Dougados, C., Monin, J. L., Magnier, E., & Martín, E. L. 2006, *A&A*, **446**, 485
- Gunn, A. G., Doyle, J. G., Mathioudakis, M., Houdebine, E. R., & Avgoloupis, S. 1994, *A&A*, **285**, 489
- Haisch, B. M., Linsky, J. L., Bornmann, P. L., et al. 1983, *ApJ*, **267**, 280
- Hanslmeier, A., Leitzinger, M., Greimel, R., et al. 2017, *Central Euro. Astrophys. Bull.*, **41**, 67
- Hawley, S. L., Allred, J. C., Johns-Krull, C. M., et al. 2003, *ApJ*, **597**, 535
- Heinze, A. N., Tonry, J. L., Denneau, L., et al. 2018, *AJ*, **156**, 241
- Hilton, E. J., West, A. A., Hawley, S. L., & Kowalski, A. F. 2010, *AJ*, **140**, 1402
- Houdebine, E. R., Foing, B. H., & Rodono, M. 1990, *A&A*, **238**, 249
- Howard, T. A. 2015, *ApJ*, **806**, 175
- Ichimoto, K., & Kurokawa, H. 1984, *Sol. Phys.*, **93**, 105
- Jones, D. O., & West, A. A. 2016, *ApJ*, **817**, 1
- Kleinman, S. J., Kepler, S. O., Koester, D., et al. 2013, *ApJS*, **204**, 5
- Koenig, X., Hillenbrand, L. A., Padgett, D. L., & DeFelippis, D. 2015, *AJ*, **150**, 100
- Korhonen, H., Vida, K., Leitzinger, M., Odert, P., & Kovács, O. E. 2017, *IAU Symp.*, **328**, 198
- Kowalski, A. F., & Allred, J. C. 2018, *ApJ*, **852**, 61
- Kowalski, A. F., Hawley, S. L., Hilton, E. J., et al. 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. 2009, *ApJ*, **704**, 531
- Kretzschmar, M. 2011, *A&A*, **530**, A84
- Kruse, E. A., Berger, E., Knapp, G. R., et al. 2010, *ApJ*, **722**, 1352
- Kuridze, D., Mathioudakis, M., Simões, P. J. A., et al. 2015, *ApJ*, **813**, 125
- Lacatus, D. A., Judge, P. G., & Donea, A. 2017, *ApJ*, **842**, 15
- Lammer, H., Lichtenegger, H. I. M., Kulikov, Y. N., et al. 2007, *Astrobiology*, **7**, 185
- Leitzinger, M., Odert, P., Hanslmeier, A., et al. 2009, *AIP Conf. Ser.*, **1094**, 680
- Leitzinger, M., Odert, P., Ribas, I., et al. 2011, *A&A*, **536**, A62
- Leitzinger, M., Odert, P., Greimel, R., et al. 2014, *MNRAS*, **443**, 898
- Leitzinger, M., Odert, P., Greimel, R., et al. 2020, *MNRAS*, **493**, 4570
- Li, Y., Ding, M. D., Hong, J., Li, H., & Gan, W. Q. 2019, *ApJ*, **879**, 30
- Luhman, K. L., Mamajek, E. E., Shukla, S. J., & Loutrel, N. P. 2017, *AJ*, **153**, 46
- Maehara, H., Shibayama, T., Notsu, S., et al. 2012, *Nature*, **485**, 478
- Mathur, S., García, R. A., Huber, D., et al. 2016, *ApJ*, **827**, 50
- McGehee, P. M. 2006, *AJ*, **131**, 2959
- Morgan, D. P., West, A. A., Garcés, A., et al. 2012, *AJ*, **144**, 93
- Moschou, S.-P., Drake, J. J., Cohen, O., Alvarado-Gomez, J. D., & Garraffo, C. 2017, *ApJ*, **850**, 191
