universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Detection of chemically peculiar stars of the upper main sequence in the CoRoT fields“

verfasst von / submitted by

Andreas Herdin, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2017 / Vienna 2017

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 861
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Astronomie
Betreut von / Supervisor:	Assoc. Prof. Dr. Ernst Paunzen
Mitbetreut von / Co-Supervisor:	

Abstract

In this Master's Thesis I search for chemically peculiar stars of the upper main sequence in 14468 precalibrated, low resolution spectra $(R = 1300)$ from stars in the CoRoT fields IRa1 (4000 spectra), LRa1 (1659 spectra) and LRa2 (8809 spectra) obtained by Sebastian et al. (2012). Low quality spectra have been removed from the three samples of spectra because they contain too low signals, too many negative counts or rectangular-like patterns. I removed bad counts that occurred systematically and non systematically from the remaining spectra and interpolated all cut out counts. Spectra where wave-like patterns are present have been detected by fitting a heuristic function to the spectral counts distribution and those spectra have been removed from the samples. The samples contained multiple spectra per observed star, so only one spectrum per star has been kept in the samples. A basic algorithm to compute the radial velocities and the rest frame wavelengths of the spectra has been applied. The spectral types are derived via `MKCLASS` (Gray & Corbally 2014) or taken from Sebastian et al. (2012). Photometric magnitudes of the spectra are computed for two similar photometric systems (Δa, $\Delta a'$), a Δa and a $\Delta a'$ analysis is carried out for each of the three samples. The statistical outliers of each sample with $|\Delta a| > 2\sigma_{\Delta a}$ or $|\Delta a'| > 2\sigma_{\Delta a'}$ are investigated manually to confirm the chemical peculiarities or reject false detections. By doing so, 15 stars which are most likely chemically peculiar stars and further 158 stars which are possible chemically peculiar stars have been found.

Contents

1 Introduction

1.1 Classification of chemically peculiar stars

The upper main sequence of the HRD (Hertzsprung-Russell diagram) exhibits an extraordinary rich variety of chemically peculiar stars (CP stars) in terms of over- or under abundances of many different chemical elements compared to the non CP or ”normal” stars with comparable spectral types in the MK system. Dviding these stars into groups according to their chemical peculiarities has been done several times, the most famous classification scheme has been invented by Preston (1974). He divided chemically peculiar stars into four groups: CP1 to CP4, following a temperature sequence ranging from 7000 K to 20000 K. Before this nomenclature has been introduced, the CP1 stars have been referred to as ”Am” stars with the suffix ”m” for ”metallic”, meaning that significantly stronger spectral absorption lines of most metals can be seen in the optical domain of stellar spectra than in spectra of stars with similar effective temperatures. Originally they were discussed as spectra with ”considerable weaker Ca II lines at 3933 Å (K line) than expected for the average metallic-type stars” (Roman et al. 1948), suggesting ”early A-type spectral types, while other metallic lines were late A or F”, with ”intermediate hydrogen lines” (Conti 1970). Am stars are described by a sequence of three spectral types based on the line types (h = Balmer lines of hydrogen, k = K line, m = iron group elements, for instance ”hA1kA4mA8”). Following this notation if the h-, k- and m-spectral types each differ by at least one subclass ($Sp(k) > Sp(h) > Sp(m)$) the star is classified as ”Am” star. CP2 stars are equivalent to ”Ap” stars, with ”p” for ”peculiar”, because those spectra show Si, Sr, Cr and rare earth elements (most famously Eu, Sc and Y) overabundances. The ”HgMn” stars became CP3 stars and finally ”He-weak” stars were labeled as CP4 stars. Preston summed up several astrophysical properties of these groups, but some of them were rather rough estimates or simply unknown, see table 1.1.

Not all chemically peculiar stars were covered by his scheme, which lead over time to several modifications of this scheme like adding more groups and redefining or refining the existing groups, like it has been done by Smith (1996), see figure 1.2. However

CP Group	Classical Name	Discovery Criteria	Other Properties: Rotation	Binary Frequency	Binary Periods	Temperature Domain
1	metallic line (Am)	wk Ca II and/or Sc II; enhanced heavy metals	slow	high	abnormal	7000–10,000°K
2	magnetic Ap	enhanced Si, Cr, Sr, Eu, et al	slow	low	abnormal	8000–15,000°K
3	HgMn	enhanced Hg II λ3984, Mn II	very slow	normal?	abnormal?	10,000–15,000°K
4	He-weak	$Q(\mathrm{Sp}) > Q(UBV)$	slow?	?	?	13,000–20,000°K?

