RAA 2021 Vol. **21** No. **11**, 288(15pp) doi: 10.1088/1674-4527/21/11/288

http://www.raa-journal.org *http://iopscience.iop.org/raa*

Research in Astronomy and Astrophysics

Discovery of 2716 hot emission-line stars from LAMOST DR5

Baskaran Shridharan[1], Blesson Mathew[1], Sabu Nidhi[1], Ravikumar Anusha[1], Roy Arun[1], Sreeja S. Kartha[1] and Yerra Bharat Kumar[2]

[1] Department of Physics and Electronics, CHRIST (Deemed to be University), Bangalore 560029, India; *shridharan.b@res.christuniversity.in*

[2] Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China

Received 2021 March 10; accepted 2021 August 17

Abstract We present a catalog of 3339 hot emission-line stars (ELSs) identified from 451 695 O, B and A type spectra, provided by LAMOST Data Release 5 (DR5). We developed an automated Python routine that identified 5437 spectra having a peak between 6561 and 6568 Å. False detections and bad spectra were removed, leaving 4138 good emission-line spectra of 3339 unique ELSs. We re-estimated the spectral types of 3307 spectra as the LAMOST Stellar Parameter Pipeline (LASP) did not provide accurate spectral types for these emission-line spectra. As Herbig Ae/Be stars exhibit higher excess in near-infrared and mid-infrared wavelengths than classical Ae/Be stars, we relied on 2MASS and *WISE* photometry to distinguish them. Finally, we report 1089 classical Be, 233 classical Ae and 56 Herbig Ae/Be stars identified from LAMOST DR5. In addition, 928 B[em]/A[em] stars and 240 CAe/CBe potential candidates are identified. From our sample of 3339 hot ELSs, 2716 ELSs identified in this work do not have any record in the SIMBAD database and they can be considered as new detections. Identification of such a large homogeneous set of emission-line spectra will help the community study the emission phenomenon in detail without worrying about the inherent biases when compiling from various sources.

Key words: stars: early-type — methods: data analysis — techniques: photometric — astronomical databases: catalogs

1 INTRODUCTION

Stars with bright emission lines in the spectrum indicating unusual activities in the stellar atmosphere have become a subject of much astrophysical research. Objects with such diverse spectra showing Balmer emission along with metallic emission lines are generally known as emission-line stars (ELSs). These lines arising from the circumstellar disk/envelope of the star are formed due to the recombination process. Based on the evolutionary stage, spectral type and mass, ELSs can be classified into various categories such as young stellar objects (YSOs), Oe/Be/Ae stars, Wolf-Rayet stars, supergiants, planetary nebulae, luminous blue variables (LBVs), Mira stars, flare stars and symbiotic stars (Kogure & Leung 2007). Hence, the study of ELSs can provide insight towards the line formation regions and various evolutionary phases of stars.

In recent years, research on ELSs (Vioque et al. 2020; Akras et al. 2019; Huang et al. 2013) has surged due to the availability of photometric surveys (*Gaia*, Two Micron All-Sky Survey (2MASS), IPHAS, etc.) and large scale spectroscopic surveys (LAMOST, SDSS, APOGEE). Merrill & Burwell (1949) were the first to undertake a survey on ELSs in which spectra of early-type stars with bright hydrogen lines were detected. Wackerling (1970) compiled a catalog of 5326 early-type ELSs in our Galaxy. Later, Allen (1973) observed nearly 250 early-type ELSs in near-infrared (NIR) wavelengths and proposed a mechanism wherein the excess flux in NIR (infrared (IR) excess) arises due to thermal radiation from the dust in the circumstellar disk and the electron *bremsstrahlung* originating from a shell of ionized gas. Following these works, Stephenson & Sanduleak (1977) identified 455 new bright Hα stars in our Galaxy using objective-prism plates. By the end of the 20th century, various surveys increased the count of ELSs considerably. One of the most extensive ELS surveys in the Hα regime was performed by Kohoutek & Wehmeyer (1999), in which they provided a catalog of 4174 Hα ELSs in the Northern Milky Way within the latitude range of –10° $< b <$ 10°. Following Kohoutek & Wehmeyer (1999), a much larger and deeper survey was performed and is now known as The INT Photometric Hα Survey of the Northern Galactic Plane (IPHAS; Drew et al. 2005). It covered the Milky Way in Hα and Sloan r' and i' bands within a magnitude range of $13 \leq r' \leq 19.5$ mag. Following the initial data release from IPHAS (González-Solares et al. 2008), a series of studies on different types of ELSs was published. Witham et al. (2008) presented the preliminary catalog of 4853 Hα emission sources from IPHAS. Various ELS classes including cataclysmic variables (Witham et al. 2007), symbiotic stars (Corradi et al. 2008, 2010; Rodríguez-Flores et al. 2014), YSOs (Vink et al. 2008; Barentsen et al. 2011) and classical Be stars (Raddi et al. 2013) were identified from the survey. Similarly, Kalari et al. (2015) studied the accretion rates of 235 classical T Tauri star candidates in the Lagoon Nebula utilizing the VST Photometric Hα Survey of the Southern Galactic Plane and Bulge (VPHAS+; Drew et al. 2014). There are quite a few studies which searched for ELSs in open clusters. For example, Mathew et al. (2008) identified 157 ELSs from the slitless spectroscopic survey of 207 open clusters in the Galaxy. Reipurth et al. (2004) and Pettersson et al. (2014) presented Hα ELS surveys in various molecular clouds relying on wide field-objective prism films. Recently, Vioque et al. (2020) identified new ELS candidates using machine learning techniques on data from *Gaia* Data Release 2 (DR2), 2MASS, *Wide-field Infrared Survey Explorer* (*WISE*) and IPHAS or VPHAS+ surveys.

The Large Sky Area Multi-Object fiber Spectroscopic Telescope (LAMOST) surveys have accumulated millions of spectra over a period of 5–6 years. Even though a huge spectral database from LAMOST has been made available to the public, very few works have identified and characterized ELS spectra to date. Lin et al. (2015) identified 192 classical Be (CBe) star candidates from LAMOST Data Release 1 (DR1), which increased the then known sample of CBe stars by about 8%. Similarly, a study on ELSs and Hα line profiles from LAMOST DR2 was conducted by Hou et al. (2016), in which 10 436 early-type ELSs were identified. Yao et al. (2017) studied Mira variable stars identified from LAMOST Data Release 4 (DR4) and Li et al. (2018) identified six new Oe stars from LAMOST Data Release 5 (DR5).

In this work, we exploit the extensive data available to us from LAMOST DR5 to identify early-type ELSs and mainly classify them into CBe stars and Herbig Ae/Be (HAeBe) stars. A CBe star is a fast rotating B-type non-supergiant with an equatorial decretion disk (Porter & Rivinius 2003). The disk is transient in nature as identified through the change in Hα emission strength over a timescale of years to decades (Štefl et al. 2003). The formation of a decretion disk in CBe stars is known as the 'Be phenomenon'. The episodic occurrence of mass loss which results in the formation of a decretion disk is still an open question. The emission strengths of Hα and other hydrogen and metallic emission lines were analyzed to understand the 'Be phenomenon' in CBe stars (Rivinius et al. 2013, for a review). Most CBe stars span the spectral range of B0 – B9, even though a few late O (Li et al. 2018) and early A type stars (Jaschek et al. 1986; Anusha et al. 2021) are included in this class. Also, CBe stars are defined as belonging to the luminosity class III to V (Jaschek & Egret 1982). In contrast to CBe stars, pre-main sequence (PMS) stars are very young objects that have not started hydrogen burning. The low-mass (0.1 to 2 $M_{\odot}$) PMS stars are known as T Tauri stars (Joy 1945), whereas those in the intermediate-mass range (2 to 8 $M_{\odot}$) are called HAeBe stars (Herbig 1960; Waters & Waelkens 1998). HAeBe and T Tauri stars share common characteristics such as a circumstellar accretion disk, IR excess, Balmer emission lines and metallic emission lines (Hillenbrand et al. 1992). For the present study, we discuss the ELSs belonging to O, B and A spectral types, whereas those belonging to later spectral types will be discussed in a follow-up study (Edwin et al. in prep).

We downloaded spectra of O, B and A type stars from LAMOST DR5 and developed a Python routine to identify spectra with Hα in emission. To re-estimate the spectral types of ELS spectra, we performed a visually aided semi-automated template matching process utilizing the MILES spectral library. Further, we made use of available photometric data from missions/surveys such as 2MASS, *WISE*, *Gaia* Early Data Release 3 (EDR3) and APASS to sub-classify our sample into CBe stars and HAeBe stars. We categorized the identified ELS spectra into various classes, as explained in Subsection 3.5.

The structure of the paper is as follows. Section 2 provides a brief introduction of the LAMOST survey and explains the data collection along with the spectral normalization procedure. In Section 3, we discuss the ELS identification and classification procedure along with the explanation of spectral-type re-estimation technique. We have also detailed the naming convention employed in this work. In Section 4, we plot IR color-color diagram (CCDm) and *Gaia* color-magnitude diagram (CMD), followed by discussion of spectral features seen in different classes identified in this work. We cross-matched our list with the SIMBAD database to verify our classifications. The major results of our work are summarized in Section 5. Description of our catalog is given in Appendix A, automated forbidden line identification criteria is provided in Appendix B and necessary information for a representative sample of ELSs identified is listed in Appendix C.