- Moschou, S.-P., Drake, J. J., Cohen, O., et al. 2019, *ApJ*, **877**, 105
- Muheki, P., Guenther, E. W., Mutabazi, T., & Jurua, E. 2020, *A&A*, **637**, A13
- Mullan, D. J., & Paudel, R. R. 2019, *ApJ*, **873**, 1
- Newton, E. R., Charbonneau, D., Irwin, J., et al. 2014, *AJ*, **147**, 20
- Ochsenbein, F., Bauer, P., & Marcout, J. 2000, *A&AS*, **143**, 23
- Odert, P., Leitzinger, M., Hanslmeier, A., & Lammer, H. 2017, *MNRAS*, **472**, 876
- Odert, P., Leitzinger, M., Guenther, E. W., & Heinzel, P. 2020, *MNRAS*, **494**, 3766
- Pettersen, B. R. 1989, *Sol. Phys.*, **121**, 299
- Pevtsov, A. A., Bertello, L., & Marble, A. R. 2014, *Astron. Nachr.*, **335**, 21
- Rebassa-Mansergas, A., Agurto-Gangas, C., Schreiber, M. R., Gänsicke, B. T., & Koester, D. 2013, *MNRAS*, **433**, 3398
- Ribas, I., Guinan, E. F., Güdel, M., & Audard, M. 2005, *ApJ*, **622**, 680
- Skiff, B. A. 2014, *VizieR Online Data Catalog*, B/mk
- Sonett, C. P., Giampapa, M. S., & Matthews, M. S. 1991, *The Sun in Time* (Berlin: Springer)
- Stassun, K. G., Oelkers, R. J., Pepper, J., et al. 2018, *AJ*, **156**, 102
- Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, *AJ*, **123**, 485
- Suárez, G., Downes, J. J., Román-Zúñiga, C., et al. 2017, *AJ*, **154**, 14
- Taylor, R. J. 1997, *The Sun as a Star* (Cambridge: Cambridge University Press)
- Theissen, C. A., West, A. A., Shippee, G., Burgasser, A. J., & Schmidt, S. J. 2017, *AJ*, **153**, 92
- Van Doorsselaere, T., Shariati, H., & Debusscher, J. 2017, *ApJS*, **232**, 26
- Vida, K., Kriskovics, L., Oláh, K., et al. 2016, *AAP*, **590**, A11
- Vida, K., Leitzinger, M., Kriskovics, L., et al. 2019, *A&A*, **623**, A49
- Villadsen, J., & Hallinan, G. 2019, *ApJ*, **871**, 214
- Watson, C. L., Henden, A. A., & Price, A. 2006, *Soc. Astron. Sci. Ann. Symp.*, **25**, 47
- Webb, D. F., & Howard, T. A. 2012, *Liv. Rev. Sol. Phys.*, **9**, 3
- West, A. A., Hawley, S. L., Bochanski, J. J., et al. 2008, *AJ*, **135**, 785
- West, A. A., Morgan, D. P., Bochanski, J. J., et al. 2011, *AJ*, **141**, 97
- Wiese, W., & Fuhr, J. 2009, *J. Phys. Chem. Ref. Data*, **38**, 565
- Xue, X.-X., Ma, Z., Rix, H.-W., et al. 2014, *ApJ*, **784**, 170
- Yashiro, S., & Gopalswamy, N. 2009, *IAU Symp.*, **257**, 233
- Yashiro, S., Akiyama, S., Gopalswamy, N., & Howard, R. A. 2006, *ApJ*, **650**, L143
- Zhang, Z. H., Pinfield, D. J., Day-Jones, A. C., et al. 2010, *MNRAS*, **404**, 1817
- Zhong, J., Lépine, S., Li, J., et al. 2015, *Res. Astron. Astrophys.*, **15**, 1154