Figure 1.1: The original classification scheme from Preston (1974)

the original nomenclature is basically kept, plus some new groups have been added, resulting in a combination of CP1-4 and other designations. Several descriptions and the ranges of astrophysical properties have been refined.

The differences between Smith's scheme and up-to-date schemes are mainly the subdivision of CP2 stars, according to their actual overabundant chemical elements, as they seem to follow a temperature sequence, where also transitions between several subgroups occur (Pöhnl et al. 2005). Recently a new approach has been suggested by Smalley et al. (2017), where the Am-type stars are split up into "classical Am-type stars" and "marginal Am-type stars", depending on whether the k- and m-spectral types differ by at least 5 subclasses. He-rich stars are counted to the CP4 group, leading to CP4a (He-weak) and CP4b (He-strong) subgroups and those subgroups split up into more groups based on their chemical abundances. Within the scope of my Master's Thesis I use a the designations of all major classes (CP1-4, Ap/Bp, Am, He-weak, λ Bootis, ...), but not the more detailed subclasses, so basically I stick to Smith's scheme.

Chemically peculiar stars of the upper main sequence

Classical name	Preston's group	Discovery criteria	Spectral types	Temperature domain
λ Boo	–	weak Mg II and weak metals	A0–F0	7500–9000
Am–Fm	CP1	weak Ca II and/or Sc II; enhanced metals	A0–F4	7000–10 000
Bp–Ap	CP2	enhanced Sr, Cr, Eu, and/or Si	B6–F4	7000–16 000
HgMn	CP3	enhanced Hg II and/or Mn II	B6–A0	10 500–16 000
He-weak	CP4	weak He I compared with colours	B2–B8	14 000–20 000
He-rich	–	enhanced He I compared with colours	B2	20 000–25 000

Figure 1.2: A modified classification scheme of chemically peculiar stars from Smith (1996)

1.2 Astrophysical properties of chemically peculiar stars

The classification of chemically peculiar is based on chemical abundances, but also several astrophysical characteristics differ between the groups, which will be discussed briefly:

- CP1/Am stars: Almost all Am-type stars appear to be in short period binary systems (Smalley et al. 2014), some of them do pulsate and no strong global magnetic field has been observed yet (Paunzen & Maitzen). They show chemical underabundances in C, O, Ca and Sc, but apart from that overabundances in all other elements (Fossati et al. 2008).
- CP2/Ap & Bp stars: These stars host strong global magnetic fields ranging up to several kG (Preston 1974) and they have surface spots of overabundant chemical elements, which appear cooler and darker (in the optical wavelength domain) than

the surrounding stellar surface (Kochukhov 2017). Due to rotation of the star the surface spots wander around periodically leading to a variability of the star for an observer. Hence by measuring the variability one can directly determine the rotational period of such a star. However generally they don't pulsate, except for the roAp (rapidly oscillating) stars (Paunzen & Maitzen). Their rotational velocities are very low compared to the non chemically peculiar stars of the same spectral type and the rotational axis and magnetic field axis is tilted as explained by the oblique rotator model (Stibbs 1950). The most famous characteristic features of CP2 spectra are flux depressions around 4100 Å, 5200 Å and 6300 Å.

- CP3/HgMn stars: Just like CP2 stars, the CP3 stars also show chemical spots on the surface (Hg, Mn), but it is still unclear whether those spots have a magnetic origin, because CP3 stars do not show strong global magnetic fields (Urpin 2016). They are too slow rotators (Preston 1974) and no significant variability has been found yet, but the surface spots appear to be variable on a very short timescale (Kochukhov et al. 2014).
- CP4/He-weak & He-strong stars: These stars exhibit strong global magnetic fields and show variability, but while the He-weak stars are on average cooler and provide He underabundances, the He-rich stars are hotter and show overabundances in He (Paunzen & Maitzen).
- λ Bootis stars: These stars are named after their prototype "λ Bootis" and have about solar abundances in C, N, O, S, but are metal poor in all other metals (Paunzen 2004).