2 DATA COLLECTION

2.1 About LAMOST Data Release 5

LAMOST is a reflecting Schmidt telescope, with an effective aperture of 3.6 m – 4.9 m and a wide field of view (FoV) of 5° (Cui et al. 2012; Wang et al. 1996). It is also known as the Guo Shoujing Telescope, and is operated and maintained by Xinglong station of the National Astronomical Observatories, Chinese Academy of Sciences. It is equipped with 4000 optical fibers in its focal plane with 16 low-resolution spectrographs and 32 charge-coupled devices (CCDs) (Zhao et al. 2012). It is designed with three major components: the correcting mirror *Ma*, the primary mirror *Mb* and a focal surface (Cui et al. 2012). Objects with r-band magnitude up to ∼20 mag are observed by LAMOST in the wavelength range of 3690 – 9000 Å at spectral resolution of R∼1800 (Cui et al. 2012). The Galactic and extragalactic surveys conducted by LAMOST are classified into two major projects: LAMOST Experiment for Galactic Understanding and Exploration (LEGUE; Deng et al. 2012) and LAMOST Extra GAlactic Survey (LEGAS; Zhao et al. 2012).

2.2 Data Collection and Spectral Normalization

During the five-year regular survey, the raw data were reduced by the LAMOST two-dimensional (2D) pipeline (Luo et al. 2015), which implemented procedures similar to those of SDSS spectro2d pipeline (Stoughton et al. 2002). A spectrum of a light beam from a fiber is imaged onto the 32 cameras using 4k×4k CCD chips (Wei & Stover 1996). The data from each CCD chip are separately reduced and then combined with other exposures to improve signal. Bias and dark frames are subtracted from each raw image. Flat-fields are then traced and extracted for each fiber. Wavelength is calibrated using the arc lamp spectra and slightly adjusted to match the known positions of certain sky lines, which are then corrected to the heliocentric frame. Further, the LAMOST one-dimensional (1D) pipeline extracts and classifies spectra into four categories (*star*, *galaxy*, *QSO* and *unknown*) where radial velocities, redshifts and other stellar parameters are calculated wherever possible (Zhao et al. 2012; Cui et al. 2012).

From 8 183 160 spectra classified as '*star*' in LAMOST DR5, we selected 451 695 spectra belonging to O, B and A spectral type as given by the LAMOST pipeline. LAMOST spectra with low signal-to-noise ratio (SNR) may affect the selection and analysis of Hα emission lines. In a study identifying Hα emission spectra from LAMOST DR2 by Hou et al. (2016), spectra with SNR of SDSS r band less than 10 (SNR$_r$ $<$ 10) were removed. Further, the automated routine to find emission lines in LAMOST spectra does not yield good accuracy if we reduce SNR$_r$ to less than 10 (SNR$_r$ $<$ 10). We observed that many noise peaks were reported as emission peaks. Hence, we adopted the criterion of SNR$_r$ $>$ 10 in order to remove noisy spectra from our sample.

The reduced and calibrated spectra were downloaded from the LAMOST DR5 website[1]. The data array of a LAMOST fits file is as follows. The first frame contains the flux information, and the second frame contains the 'inverse variance' ($1/\sigma^2$, where σ is the uncertainty). The third frame stores wavelength in Å whereas the fourth and fifth frames contain 'andmask' and 'ormask' data respectively. Further information about the fits description and data structure is accessible online[2]. To normalize the spectra, we employed `laspec`, a LAMOST spectral kit developed by Bo Zhang (Zhang et al. 2020). This Python-based package is open-source and available in GitHub[3]. The entire spectrum is first smoothed applying a smoothing spline and the pixels lying away from the 1.5σ threshold are removed. The remaining pixels are smoothed to obtain the pseudo-continuum, which is then used to obtain a normalized spectrum. Refer to section 2.1 of Zhang et al. (2020) for a detailed explanation regarding pre-processing of LAMOST spectra.

3 ANALYSIS

3.1 Identification of Emission-Line Stars from LAMOST DR5

In this section, we explain the method to identify the ELSs with Hα in emission in the Galaxy from the large data set of LAMOST DR5. For the purpose of identifying ELSs with Hα in emission, we developed a Python code based on `find_peaks` module in the `scipy.signal` package (Virtanen et al. 2020). The important parameter in `scipy.signal.find_peaks` is "Width", which corresponds to the full width at half maximum (FWHM) of the emission profile. Spikes in the spectrum caused by instrumental defects will be very narrow and can be mistaken for a narrow emission line. To remove such lines, we applied a width cutoff of three sampling points. This reduces the possibility of false emission (caused by instrumental defects or noise) near Hα being reported as

[1] dr5.lamost.org
[2] http://dr5.lamost.org/v3/doc/data-production-description
[3] https://github.com/hypergravity/laspec

real emission in the spectrum. We consider Hα emission to be true only if FWHM is at least three sampling points wide. Even though it affects the completeness of ELSs from DR5, the probability of reporting false emission is reduced compared to selection criteria used in Hou et al. (2016). The scope of this work is to provide a comprehensive catalog of stars with Hα in emission which justifies the use of our stringent selection criteria. Out of 451 695 O, B and A spectra from LAMOST DR5, we identified 5437 spectra with Hα in emission. Through a visual check, those without any photospheric features and broken spectra were removed, resulting in 4138 good quality spectra with Hα in emission from LAMOST DR5. The selection scheme and the classification procedure we followed to identify 4138 emission-line spectra from LAMOST DR5 are presented in the form of a flowchart in Figure 2. It should be noted that if the FWHM criterion is relaxed to a lower value, say to 1 sampling point, the number of ELSs found would increase drastically but they may consist of many false detections.

From Figure 1, it can be seen that most of the ELSs are in the direction of the Galactic anti-center but a few objects are within $60° \leq l \leq 90°$. Yuan et al. (2015) reported that the LAMOST Spectroscopic Survey of the Galactic Anti-center (LSS-GAC) survey was initially aimed at observing the Galactic thin/thick disks and halo for a contiguous sky area which is centered on the Galactic anti-center direction ($150° \leq l \leq 210°$). Also, for large distances, the inverse relation between *Gaia* EDR3 (Lindegren et al. 2021) parallax and distance fails. This was resolved by Bailer-Jones et al. (2021) following a probabilistic approach to determine the distance from *Gaia* EDR3 parallax. The distance range observed here affirms that the sample of ELSs observed is spread over 10 kpc with about 90% of the objects within 5.5 kpc; 53 objects are present in the high Galactic latitude region ($|b| \geq 40°$) which may be of further interest to readers, but is out of the scope of this work.

3.2 Spectral Type Re-estimation

The spectral type of each LAMOST DR5 spectrum is estimated using the LAMOST Stellar Parameter Pipeline (LASP; Wu et al. 2011). The raw CCD images are processed through the 2D reduction pipeline (Cui et al. 2012) and the output is in the form of a 1D spectrum for each star. The spectral type of the star is estimated through template matching with standard spectral templates and line recognition algorithms (Luo et al. 2015). A sample of 183 stellar spectra from the LAMOST pilot survey and LAMOST DR1 were selected as standard spectral templates (Wei et al. 2014; Kong et al. 2019). For classifying O and B type stars, the LAMOST pipeline has used only O, B, B6 and B9 stellar templates. The presence of the veiled continuum and emission lines that formed outside the photosphere complicates the spectral classification in early-type stars (Hernández et al. 2004). For this reason, automated spectral type estimation for ELSs may give erratic results. This was validated by our finding that, despite eliminating bad spectra from our sample, spectra classified as A-type by LASP feature HeI absorption lines. Hence we need to re-estimate the spectral types for all the spectra. Thus, one has to resort back to the semi-automated template matching technique where the user needs to select the best fit considering various spectral lines. In order to distinguish O, B and A type spectra, we can rely on the profile of Balmer absorption lines. Since we are dealing with emission-line spectra, we cannot use Hα, Hβ or even Hγ to some extent because the emission mechanism can fill in these absorption lines. Hence as a first check, we matched the Hδ and Hϵ profiles to estimate the spectral type coarsely. To estimate the spectral sub-class, we relied on absorption lines such as HeII, HeI and Mg II. Theoretically, one expects ionized HeII lines in O-type stars. The HeII lines fade as we move towards B-type stars where HeI absorption lines are stronger and further HeI lines slowly disappear in early A-type stars (Gray & Corbally 2009). Also, since CBe stars are fast rotators (200–300 km s^{-1}), the depth or profile shape varies depending on the $v\sin i$ parameter. This will in turn affect the spectral type re-estimation process if we focus on the depth of lines (Slettebak et al. 1980). Keeping this in mind, we matched the wing profile of Hϵ, Hδ, HeI and MgII lines rather than the depth of the line. However, estimating accurate spectral type for CBe stars considering low-resolution spectra is difficult and one can expect up to an error of two sub-classes in B-type stars. To see the comparison between LAMOST candidate spectra and MILES template spectra, please refer to figure 2 of Anusha et al. (2021).