Appendix A: Missing flares of Hilton et al. (2010)

Using notation from Hilton et al. (2010), Table 3, we identified the stars by their right ascension and declination coordinates. We list them below

Star RA 355.347630 and Dec -0.646637. This M 1 star in the high S/N bin probably showed too little variation in the $H\alpha$ line height for detection. It was probably not found by the high emission line algorithm due to the high continuum flux compared to the $H\alpha$ line height.

Star RA 128.514690 and Dec 23.718951. This M 3 star in the high S/N bin shows flaring activity in later Balmer lines, but not in $H\alpha$.

Star RA 134.892240 and Dec 28.224011. This M 3 star in the high S/N bin shows small variation (decay) in one spectrum. It was probably not detected due to too small temporal changes in $H\beta$.

Star RA 154.133280 and Dec 36.066472. This M 3 star in the high S/N bin shows slowly decaying emission lines. The changes in $H\alpha$ are most likely too small for a detection.

Star RA 179.329230 and Dec 36.899333. This M 3 star in the high S/N bin shows too little changes in $H\alpha$. The spectrum showing the highest differences in the in $H\alpha$ emission line does not show the same behavior in $H\beta$.

Star RA 8.116274 and Dec -0.669970. This M 4 star in the high S/N bin shows a significantly high rising flare. There was probably no detection due to a cosmic-ray hit near the peaking spectrum in $H\beta$, which caused the Gaussian fit to fail.

Star RA 123.544780 and Dec 7.833261. This M 4 star in the medium S/N bin shows variation of the $H\alpha$ and $H\beta$ lines, but too little to be detected with our algorithms. The change from one time instant to another between flaring spectra in $H\beta$ appears to be too small.

Star RA 237.126050 and Dec 51.509009. This M 4 star in the high S/N bin shows a huge flare visible in all chromospheric lines. The flare was not detected due to missing data points in the wing of the $H\alpha$ line.

Star RA 329.517010 and Dec -8.355448. This M 6 star in the low S/N bin shows its peak Balmer line enhancement in its last spectrum. It was probably not detected because the previous spectrum in $H\beta$ showed a cosmic-ray hit, giving rise to a faulty comparison between the flaring and the previous spectra.

Star RA 182.070410 and Dec 8.757879. This M 9 star in the lowest S/N bin shows a clear flare in the Balmer lines in its last spectrum. It was probably not detected due to a cosmic-ray hit near the Balmer line in the previous spectrum in $H\alpha$.

Appendix B: Special cases in the flaring list

M stars. Several stars were identified to be cataclysmic variables, showing extreme shifts in the spectra and chromospheric line height variations. Together with other, possible close (and eclipsing) white dwarf + late-type binaries with extreme line shifts, they were excluded from the list.

After further inspection, several stars still present in the list are also possible WD+dM binaries. One object (SDSS

J045325.65-054459.2) was found to be most likely a late-type M star in front of a more distant star of an earlier type. This becomes clear from the available *Gaia* distances, which reveal two sources within two arcseconds. Rebassa-Mansergas et al. (2013) came to the same conclusion, stating the object as a superposition of two main-sequence stars.

Other stars given as possible binaries were still included if there was no sign of significant shift or unusual behavior of the lines, which would otherwise indicate close binaries. The object SDSS J005506.77-005702.4 is given as either a detached or (with low confidence) semidetached star by Becker et al. (2011). With only three consecutive spectra, the possible shift due to being a binary system in the data is too low to make concrete assumptions; the object is therefore kept in the list with the flaring attribute “probable”.

Object SDSS J141528.99+123127.4 was found to be in front of a galaxy. Manual inspection revealed that the *Gaia* DR2 parallax closest to the stellar coordinates obtained by cross-matching gave the value of the galaxy (parallax around 0 mas). The distance was replaced with the real stellar distance value. Object SDSS J161237.41+164556.0 also appears to be in front of another object, resulting in two *Gaia* DR2 distances. Both distances are similar (225 versus 256 parsec, Bailer-Jones et al. 2018).

The spectrum of object SDSS J162027.54+364002.8 consists of two stars (M 4 and M 3) that lie too close to each other to be resolvable as independent sources because the entire diameter in SDSS plates is larger than their angular distance. The distances given by *Gaia* DR2 are 300 and 350 parsecs, respectively.