1.3 The Δa photometric system

The Δa photometric system has been invented by Maitzen (1976) and it essentially consists of three narrow band filters (g_1, g_2, y), which cover a wavelength range from 5000 Å to 5500 Å. The classic Δa photometric system, which has undergone some minor adaptions and the modified version both are displayed in figure 1.3.

The original filter curves (g_1 & g_2) are chosen by Maitzen such that the g_2, centered at

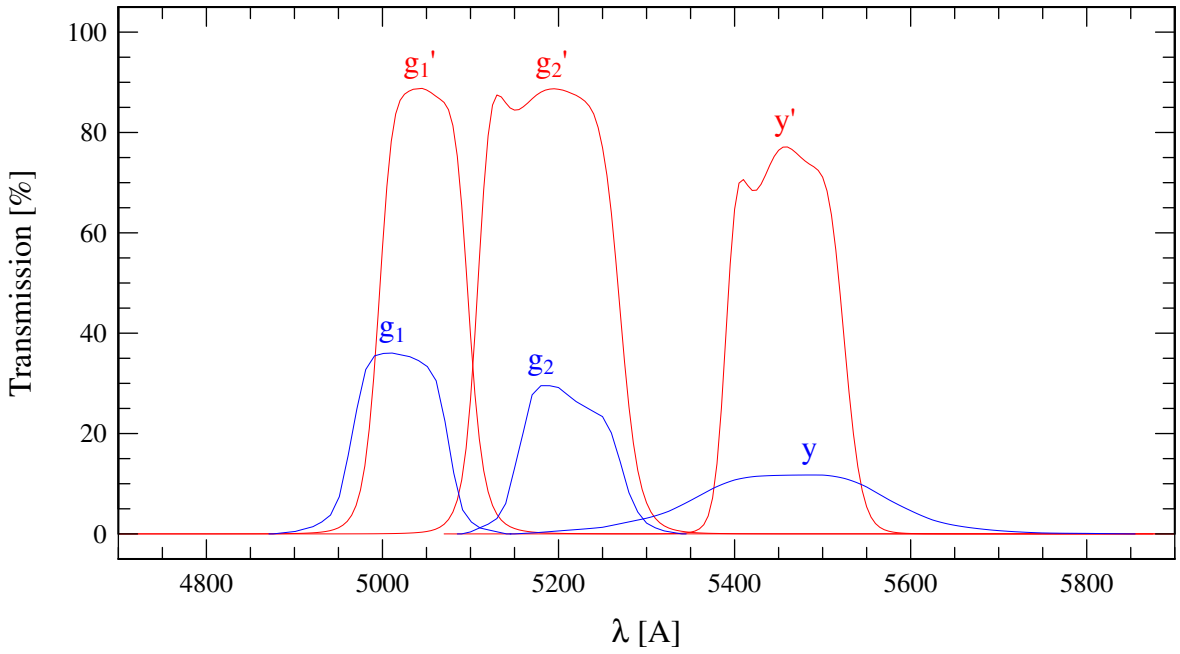

Figure 1.3: The filter transmission curves for all six filters. g_1, g_2, y form the classic Δa system, g_1', g_2' and y' the modified version.

5205 Å, covers the famous flux depression around 5200 Å, while the other two lie left (g_1, centered at 5027 Å) and right (y from the $uvby\beta$ Stroemgren-Crawford photometric filter system, centered at 5476 Å) to it. The measured fluxes in those filters are converted into magnitudes and compared on the magnitude scale:

$$a = g_2 - \frac{g_1 + y}{2} \tag{1.1}$$

This equation computes the difference between the g_2 magnitude and the arithmetic mean of the g_1 and y magnitudes. The result is referred to as a and this quantity is a measure of the total flux depression. However, it also depends on temperature and on average the depth of the flux depression increases towards cooler stars. Hence a set of reference stars along the temperature sequence (or spectral sequence) has to be defined: Those stars have to be "normal" by means of no chemical peculiarities or variabilities are allowed and they have to be main sequence stars. Their a magnitudes are used to derive the so called "normality line", a linear relation between colors $(g_1 - y)$ and a_0 magnitudes by performing a linear regression analysis on $(g_1 - y)$ and a of the reference stars. Now