As explained above, we performed spectral type re-estimation implementing the semi-automated template matching method. For this, we used the MILES stellar spectral library (Sánchez-Blázquez et al. 2006) since it has similar resolution (R∼2000) to LAMOST DR5 spectra (R∼1800). The homogeneously calibrated 985 spectra in the MILES spectral library were obtained from the 2.5 m Isaac Newton Telescope (INT), covering a range of 3525–7500 Å (Falcón-Barroso et al. 2011). The template stars span a wide parametric space, i.e. $T_{\rm eff} = 3000 - 40\,000$ K and $\log(g) = 0.2 - 5.5$. Out of 985 MILES templates, 79 of them are in the spectral range B0 – A9 covering luminosity classes III – V. Rest-frame corrections to LAMOST DR5 spectra were made applying the redshift

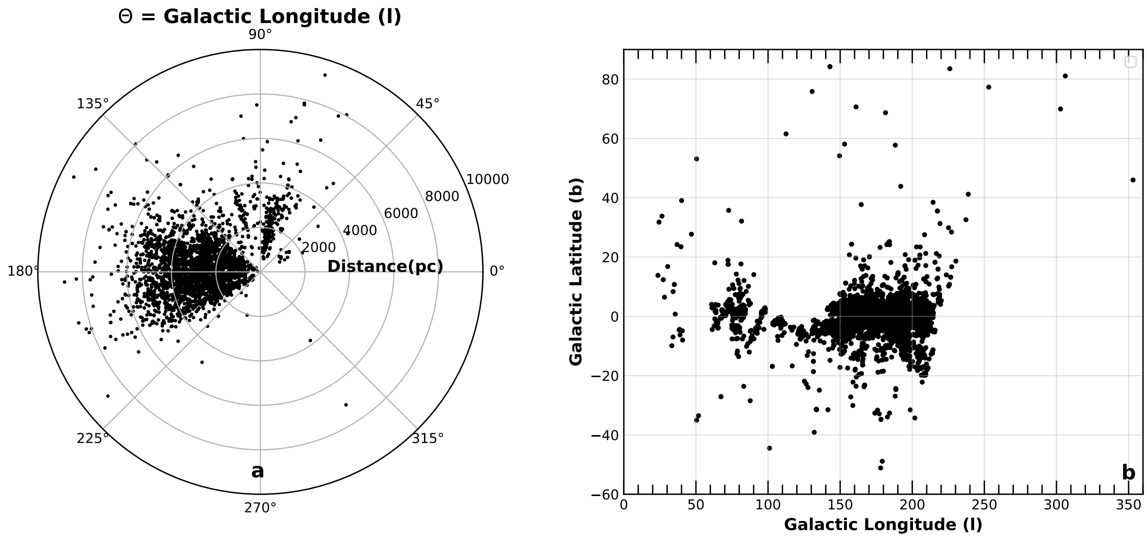

Fig. 1 The polar and spatial distributions of 3339 ELSs in the Galaxy identified from LAMOST DR5 are displayed. Fig. 1(a) (*left*) features the polar distribution of ELSs in the Galactic longitude (l) vs. distance (d) plane. Distances are taken from Bailer-Jones et al. (2021). The ELSs are concentrated in the Galactic anti-center direction due to the observational strategy adopted by LAMOST (Hou et al. 2016). Fig. 1(b) (*right*) represents the spatial distribution of ELSs in the Galactic latitude (b) vs. Galactic longitude (l) plane. It can be seen that most of the ELSs are found to be within $150° \leq l \leq 200°$ and $|b| \leq 20°$.

values provided by LASP. We developed a Python code, which for a given spectrum identifies three best-fitting templates (based on the chi-squared value with a lower chi-squared value yielding a better fit) after iterating over the entire MILES library. To ensure the HeI and MgII absorption lines fit well, we over-plotted each spectrum with the top three template matches and visually identified the best among them. In most cases, the template with the least chi-squared value fitted the spectra the best, but in a few cases, the template with the second or third least chi-squared value fitted the higher-order Balmer lines (Hδ and Hϵ), HeI 4471 Å and MgII 4481 Å lines better, which justifies the time-intensive spectral typing procedure. Out of 4138 ELS spectra, we were able to ascertain spectral types to 3307 spectra. At this stage, spectra showing nebular forbidden lines such as [NII], [SII] and [OI] were identified using an automated Python routine, which is explained in Appendix B. These spectra are classified as either A[em] or B[em][4] based on re-estimated spectral types. We could not assign a spectral type to 831 spectra because the blue ends of the spectra were noisy and they are classified as “Em” or “Em[em]” (if forbidden lines are present).

3.3 Evolved Star Candidates

The evolved and highly luminous stars (post-main sequence/supergiants) with early spectral type are reported to have Hα in emission (Rosendhal 1973). The Hα emission in evolved stars primarily originates from wind driven mass loss (Weymann 1963; Hutchings 1970). The individual reddening for each spectral coordinate was obtained utilizing the `dustmaps` open source package (Green 2018), which provides probabilistic reddening measurements based on *Gaia* parallax and stellar photometry from 2MASS and PanSTARRS1. The reddening $E(g-r)$ obtained using `dustmaps` is converted to $E(B-V)$ applying a general correction of 0.884[5]. Relying on these reddening measurements, we calculated the extinction in the V band by applying the relation $A_V = 3.1 \times E(B-V)$. The extinction in photometric bands of *Gaia* and 2MASS is then determined by adopting the conversion relations from Wang & Chen (2019).

To identify evolved stars present in our sample, two methods, spectroscopic (first) and photometric (second), are applied. First, 42 spectra with luminosity class Ia/b & II estimated using the MILES library are classified as “Evolved*”[6]. In addition, there can be some evolved stars whose luminosity class is misclassified. Hence as

[4] It should be noted that [em] signifies spectra with forbidden lines instead of [e] because [e] stars are conventionally defined to have [FeII] lines (Swings 1973)

[5] *argonaut.skymaps.info/usage*

[6] ‘*’ here denotes that they are evolved star candidates, a convention we have followed throughout this work.

a second criterion, we set the bolometric luminosity ($L_{\rm bol}$) limit adopted from Lamers et al. (1998). The criterion of $\log(L_{\rm bol}/L_\odot) > 4.5$ is applied to all the remaining stars. We used *Gaia BP* (G_{BP}) magnitude and performed necessary bolometric corrections to calculate $\log(L_{\rm bol}/L_\odot)$. The equation for bolometric magnitude utilizing G_{BP} is given below,

$$M_{bol} = G_{BP} - 5\log(d) + 5 + BC - A_{G_{BP}}\,, \quad (1)$$

where BC is the bolometric correction taken from Jordi et al. (2010) for the re-estimated spectral types. $A_{G_{BP}}$ is calculated using the conversion coefficient taken from Wang & Chen (2019), which is stated as $A_{G_{BP}} = 1.002 \times A_V$. The $(L_{\rm bol}/L_\odot)$ is calculated with the following equation,

$$\frac{L_{\rm bol}}{L_\odot} = \frac{L_o}{L_\odot} \times 10^{-0.4M_{\rm bol}}\,, \quad (2)$$

where L_o is the zero-point luminosity equaling 3.0128×10^{28} W. Using the above equation, we classified four additional spectra as "Evolved*" as they satisfy the photometric criteria. In total, 46 spectra are classified as 'Evolved*', of which 26 spectra belong to B2 or earlier spectral type.

3.4 Estimation of IR Excess in ELS

The spectral energy distribution (SED) of main sequence and PMS ELSs shows excess flux over the blackbody distribution in the IR. This is indicative of the presence of a circumstellar disk in ELS (Natta et al. 1993). IR continuum excess observed in CBe stars is due to thermal *bremsstrahlung* emission from the circumstellar disk (Gehrz et al. 1974). For a CBeAe, the excess is predominantly concentrated in the NIR region. For HAeBe stars, the IR excess is attributed due to thermal emission from dust in the accretion disk (Hillenbrand et al. 1992; Malfait et al. 1998).

In order to differentiate main sequence ELS (CBeAe) and PMS stars (HAeBe), we relied on two proxies of IR excess. Firstly, we chose all the objects with 2MASS (Cutri et al. 2003) magnitudes satisfying the criterion that reddening corrected $(H - K_S)_0$ color is greater than 0.4, which is a well-established cut-off to select HAeBe stars (Finkenzeller & Mundt 1984). Secondly, we calculated two spectral indices (Lada 1987; Wilking 1989; Greene et al. 1994) from the SEDs as proxies to quantify NIR and mid-infrared (MIR) excess. NIR spectral index (n_{J-K_S}) is calculated using 2MASS J and K_S magnitudes and MIR spectral index (n_{K_S-W2}), with 2MASS K_S and *WISE* (Cutri et al. 2014) $W2$ magnitudes. The equation defining the spectral index is calculated using the definition shown as Equation (3). λ_1 and λ_2 represent the wavelengths of two bands for which spectral index is calculated. F_{λ_1} and F_{λ_2} correspond to the extinction corrected absolute flux measured at λ_1 and λ_2.

$$n_{\lambda_1-\lambda_2} = \frac{\log(\frac{\lambda_2 F_{\lambda_2}}{\lambda_1 F_{\lambda_1}})}{\log(\frac{\lambda_2}{\lambda_1})}\,. \quad (3)$$

Based on the availability of IR photometric values for 3339 unique objects, n_{J-K_S} was calculated for 2978 stars and n_{K_S-W2} index was calculated for 2743 stars. Figure 3 displays the classification of ELSs based on n_{J-K_S} and n_{K_S-W2}. The remaining 597 stars did not have 2MASS or *WISE* magnitudes to calculate indices. We considered objects with $(n_{K_S-W2}) > -1.5$ to have MIR excess (Andre et al. 1993; Arun et al. 2019). In addition, we also segregated objects having $(n_{J-K_S}) > -1.5$ but $(n_{K_S-W2}) < -1.5$ into a separate class named "Ae/Be*-onlynir". These objects have excess in NIR but do not show excess in MIR wavelengths, which needs to be studied in detail.