K stars. The star SDSS J161924.09+061554.2, marked as subtype K3 in SDSS, had a flare that changed its Balmer absorption lines into emission lines. It was identified to be an RS Canum venaticorum star (Drake et al. 2014).

The star SDSS J011707.01+250359.8, marked as subtype K3 in SDSS, showed a new emission line in $H\alpha$ and $H\beta$ at one time instance. It was identified to be a eclipsing binary of type W UMa (Drake et al. 2014).

The star SDSS J075835.12+482523.6 is marked as a K5 star in the SDSS and shows high chromospheric emission lines with variation over time. The spectra show clear shifts in the data, which is an indication of a close binary object. This assumption is confirmed by Kleinman et al. (2013), who declared it a white dwarf+M0e binary, and by Watson et al. (2006), who recognized the object as a cataclysmic variable.

The star SDSS J215455.17+011414.5, marked as subtype K5 in the SDSS, shows significantly higher emission lines in $H\alpha$, $H\beta$, and $H\gamma$ and especially high Ca H and K emission lines at two time instances. The spectrum as a whole is redshifted by $\sim 2 \text{ \AA}$ in these instances. Even without the flare, the star showed these lines in emission. Xue et al. (2014) determined the star to be a K giant star on the RGB, while Kowalski et al. (2009) in their search for flaring M dwarfs used the spectral classification dM0. When data from *Gaia* are used, the star appears to be on the upper part of the main sequence, while *Gaia* data also give information on the radius ($2.1972 R_{\odot}$) and luminosity ($2.153 L_{\odot}$) (*Gaia* DR2). This clearly suggests that the star is not an M dwarf star. Figure 7 confirms this assumption because it is the only flaring star above the Sun-like star line. We kept the flare in the final flaring list as a possible K giant or sub-giant. The high emission lines might also indicate a T Tauri type, but there is no literature result to further strengthen this claim.

The star SDSS J000500.22-043228.8 is marked as a K5 star in the SDSS, while an unpublished work found in Skiff (2014)

gave it the spectral type M0Ve. [Heinze et al. \(2018\)](#) declared the object a distant binary. The spectral type K5 appears to be more accurate according to the stellar position in Fig. 7. The SDSS spectra show a possible flare evolution on both $H\alpha$ and $H\gamma$. The whole spectrum is shifted over time, which is an indicator for binaries. Flags in the red spectrum range and missing Balmer lines of a higher order resulted in the “Probable” flare categorization of the event.

Appendix C: Complete list of flaring late-type stars

The complete list of flares found by this work is given in Table C.1 (available at CDS). The columns included in this table are listed below.

RAdeg, DEdeg. SDSS J2000.0 right ascension and declination of the object in degrees.

P-MJD-F. SDSS Plate-MJD-Fiber Identifier.

Method. The method that yielded the detection of this object. The following abbreviations are used:

HTM: high-threshold method (see Sect. 3.1.1)

LTM: low-threshold method (see Sect. 3.1.1)

ELM: emission line method (see Sect. 3.1.2)

R: reevaluated flares also found by [Hilton et al. \(2010\)](#) that were originally rejected after visual inspection.

S/N Bin. Categorization of object into specific S/N bin (see Sect. 2.3).

SpType. Spectral type of object based on a literature search (see Sect. 2.2).

SpTypeSource. Source of the spectral type in the previous column. The following abbreviations are used:

S1: [West et al. \(2011\)](#); S2: [West et al. \(2008\)](#); S3: [Kleinman et al. \(2013\)](#); S4: [Cook et al. \(2016\)](#); S5: [McGehee \(2006\)](#); S6: [Guieu et al. \(2006\)](#); S7: [Newton et al. \(2014\)](#); S8: [Briceño et al. \(2019\)](#); S9: [Kraus & Hillenbrand \(2009\)](#); S10: [Luhman et al. \(2017\)](#); S11: [Zhang et al. \(2010\)](#); S12: [Morgan et al. \(2012\)](#); S13: [Kowalski et al. \(2009\)](#); S14: [Zhong et al. \(2015\)](#); S15: [Skiff \(2014\)](#); SDSS: Spectral type from SDSS file if no literature data were available.