Δa is defined as the difference between the measured a of a star and a_0 or in other words it can be described as flux depression compared to a typical flux depression at a given color (or temperature) and this quantity and is computed for all stars of the sample.

$$\Delta a = a - a_0 \tag{1.2}$$

The final result is a color versus Δa plot, where the Δa indices of the sample of stars scatter around zero and statistical outliers, which are out of a $|\Delta a| > 3\sigma_{\Delta a}$ range, are claimed to be chemically peculiar stars (mostly CP2 stars). These stars show either a large flux depression ($\Delta a > 0$) or a flux excess ($\Delta a < 0$) around 5200 Å compared to the other stars of the same sample. Since always a sample of stars is used, Δa can not be computed for a single star or very few stars and it is an intrinsic photometric system.

2 Basic numerical concepts

2.1 Line strength & spectral type estimator

Line strength estimators (LSEs) are used to determine a rough stellar spectral type and obtain an one-dimensional (1D) spectral type sequence, which is approximately based on an effective temperature sequence. This method provides an overview of the spectral type at hand and it is only used to distribute the spectra to different numerical routines, but within this Master's Thesis LSE are not considered as results of spectral type classification. In order to derive accurate spectral and luminosity classes the spectral classification routine `MKCLASS` is used and also the already published spectral types will be taken into account. The idea behind a LSE is to obtain a rough estimate of a spectral type in terms of purely numerical values from an even reddened (due to interstellar extinction) and radial velocity uncorrected spectrum. The aim is to compute the interdecile range of relative counts (or relative fluxes) within a sufficient large wavelength domain around the desired spectral line rest frame wavelength. This quantity is labeled as LSE and for each spectrum several LSEs are derived, based on the amount of spectral lines of interest. For one spectrum alone no scaling (absolute counts to relative counts) would have been needed, but due to the comparison of many spectra it is essential to use relative instead of absolute counts. In this Master's Thesis I scale each spectrum to its counts level at 5200 Å before computing the LSEs.

A LSE is computed in the following way: Assume $\lambda_{\circ,l}$ to be the rest frame wavelength of the spectral line l, then the wavelength range L will contribute to $\lambda_{\circ,l} - \Delta\lambda_l \leq L \leq \lambda_{\circ,l} + \Delta\lambda_l$, with a predefined line shift $\Delta\lambda_l$ (i.e. the Doppler shift). The relative counts within this range are $\tilde{y}_L$, the p-quantile function (with p between 0 and 1) is $q_p(\tilde{y}_L)$, so the interdecile range and LSE of the spectral line l is defined as

$$\mathrm{LSE}_l = q_{0.9}(\tilde{y}_L) - q_{0.1}(\tilde{y}_L) \tag{2.1}$$

In this Master's Thesis the `R` built-in quantile function `quantile` with `type = 7` (default mode) is always used. To obtain a 1D spectral sequence I chose the wavelength $\lambda =$

5175 Å and a range of $\Delta\lambda_{Mg_b} = 12.5\,A$, which is enclosing the temperature dependent Mg line triplet (5167 Å, 5172 Å, 5184 Å) and computed LSE_{Mg_b} for all spectra within a sample. Since the spectra are scaled to the counts at 5200 Å, the slope of the continuum around this line triplet is always around 1, hence the effect of interstellar reddening can be neglected. The broad range of 25 Å (twice $\Delta\lambda_{Mg_b}$) allows to compute the LSE even for stars with high radial velocities. In the next step the remaining LSEs of all spectra are computed and sorted by LSE_{Mg_b} in ascending order. Ascending LSE_{Mg_b} corresponds to a rough temperature sequence and the results are plotted in figure 2.1.