3.5 Naming and Classification Scheme

Even though we treated 4138 spectra as independent entities, many stars were observed multiple times during the 5-year observation cycle. Hence it is necessary to identify and name each spectrum based on its observed coordinate. We sorted 4138 spectra based on their $snrr$ value in descending order and assigned the name of first (highest SNR) spectrum as "LEMC 0001". LEMC is abbreviated as "LAMOST Emission-Line Catalog (LEMC)". If another spectrum has RA and DEC within 1″ of "LEMC 0001", then the new spectrum is named "LEMC 0001_2" and the next observation of the same object (if present) is named "LEMC 0001_3", and so on. The same process is iterated over the entire list and spectra are named from LEMC 0001 to LEMC 3339.

Coming to the classification of 4138 spectra, we created various classes to accommodate the non-availability of photometry values in some cases. If a spectrum classified as B-type does not have *Gaia* EDR3 or 2MASS photometry within a search radius of 3″, then it is classified as "Be**". Further if the same spectrum has *Gaia* EDR3 and 2MASS photometry but was not detected in *WISE*, we classified it as "Be*" since Lada spectral indices cannot be calculated. We classify a spectrum as "CBe" only if it satisfies all three conditions, i.e., Hα in emission, $(H - K_S)_0 \leq 0.4$ and the criteria set by Lada indices. This makes our classification robust.

We noted that in a few cases the multi-epoch observations have conflicting classifications. For example, LEMC 265 has three observations until DR5 which are classified as HBe, HAe and Em in each case. This is

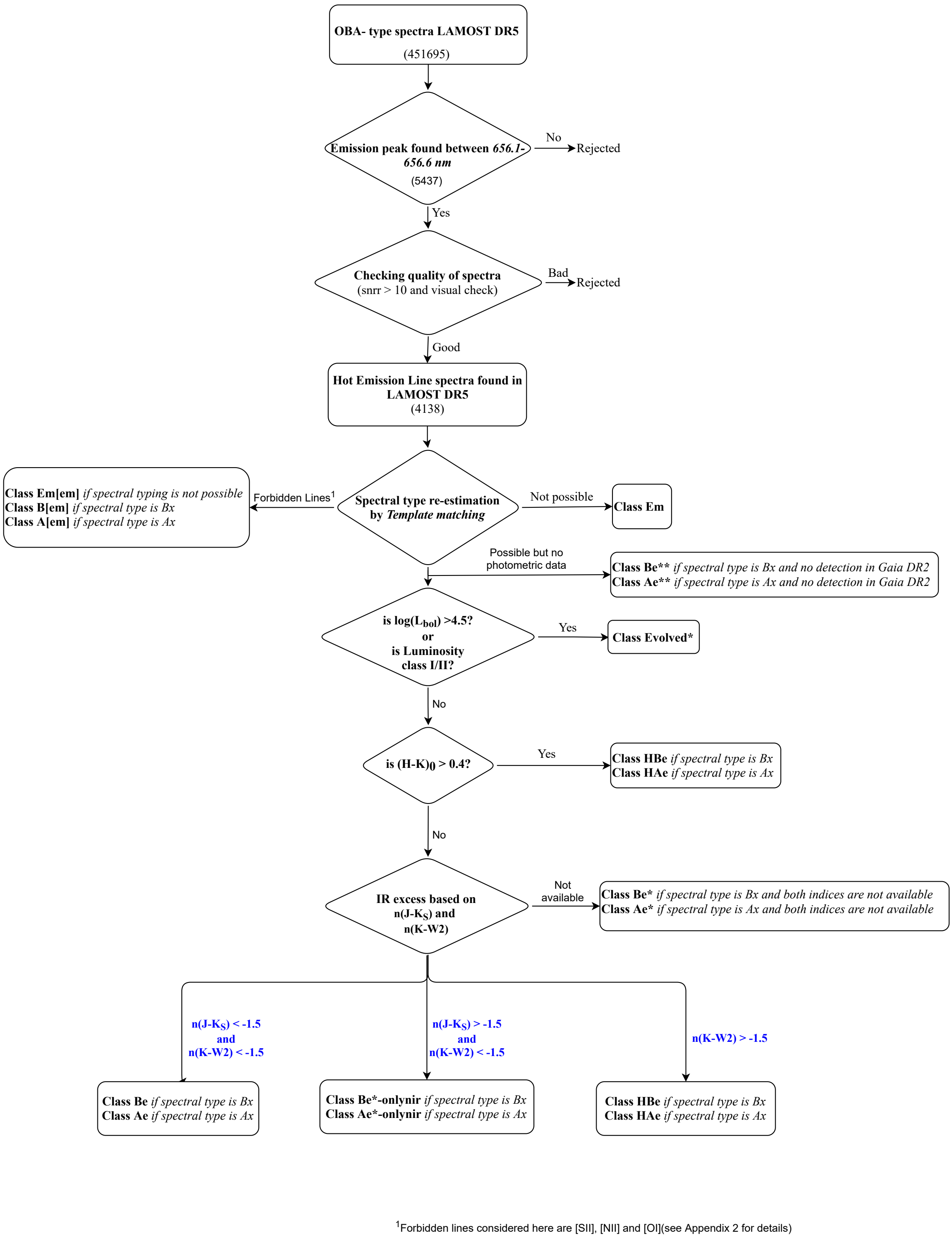

Fig. 2 The classification procedure followed in this work is explained as a flowchart here. Various classes that are used in each step are explained in detail.

because spectra classified as "Em" are noisy in the blue-end which make the spectral type estimation not possible. The spectral types of two other observations are estimated here to be B9 and A1V, which led to final classifications of HBe and HAe respectively. In another case, seven multi-epoch observations of LEMC 0013 are classified into "CBe" (four spectra) and "Evolved*" (three spectra). Since the spectrum of LEMC 0013 with highest $snrr$ values (LEMC $0013_{-}1$) is classified as "CBe", the star LEMC 0013 is put into the list of 1089 unique CBe stars. Hence, care should be taken to select the best spectrum out of multiple observations.

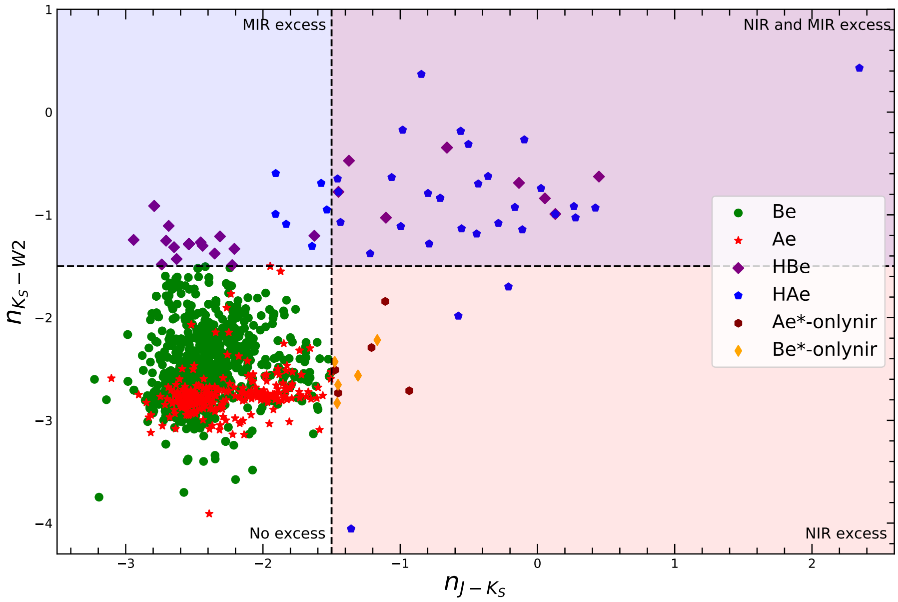

Fig. 3 The plot illustrates classification of different ELSs into categories based on their NIR and MIR Lada indices (Lada 1987). The *black dashed lines* represent $(n_{J-K_S}) = -1.5$ and $(n_{K_S-W2}) = -1.5$. Classes such as Be (*green filled circles*), Ae (*red filled stars*), HBe (*purple filled diamonds*), HAe (*blue filled pentagons*), Ae*-onlynir (*brown filled hexagons*) and Be*-onlynir (*yellow solid thin diamonds*) are marked in this figure.

4 RESULTS

We classified 4138 spectra of 3339 unique ELSs identified from LAMOST DR5 into various classes based on available optical/IR photometry and presence of forbidden emission lines. A detailed classification algorithm is illustrated in the form of a flowchart in Figure 2.

4.1 2MASS-*WISE* Color-Color Diagram

Presence of dust in the circumstellar environment/accretion disk of an ELS can be identified from the excess flux emission in the IR wavelengths (Hartmann et al. 2005). We plotted the $(H-K_S)_0$ versus $(J-H)_0$ for CBe, CAe and HAeBe stars in Figure 4. As observed in the 2MASS CCDm (see Fig. 4), CBe stars are more clustered together than CAe stars. Also, note that most of the HAeBe stars have $(H-K_S)_0 > 0.4$. It is seen that HAe stars more strictly obey the $(H-K_S)_0 > 0.4$ cutoff than HBe stars as suggested in Finkenzeller & Mundt (1984). IR excess in HBe stars is visibly more evident in the (n_{K_S-W2}) index than $(H-K_S)_0$ value.