SpTypeSDSS. Original spectral type designation from the .fits file header provided by SDSS.

Dist. Reported distance of star in parsec.

DistSource. Source of the distance used in the previous column. If no specific source is given, [Gaia Collaboration \(2018\)](#) was used (see Sect. 2.5). The following abbreviations are used:

D1: [Jones & West \(2016\)](#); D2: [West et al. \(2011\)](#); D3: [Stassun et al. \(2018\)](#); D4: [Theissen et al. \(2017\)](#).

Quiet. Indicates whether the reference spectrum for calculating the flare energy was quiet.

1: Quiet reference

2: Reference spectrum probably quiet

0: Reference spectrum has problems (e.g. cosmic-ray hit), the energy calculation is not reliable.

FlareEvol. Estimates of the evolution visible in the spectra (see Sect. 4.2.3 and the discussion in Sect. 5.6). The following abbreviations are used:

“rise” or “p. rise” ... flaring rise or probably rise

“decay” or “p. decay” ... flaring decay or probably decay

“unclear” ... activity visible, but impossible to declare evolution

“diff. activity” ... the flare occurs at a different time, no evolution visible

“full” ... rise and decay as well as a peak are visible.

Consider. Indicates how reliable the flare is using either “Yes” or “Probable” (see Sect. 3.1.3 and the discussion in Sect. 5.6).

Hilton2010. Indicates whether the flare was also found by [Hilton et al. \(2010\)](#).

Platequal. Read out from the SDSS .fits file header. Plate quality can either be “good”, “marginal” or “bad”.

HaFlareEnerg. Calculated $H\alpha$ flare energy in units of [Ws] (see Sect. 4.2.2).

HaFlareEnergErr. Calculated $H\alpha$ flare energy error in units of [Ws] (see Sect. 4.2.2).

HaFlareLum. Calculated peak $H\alpha$ flare luminosity in units of [W] (see Sect. 4.2.2).

HaFlareLumErr. Calculated peak $H\alpha$ flare luminosity error in units of [W] (see Sect. 4.2.2).

Nexp. Number of available exposures (single spectra) in the SDSS spectrum that include the $H\alpha$ line.

NflareSpec. Number of single spectra in the SDSS spectrum that are in flaring state.

Refspec. Number of the spectrum that was used as a reference for the calculation. The number refers to the chronological order: 1 refers to the first spectrum that was observed. A zero indicates that the coadded spectrum was used as reference.

Peakspec. Number of the spectrum that shows the peak of the flare. The number refers to the chronological order: 1 refers to the first spectrum that was observed.

Flaretimeall. Summed time of all flaring single spectra in the SDSS spectrum in minutes.

Comments. Subjective comments added after confirming the flares. Frequently used abbreviations and notes:

(Subjective estimated) significance and size of the flare using the designations “large”, “medium”, “small”, and “mini”.

cons: indicates whether the flare is in a consecutive series of spectra.

Notes on wing variability: for example, “ha hb hg hd both wings” means enhanced wings in $H\alpha$, $H\beta$, $H\gamma$, and $H\delta$. When a Balmer line wing is mentioned, it indicates the possibility of an enhancement. Possible absorptions are mentioned separately. “CME?” indicates that this event was analyzed in depth to conclude whether a CME did occur.

“New emission line” or “new em. line” indicates whether the flare produced a Balmer emission line where no emission line was previously visible (good indicator for flaring activity).

Other notes include: visible shifts in the data, high noise values, “cont. em?” as the possibility of enhanced continuum emission due to the flare, and many available spectra (as “many data points”).