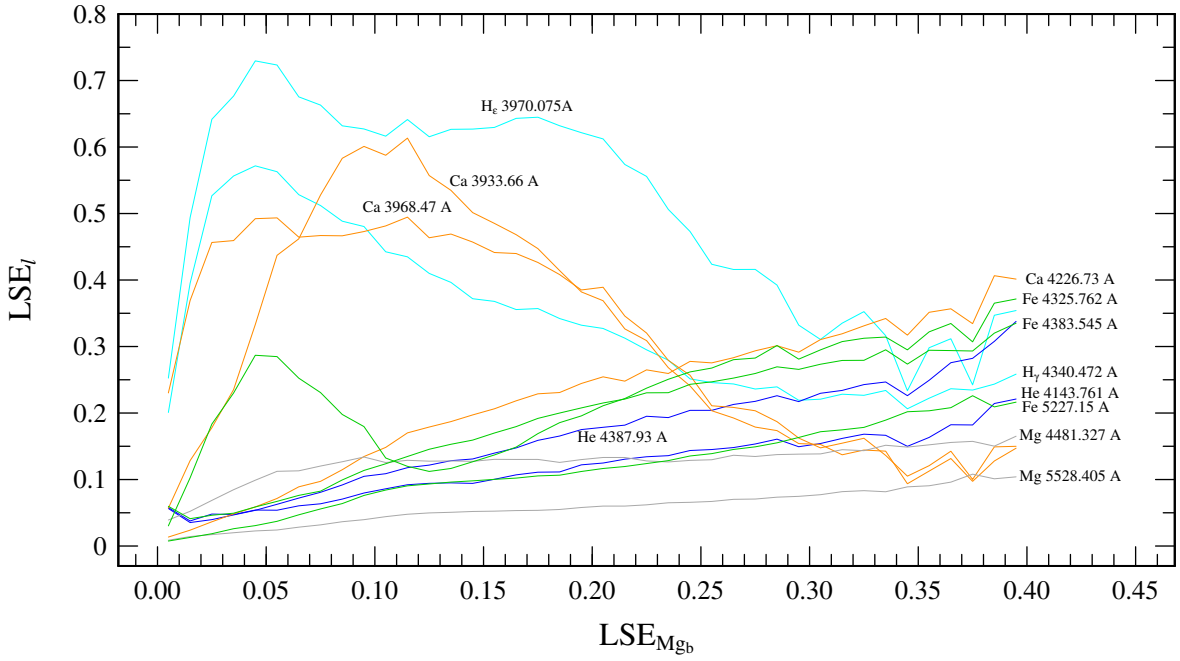

Figure 2.1: LSE_{Mg_b} plotted against LSE_l of a couple of spectral absorption lines from table 3.4 for the LRa2 sample. H is turquoise, He is blue, Ca orange, Mg is grey and Fe green. LSE_{Mg_b} is computed for all spectra of a sample and binned in steps of 0.01 to calculate the median of LSE_l within each bin. The vertical spread within a bin is typically in the range of 0.05 to 0.10, but for clarity it is not shown in the figure. Low LSE_{Mg_b} values correspond to hot stars with He and H lines, while high LSE_{Mg_b} values resemble cool stars with Ca and Fe lines. Some spectral lines are close to each other or overlap, like the Ca II line at 3968.47 Å and the H_ϵ line at 3970.075 Å, so caution has to be taken.

2.2 Spectral line fitting technique

Spectral lines are fitted by the function f

$$f(\lambda|\hat{\lambda},\sigma,h,d) = d \cdot [1 - h \cdot G(\lambda)] = d \cdot \left[1 - \frac{h}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(\lambda-\hat{\lambda})^2}{2\sigma^2}\right)\right] \tag{2.2}$$

with four parameters $\hat{\lambda}, \sigma, h$ and d as model parameters. It contains a Gaussian $G(\lambda)$ with a mean value $\hat{\lambda}$ and a standard deviation σ. h is an amplitude and d the level of the surrounding continuum. Basically it corresponds to a Gaussian flipped along the y-axis and scaled by a constant. This formula is a simplification, because usually spectral lines have to be fitted by a convolution of a Gaussian and Lorentzian (the so called Voigt profile), thus taken into account rotational broadening, the instrumental line profile and some other astrophysical effects, but in this Master's Thesis I am only interested in the position of the minimum of a spectral line and not in exact line measurements (line depth, line width, etc.), as with this formula I am not going to derive any astrophysical parameters apart from the radial velocity. For that reason this simplification will provide results with sufficient accuracy.
One might notice that $\frac{h}{\sigma\sqrt{2\pi}}$ is the normalised line depth of the minimum of f. In spectral line analysis sometimes a different parameter is used instead of this term, $\epsilon := \epsilon(\sigma, h) = \frac{h}{\sigma\sqrt{2\pi}}$, where ϵ is referred to as amplitude, instead of h. However this implies that σ and h can not be determined independently anymore, since a set of four new parameters $(\hat{\lambda}, \sigma, \epsilon, d)$ is used and this might influence the optimization process of a function containing a Gaussian. Of course one could use a function like