Also, we have constructed a 2MASS-*WISE* CCDm where $(W1-W2)_0$ is plotted against $(H-K_S)_0$ in Figure 4. Almost all HAeBe stars occupy a distinct region compared to CBeAe stars, which justifies our classification technique.

4.2 *Gaia* Color-Magnitude Diagram

Gaia EDR3 provides G, G_{BP} and G_{RP} magnitudes for 3017 ELSs present in our sample. The absolute G magnitude (M_G) is calculated using the distance values taken from Bailer-Jones et al. (2021). Further, the extinction corrected M_G and $(G_{BP}-G_{RP})_0$ colors are plotted in Figure 5 for CBe, CAe and HAeBe stars. The zero-age main sequence (ZAMS) defined by Pecaut & Mamajek (2013) does not extend beyond B9 spectral type for M_G and $(G_{BP}-G_{RP})_0$ values. Hence we plotted the 60 Myr isochrone from the Modules for Experiments in Stellar Astrophysics (MESA) Isochrones and Stellar Tracks (MIST[7]; Choi et al. 2016; Dotter 2016) in the *Gaia* CMD (Fig. 5), which closely matches the ZAMS defined by Pecaut & Mamajek (2013). The MIST archive has provided an isochrone corresponding to $(V/V_{crit} = 0.4)$ which is the only model available in the database for a rotating system. We adopted the metallicity $[\mathrm{Fe/H}] = 0$ for the respective isochrone. It is clear that the CAe and CBe stars occupy overlapping regions whereas HAe and HBe are separated well in the *Gaia* CMD.

4.3 Spectral Features of Various Classes

In this subsection, we discuss the spectral features present in the stars in various evolutionary phases, as identified in

[7] http://waps.cfa.harvard.edu/MIST/

Table 1 Statistics of Classified Spectra

Classification	Number of unique objects	Total number of spectra available	Classification	Number of unique objects	Total number of spectra available
CBe	1089	1523	CAe	233	286
HAe	33	41	HBe	23	46
A[em]	605	637	B[em]	323	391
Em	196	307	Em[em]	490	524
Ae*	17	19	Be*	38	42
Ae**	17	18	Be**	168	182
Be*-onlynir	5	7	Ae*-onlynir	4	4
Evolved*	36	46	F[em]?	58	61
HFe?	1	1	Fe?*	3	3

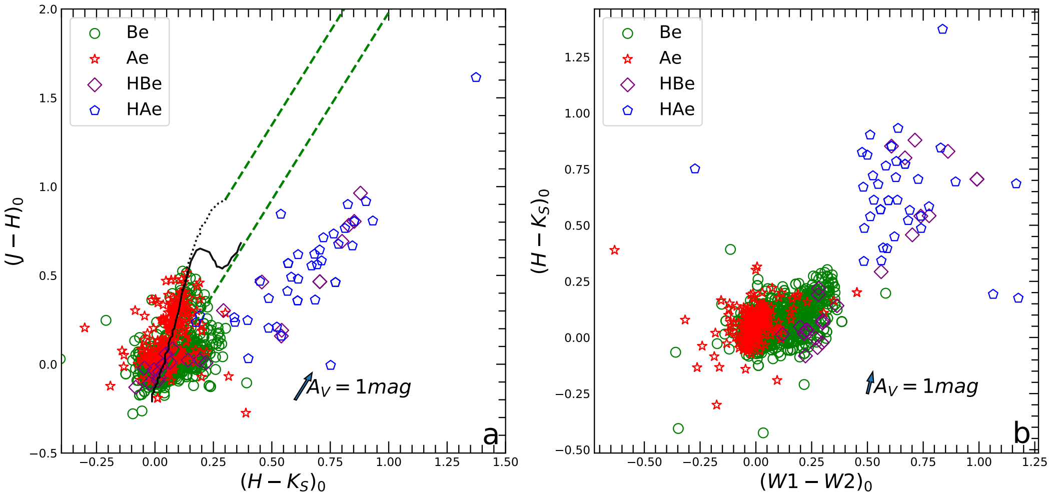

Fig. 4 (a). Reddening corrected 2MASS CCDm marking positions of CBe (*green open circles*), CAe (*red open stars*), HBe (*crimson red open squares*) and HAe (*blue open pentagons*) stars. (b) Reddening corrected 2MASS-*WISE* CCDm representing positions of CBe (*green open circles*), CAe (*red open stars*), HBe (*crimson red open squares*) and HAe (*blue open pentagons*) stars. The separation of main-sequence stars (CBe and CAe) from PMS stars (HBe and HAe) is very evident in the above plots. An *arrow* in both (a) and (b) signifies the direction in which a star will move for an extinction of $A_V = 1$ mag.

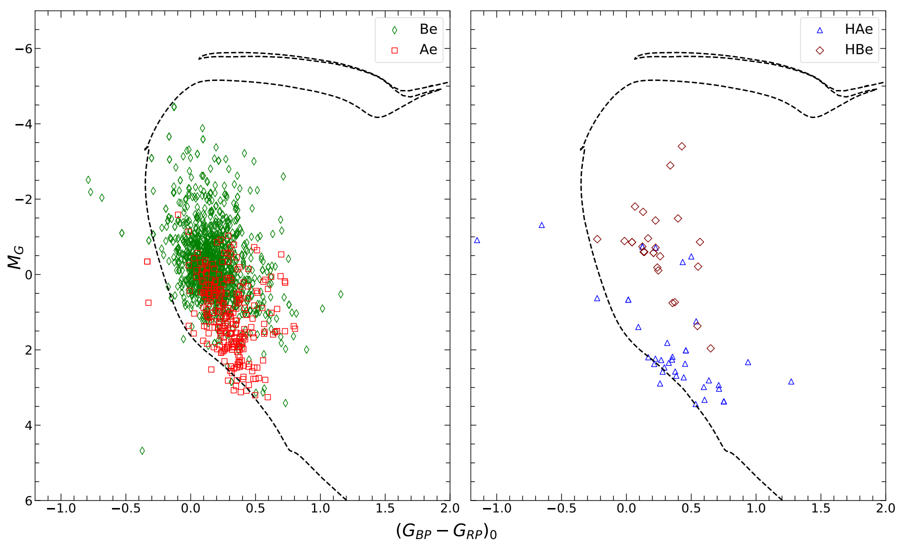

Fig. 5 *Gaia* CMDs for CBeAe (*left*) and HAeBe stars (*right*) are drawn. CBe (*green open thin diamonds*), CAe (*red open squares*), HBe (*crimson red open diamonds*) and HAe (*blue open triangles*) are included. The dotted line in both panels represents the isochrone of 60 Myr plotted in the *Gaia* CMD with $(V/V_{crit} = 0.4)$ and $[\mathrm{Fe/H}] = 0$.

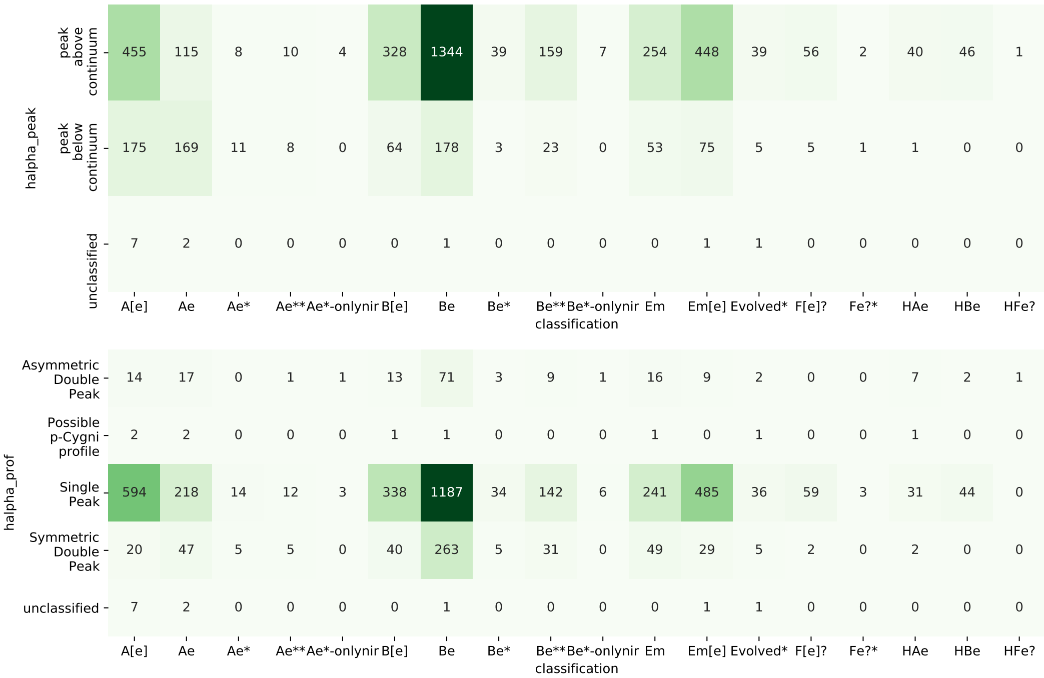

Fig. 6 (*Top*) A heatmap representing the statistical distribution of spectra based on classification of identified ELS spectra and position of Hα peak in the spectra. (*Bottom*) A heatmap representing the statistical distribution of spectra based on classification of identified ELS spectra and the morphology of Hα peak in the spectra. The number of spectra belonging to each class is expressed inside the respective boxes.

this work. A breakdown of 3339 ELSs identified is featured in Table 1. detailed examination of these lines helps us understand the properties of the gaseous disk.

4.3.1 Morphology of Hα

In addition to classification of spectra explained in Section 3, we visually examined the spectra and noted different morphologies of Hα emission seen in our sample. Out of 4138 ELS spectra identified in this work, 81% of the spectra show strong emission with the Hα peak above the continuum and remaining 19% show weaker emission where the Hα peak is seen inside the absorption core (emission-in-absorption). Further, 83% of the spectra display a single peak profile, 12% exhibit a symmetric double peak emission profile and 4% show an asymmetric double peak emission profile. Statistical breakdown of 4138 spectra based on Hα profile morphology, Hα peak and classification is depicted in Figure 6 as a heatmap.

4.3.2 Classical Be stars

We identified a homogeneous sample of 1523 spectra belonging to 1089 CBe stars in this study. In addition to the photospheric absorption lines, the spectra of CBe stars generally display emission lines of different elements such as hydrogen, helium, oxygen, iron, calcium, etc. A detailed examination of these lines helps us understand the properties of the gaseous disk.

As mentioned in previous literature (Banerjee et al. 2021; Rivinius et al. 2013; Saad et al. 2012; Mathew & Subramaniam 2011), our sample of CBe stars also exhibits diverse profiles in the Balmer lines. About 88% of spectra classified as CBe show Hα in central emission above continuum and 12% feature emission-in-absorption; 78% of spectra display single peak emission and 22% of the spectra show double peak emission. Around 45% of the spectra show Hβ in weak emission-in-absorption. All of our sample stars display Hγ and higher-order Balmer lines in absorption. Interestingly, we also observe variability in line profiles/intensity in some cases. Apart from hydrogen, we observe that few CBe stars in our sample also show O I and Fe II lines in emission. Studies on the emission line features of FeII lines are sparse in literature when compared to Hα, due to the fact that Fe II lines are difficult to detect. Additionally, a closer view on the variability of Fe II lines compared to Balmer lines provides immense aid in understanding the kinematics of Be star disks. A dedicated analysis on these major line features observed in CBe stars will be presented as a follow-up work in Anusha et al. (in prep).

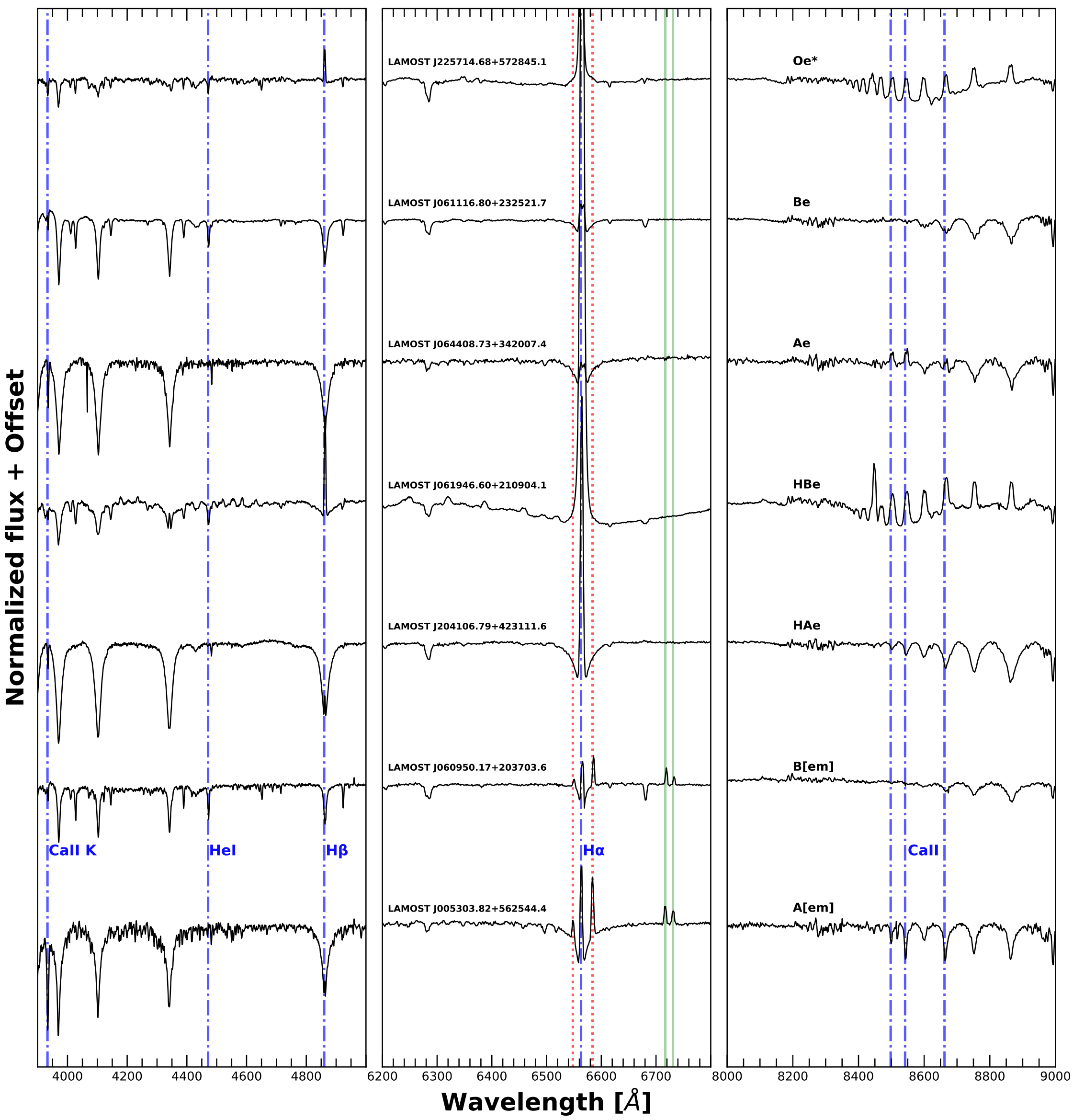

Fig. 7 This plot displays a representative spectrum belonging to each of Oe*, CBe, CAe, HBe, HAe, B[em] and A[em] classes. *Red dotted lines* signify [NII] lines at 6548 and 6584 Å and *green solid lines* correspond to [SII] lines at 6717 and 6731 Å. The spectra are normalized using the `laspec` (Zhang et al. 2020) module.

4.3.3 Classical Ae stars

We classified 286 spectra as CAe in this work with 169 spectra showing emission-in-absorption profile and 115 spectra exhibiting emission above continuum. Since A-type stars have the strongest Balmer absorption lines, we observe more spectra with the Hα emission peak below the continuum; 76% of the CAe spectra display single-peak emission, 16% exhibit symmetric double-peak emission and only 6% manifest asymmetric double-peak emission.

Out of 233 unique CAe stars identified, 159 CAe stars were analyzed in detail by Anusha et al. (2021). The mismatch in number of CAe stars reported in their work and in our sample is due to a more stringent criterion of considering V-[12] magnitude color excess, where [12] is the IRAS 12 μm magnitude from the IRAS Point Source Catalog and V queried from APASS. The CAe stars are expected to show weak Hα emission compared to CBe stars (Rivinius et al. 2013). For the sample of 159 CAe stars, they observed that Hα equivalent widths (EWs) corrected for the underlying photospheric absorption are within the range of –0.2 to –23.6 Å, which is comparatively lower than that reported for CBe stars. They also identified other emission lines such as FeII, OI, CaII triplet and Paschen series in hydrogen lines from their sample of CAe stars.

4.3.4 HAeBe stars

We identified 87 spectra belonging to 56 HAeBe stars in this work; 86 out of 87 spectra display Hα emission above the continuum; 87% of HAeBe spectra show single-peak emission, 10% exhibit asymmetric double-peak emission and less than 2% feature symmetric double peak emission. It is interesting to see that 10% of HAeBes display asymmetric double-peak emission whereas only 4% of CBeAe spectra show asymmetric double-peak emission.

Apart from the presence of Balmer lines, HAeBe stars share spectral similarities with those of CBeAe, which include OI lines, CaII IR triplet and the Paschen series of HI (Persson & McGregor 1985; Hamann & Persson 1992) are also seen in our sample of spectra. Other emission lines present in the spectra of HAeBe stars are CaII H and K lines, NaI, and MgII. Since we have removed the spectra with nebular forbidden lines from the analysis, the sample of HAeBe stars is not complete. Nidhi et al. (in prep) will consider this caveat and present an analysis of a bigger sample of intermediate-mass HAeBe stars from LAMOST DR5 including F0–F5 spectral types.

4.4 SIMBAD Crossmatch

Finally, the newly identified ELSs are cross matched with the SIMBAD database. Out of 3339 unique ELSs, 623 objects have a record in the SIMBAD database, among which 174 objects are marked as "Em*" (Emission-line star) and 17 stars are recorded as "Be*" (Be star) or "Ae*" (HAeBe star); 356 objects are represented as "*" (Star) with no specific classification; 22 objects are recorded as "Y*O" (Young Stellar Object) or "Y*?" (Young Stellar Object Candidate) and seven objects are reported as "WD?" (White Dwarf candidate) by Zhang et al. (2013). Other types that are seen are "IR" (IR source), "Ir*" (Variable Star of irregular type),"V*" (Variable star), "Ro*" (Rotationally Variable star) or "HB*" (Horizontal Branch star). All of the objects marked in SIMBAD as "Be*" are classified by us as "Be" and objects marked as "Y*O" are classified either as "HAeBe" or "A[em]/B[em]". Most of our classification is in accordance with what is reported in SIMBAD, the only notable differences are the seven stars that are classified as "WD?" by Zhang et al. (2013). In short, through this work on identifying the ELSs from LAMOST DR5, we report 2716 new ELSs which do not have any record in the SIMBAD database. Figure 7 depicts a representative sample of spectra from some of the classifications discussed in this work.

5 CONCLUSION

We provide a catalog of 3339 hot (O, B and A spectral type) ELSs identified from LAMOST DR5, of which 2716 are newly reported. Utilizing an automated Python routine, we identified 5437 spectra with Hα in emission and visually selected 4138 good quality spectra for further analysis. We noticed that the spectral type provided by LAMOST is not accurate, deviating by a few sub-classes in some cases. It is possible that emission features in the spectra are the reason for such misclassifications. We re-estimated the spectral types of 3307 O, B and A type spectra by a semi-automated Python routine that identifies three best template fits to the spectra in the blue region including higher-order Balmer lines such as Hδ and Hϵ and then the best fit of the three templates (giving more importance to HeI and MgII lines) was selected visually. Our re-estimated spectral type matches are better with spectral type given in SIMBAD than spectral type provided by LASP, which validates our technique. Once the spectral types were re-estimated, we distinguished stars with excess in NIR and MIR bands based on 2MASS $(H - K_S)_0$, (n_{J-K_S}) and (n_{K_S-W2}) values. IR excess in NIR and MIR colors indicates the presence of dust in the circumstellar disk of PMS stars. CBe stars show low excess in NIR whereas HAeBe stars manifest considerable excess in NIR and sometimes a large excess in MIR as seen in Figure 4. We categorized the Hα profiles of all 4138 emission-line spectra visually and they are represented as a heatmap with respect to each classification, as depicted in Figure 6.

We present in context a caveat associated with our classification. As explained in Subsection 3.5, in some cases stars with multiple observations can end up having multiple classifications. This occurs due to the inability of distinguishing B8/9 type from A0/1 type spectra owing to the low-resolution of LAMOST DR5 spectra. LAMOST Mid-Resolution Spectra (MRS) in Data Release 6/7 (DR6/7) will be helpful to resolve this confusion.

In this work, we report a homogeneous list of 1089 CBe, 233 CAe, 56 HAeBe stars and 686 stars categorized as "Em" or "Em[em]". In addition to these classes, we describe 240 CBeAe candidates as "Be*/**" and "Ae*/**". It may be noted that 159 CAe stars identified from this work are analyzed in detail in Anusha et al. (2021). We classified 928 objects either as "B[em]" or "A[em]" due to presence of [SII], [NII] and/or [OI] forbidden lines. Such a homogeneous list from the emission-line catalog will help the community when studying ELSs in detail without worrying about the bias associated with resolution and classification schemes when spectra are compiled from various sources. A sample of identified ELSs is provided in Appendix C and the entire catalog will be made available online in machine-readable format.

Acknowledgements We would like to thank the Science & Engineering Research Board (SERB), a statutory body of Department of Science & Technology (DST), Government of India, for funding our research under grant number CRG/2019/005380. We thank the Center for Research, CHRIST (Deemed to be University), Bangalore, India, for funding our research under the grant number MRP DSC-1932. We thank our colleagues Edwin Das and Robin Thomas for their valuable comments on the manuscript. This work has made use of data products from the Guo Shoujing Telescope (the Large Sky Area Multi-Object Fiber Spectroscopic Telescope, LAMOST). 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. We thank the SIMBAD database and the online VizieR library service for helping us in literature survey and obtaining relevant data.

Appendix A: CATALOG DESCRIPTION

Appendix B: FORBIDDEN LINES

To distinguish forbidden lines from noisy features in low-SNR spectra, we used the 'prominence' feature in the 'find_peaks' module. We flagged a forbidden line as "yes" only if the prominence is greater than 0.2; if the prominence value is between 0.1 and 0.2, it is flagged as "maybe" since it can be difficult to distinguish it from a noise feature if SNR is low and if prominence value is less than 0.1, the line is flagged as "no". As CBeAe stars do not show forbidden lines such as [SII] $\lambda\lambda$ 6717,6731, [NII] $\lambda\lambda$ 6548, 6584 and [OI] $\lambda\lambda$ 6300, 6363 in their spectra, it is necessary to identify and remove spectra with the above mentioned forbidden lines.

In addition to each forbidden line's flag, the final decision was taken by visually selecting the spectra without any forbidden lines. We included the flags keeping in mind the future works, where one can rely on the flags given to select a subset of spectra with specific forbidden lines.

Appendix C: SAMPLE OF ELS IDENTIFIED

References

Akras, S., Guzman-Ramirez, L., Leal-Ferreira, M. L., & Ramos-Larios, G. 2019, ApJS, 240, 21
Allen, D. A. 1973, MNRAS, 161, 145
Andre, P., Ward-Thompson, D., & Barsony, M. 1993, ApJ, 406, 122
Anusha, R., Mathew, B., Shridharan, B., et al. 2021, MNRAS, 501, 5927
Arun, R., Mathew, B., Manoj, P., et al. 2019, AJ, 157, 159
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
Banerjee, G., Mathew, B., Paul, K. T., et al. 2021, MNRAS, 500, 3926
Barentsen, G., Vink, J. S., Drew, J. E., et al. 2011, MNRAS, 415, 103
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Corradi, R. L. M., Rodríguez-Flores, E. R., Mampaso, A., et al. 2008, A&A, 480, 409
Corradi, R. L. M., Valentini, M., Munari, U., et al. 2010, A&A, 509, A41
Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, RAA (Research in Astronomy and Astrophysics), 12, 1197
Cutri, R. e., et al. 2014, VizieR Online Data Catalog, II
Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, VizieR Online Data Catalog, II/246
Deng, L.-C., Newberg, H. J., Liu, C., et al. 2012, RAA (Research in Astronomy and Astrophysics), 12, 735
Dotter, A. 2016, ApJS, 222, 8
Drew, J. E., Greimel, R., Irwin, M. J., et al. 2005, MNRAS, 362, 753
Drew, J. E., Gonzalez-Solares, E., Greimel, R., et al. 2014, MNRAS, 440, 2036
Falcón-Barroso, J., Sánchez-Blázquez, P., Vazdekis, A., et al. 2011, A&A, 532, A95
Finkenzeller, U., & Mundt, R. 1984, A&AS, 55, 109
Gehrz, R., Hackwell, J., & Jones, T. 1974, ApJ, 191, 675
González-Solares, E. A., Walton, N. A., Greimel, R., et al. 2008, MNRAS, 388, 89
Gray, R. O., & Corbally, C. J. 2009, Stellar Spectral Classification (Princeton university press)
Green, G. 2018, The Journal of Open Source Software, 3, 695
Green, G. M., Schlafly, E., Zucker, C., Speagle, J. S., & Finkbeiner, D. 2019, ApJ, 887, 93
Greene, T. P., Wilking, B. A., Andre, P., Young, E. T., & Lada, C. J. 1994, ApJ, 434, 614
Hamann, F., & Persson, S. E. 1992, ApJS, 82, 285
Hartmann, L., Megeath, S. T., Allen, L., et al. 2005, ApJ, 629, 881
Herbig, G. H. 1960, ApJS, 4, 337
Hernández, J., Calvet, N., Briceño, C., Hartmann, L., & Berlind, P. 2004, AJ, 127, 1682

Table A.1 Description of Columns Present in the Catalog

Columns	Description
$specname$	Name of FITS file as given by LAMOST
$classification$	Class that the spectrum belongs to
$catalog\_name$, $unique\_name$	Name assigned to the spectra in this work. $unique\_name$ gives the name of ELS star to which the spectrum belongs. In addition to $unique\_name$, $catalog\_name$ contains an extra number to identify the spectrum with highest $snrr$ value. That is, if an ELS has multi-epoch spectra, then the highest $snrr$ spectrum will be denoted by "_1" in its $catalog\_name$
$spectral\_type$	Spectral type assigned to the spectrum. The spectral type of best fit template from MILES library is provided.
$halpha\_profile\_comments$	Comments on H-alpha emission profile, visually checked
$forb\_visual\_comments$	Comments on forbidden lines present in spectra, visually checked
$l6\,300$, $l6\,363$, $l6\,548$, $l6\,584$, $l6\,717$, $l6\,731$	"yes" or "no", based on an automated line finder routine. The code and threshold used are explained in Appendix B.
$lamost\_design\_id$	The ID as provided in LAMOST DR5
$\_RAJ2000$, $\_DEJ2000$	Observed RA and Dec in degrees, as given in FITS header
$subclass$	Spectral type assigned by LASP, as given in LAMOST DR5.
$snru$, $snrg$, $snrr$, $snri$, $snrz$	SNR of the spectrum in different SDSS bands, as given in LAMOST DR5.
$2MASS$ to Rfl	2MASS photometry
$ALLWISE$ to qph	ALLWISE photometry
$B-V$ to u_e_Bmag	APASS photometry
Av, Av_lower, Av_upper	Extinction in V band queried from Green et al. (2019). Upper and lower bounds are given based on $r\_low\_photogeo$ and $r\_hi\_photogeo$ respectively
$n(J-K)$, $n(K-W2)$	Lada indices calculated as defined in Sect. 3.4
$(J-H)_0$ to $e\_(H-K)_0$	Extinction corrected 2MASS and *WISE* colors
abs_Vmag to $(B-V)_0$	Extinction corrected APASS absolute V magnitude along with errors and $(B-V)$ color
$gaiaedr3\_source\_id$ to $phot\_rp\_mean\_mag\_error$	*Gaia* EDR3 photometry
r_med_geo to $r\_hi\_photogeo$	Distances (geometric and photo-geometric) to the star along with upper and lower bounds based on *Gaia* EDR3 taken from Bailer-Jones et al. (2021); "geometric" distance is calculated utilizing parallax and a direction dependent distance prior, whereas "photo-geometric" uses an additional color and apparent magnitude of the star.

Hillenbrand, L. A., Strom, S. E., Vrba, F. J., & Keene, J. 1992, ApJ, 397, 613
Hou, W., Luo, A. L., Hu, J.-Y., et al. 2016, RAA (Research in Astronomy and Astrophysics), 16, 138
Huang, Y. F., Zeng Li, J., Rector, T. A., & Mallamaci, C. C. 2013, AJ, 145, 126
Hutchings, J. B. 1970, MNRAS, 147, 161
Jaschek, M., & Egret, D. 1982, in IAU Symposium, Vol. 98, Be Stars, eds. M. Jaschek & H. G. Groth, 261
Jaschek, M., Jaschek, C., & Egret, D. 1986, A&A, 158, 325
Jordi, C., Gebran, M., Carrasco, J. M., et al. 2010, A&A, 523, A48
Joy, A. H. 1945, ApJ, 102, 168
Kalari, V. M., Vink, J. S., Drew, J. E., et al. 2015, MNRAS, 453, 1026
Kogure, T., & Leung, K. C. 2007, A Book on Astrophysics of Emission-Line Stars, Astronomical Society of the Pacific Conference Series, 362, 260
Kohoutek, L., & Wehmeyer, R. 1999, A&AS, 134, 255
Kong, X., Luo, A., et al. 2019, in Astronomical Data Analysis Software and Systems XXVII, 523, 91
Lada, C. J. 1987, in IAU Symposium, Vol. 115, Star Forming Regions, eds. M. Peimbert & J. Jugaku, 1
Lamers, H. J., Zickgraf, F.-J., de Winter, D., Houziaux, L., & Zorec, J. 1998, A&A, 340, 117
Li, G.-W., Shi, J.-R., Yanny, B., et al. 2018, ApJ, 863, 70
Lin, C.-C., Hou, J.-L., Chen, L., et al. 2015, RAA (Research in Astronomy and Astrophysics), 15, 1325
Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021, A&A, 649, A2
Luo, A.-L., Zhao, Y.-H., Zhao, G., et al. 2015, RAA (Research in Astronomy and Astrophysics), 15, 1095
Malfait, K., Bogaert, E., & Waelkens, C. 1998, A&A, 331, 211
Mathew, B., & Subramaniam, A. 2011, preprint (arXiv:1108.5850)
Mathew, B., Subramaniam, A., & Bhatt, B. C. 2008, MNRAS, 388, 1879
Merrill, P. W., & Burwell, C. G. 1949, ApJ, 110, 387
Natta, A., Prusti, T., & Krugel, E. 1993, A&A, 275, 527
Pecaut, M. J., & Mamajek, E. E. 2013, ApJS, 208, 9
Persson, S. E., & McGregor, P. J. 1985, AJ, 90, 1860
Pettersson, B., Armond, T., & Reipurth, B. 2014, A&A, 570, A30
Porter, J. M., & Rivinius, T. 2003, PASP, 115, 1153

Table C.1 Details of Sample of ELSs Identified in Our Work

lamost_design_id	_RAJ2000	_DEJ2000	classification	spectral_type	catalog_name
LAMOST J043018.87+071233.5	67.578651	7.209331	A[em]	A1V	LEMC 2910
LAMOST J043020.98+455603.2	67.58745	45.934232	Be	B8III	LEMC 0675
LAMOST J043023.15+550408.8	67.596482	55.069113	Be*-onlynir	B8III	LEMC 0495_1
LAMOST J043025.78+490355.8	67.607458	49.0655	Em		LEMC 2122
LAMOST J043030.15+533334.7	67.625642	53.559666	Be	B5V	LEMC 0493
LAMOST J043041.64+432544.5	67.673508	43.429055	Be	B5V	LEMC 1863
LAMOST J043058.53+373855.7	67.743908	37.64883	Evolved*	A2Ib/II	LEMC 0458_1
LAMOST J043104.88+520145.7	67.770358	52.029388	Em[em]		LEMC 1862
LAMOST J043111.79+322932.7	67.79914	32.492425	Be	B9III	LEMC 0896_1
LAMOST J043131.26+475750.7	67.880284	47.964108	Be	B2IIIvar	LEMC 0204_1
LAMOST J043133.92+523039.2	67.891338	52.510914	A[em]	A0III	LEMC 3203
LAMOST J043134.99+493650.1	67.895833	49.613944	Em		LEMC 2975
LAMOST J043153.91+430248.7	67.974655	43.046862	Be	B6IV	LEMC 2246
LAMOST J043220.04+513458.1	68.083504	51.582813	Ae	A1V	LEMC 2131
LAMOST J043226.98+343056.2	68.11243	34.515624	B[em]	B8V	LEMC 0701
LAMOST J043229.03+540352.2	68.120962	54.064511	Be	B8III-IV	LEMC 0408_1
LAMOST J043230.94+531036.0	68.128957	53.176687	Be	B5V	LEMC 1774_1
LAMOST J043244.07+554120.0	68.183645	55.688909	Be	B8	LEMC 0095_1
LAMOST J043247.86+455849.2	68.199448	45.980349	B[em]	B8V	LEMC 1947
LAMOST J043308.59+382811.8	68.285831	38.469964	Em[em]		LEMC 1628_1
LAMOST J043324.72+405849.1	68.353	40.980333	Be	B8	LEMC 1382_1
LAMOST J043327.11+512242.3	68.362983	51.37842	Be**	B8	LEMC 0020
LAMOST J043327.49+512242.4	68.364582	51.378454	Be**	B8	LEMC 0909
LAMOST J043347.36+330254.0	68.44737	33.048357	Em[em]		LEMC 2725
LAMOST J043431.86+442815.7	68.632762	44.471051	Be	B8V	LEMC 1205
LAMOST J043434.43+505322.9	68.643483	50.88972	Be	B9p+...	LEMC 1663_1
LAMOST J043434.61-025615.8	68.644226	-2.937746	Ae*	A1IV	LEMC 2382

Raddi, R., Drew, J. E., Fabregat, J., et al. 2013, MNRAS, 430, 2169

Reipurth, B., Pettersson, B., Armond, T., Bally, J., & Vaz, L. P. R. 2004, AJ, 127, 1117

Rivinius, T., Carciofi, A. C., & Martayan, C. 2013, A&A Rev., 21, 69

Rodríguez-Flores, E. R., Corradi, R. L. M., Mampaso, A., et al. 2014, A&A, 567, A49

Rosendhal, J. D. 1973, ApJ, 186, 909

Saad, S., Hamdy, M., & Abolazm, M. 2012, NRIAG Journal of Astronomy and Geophysics, 1, 97

Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, MNRAS, 371, 703

Slettebak, A., Kuzma, T. J., & Collins, G. W., I. 1980, ApJ, 242, 171

Stephenson, C., & Sanduleak, N. 1977, ApJS, 33, 459

Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, AJ, 123, 485

Swings, J. P. 1973, A&A, 26, 443

Vink, J. S., Drew, J. E., Steeghs, D., et al. 2008, MNRAS, 387, 308

Vioque, M., Oudmaijer, R. D., Schreiner, M., et al. 2020, A&A, 638, A21

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Štefl, S., Baade, D., Rivinius, T., et al. 2003, A&A, 411, 167

Wackerling, L. R. 1970, MNRAS, 149, 405

Wang, S., & Chen, X. 2019, ApJ, 877, 116

Wang, S.-G., Su, D.-Q., Chu, Y.-Q., Cui, X., & Wang, Y.-N. 1996, Appl. Opt., 35, 5155

Waters, L. B. F. M., & Waelkens, C. 1998, ARA&A, 36, 233

Wei, M., & Stover, R. J. 1996, in Solid State Sensor Arrays and CCD Cameras, 2654, International Society for Optics and Photonics, 226

Wei, P., Luo, A., Li, Y., et al. 2014, AJ, 147, 101

Weymann, R. 1963, ARA&A, 1, 97

Wilking, B. A. 1989, PASP, 101, 229

Witham, A. R., Knigge, C., Drew, J. E., et al. 2008, MNRAS, 384, 1277

Witham, A. R., Knigge, C., Aungwerojwit, A., et al. 2007, MNRAS, 382, 1158

Wu, Y., Luo, A.-L., Li, H.-N., et al. 2011, RAA (Research in Astronomy and Astrophysics), 11, 924

Yao, Y., Liu, C., Deng, L., de Grijs, R., & Matsunaga, N. 2017, ApJS, 232, 16

Yuan, H. B., Liu, X. W., Huo, Z. Y., et al. 2015, MNRAS, 448, 855

Zhang, B., Liu, C., & Deng, L.-C. 2020, ApJS, 246, 9

Zhang, Y.-Y., Deng, L.-C., Liu, C., et al. 2013, AJ, 146, 34

Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, RAA (Research in Astronomy and Astrophysics), 12, 723