$$f_{mod}(\lambda|\hat{\lambda},\sigma,\epsilon,d) = d \cdot \left[1 - \epsilon \cdot \exp\left(-\frac{(\lambda-\hat{\lambda})^2}{2\sigma^2}\right)\right] \tag{2.3}$$

treat σ and ϵ as independent variables, but mathematically speaking f_{mod} does not contain a Gaussian. On the other hand a Gaussian or (density of a normal distribution) can be larger than one, because the scaling factor $\frac{1}{\sigma\sqrt{2\pi}}$ ensures that only the area below the density is always equal to one. This means that $1 - G(\lambda)$ and therefore also f can be negative and to avoid this behavior one has to introduce the condition $h \leq \sigma\sqrt{2\pi}$. Again

we end up with a relation between σ and h similar to ϵ, but in this case the advantage is that this condition is used only before or after, but not during the optimization process. The data to fit consist of wavelengths in a specific wavelength domain l and their corresponding (relative) counts (λ_l, y_l) or (λ_l^k, y_l^k), if referring to as discrete data points with k as running index of all data points within the interval l. A linear least squares minimization method of f with given $(\hat{\lambda}, \sigma, h, d)$ and the data (λ_l, y_l) is performed

$$\sqrt{\sum_k \left(\left(y_l^k - f(\lambda_l^k | \hat{\lambda}, \sigma, h, d) \right)^2 \right)} \rightarrow \min \tag{2.4}$$

by applying `optim`, a `R` built-in general-purpose optimization routine on f and equation 2.4. The four parameters $(\hat{\lambda}, \sigma, h, d)$ are used as initial values for `optim` and will be iterated according to the optimization method used (here "BFGS", the Broyden-Fletcher-Goldfarb-Shanno algorithm[1]). The best-fit parameters $(\hat{\lambda}_{min}, \sigma_{min}, h_{min}, d_{min})$ are returned (next to some other parameters) once `optim` either has found a minimum of equation 2.4 or the iteration stops once it has reached the maximum number of calls of f.

The important issue is how to derive accurate initial parameters from the data (λ_l, y_l): d can be computed directly from y_l by

$$d = 1.05 \cdot \mathrm{q}_{0.95}(y_l) \tag{2.5}$$

with $q_{0.95}$ as 95% quantile of y_l. Since by definition $\mathrm{q}_{0.95}(y_l) < \max(y_l)$, the factor 1.05 scales d to the suspected continuum level and this method is more robust to statistical outliers within the range of l than simply computing $\max(y_l)$.
$\hat{\lambda}_{min}$ and σ_{min} are both determined by a smooth natural spline fitted to (λ_l, y_l). This spline estimates the position of the minimum of the spectral line and the HWHM (half width half maximum) b. The HWHM of a Gaussian amounts to $b = \sigma\sqrt{2 \cdot \log_e(2)} \cdot$ and it is the same for f, if one adapts the definition of the HWHM: One searches for λ such that the

[1] This algorithm has been developed independently by the four mathematicians C. G. Broyden, R. Fletcher, D. Goldfarb, D. F. Shanno. It belongs to the class of quasi-Newton methods and is designed to solve unconstrained non-linear optimization problems iteratively.

midpoint of the maximum and minimum of $f(\lambda|\hat{\lambda},\sigma,h,d)$ intersects $f(\lambda|\hat{\lambda},\sigma,h,d)$.

$$\frac{\max(f(\lambda|\hat{\lambda},\sigma,h,d)) + \min(f(\lambda|\hat{\lambda},\sigma,h,d))}{2} = f(\lambda|\hat{\lambda},\sigma,h,d) \tag{2.6}$$

This equation can be solved for λ:

$$\frac{d + d - d \cdot h/\sigma\sqrt{2\pi}}{2} = d - d \cdot \frac{h}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(\lambda-\hat{\lambda})^2}{2\sigma^2}\right) \tag{2.7}$$

$$d \cdot \frac{h}{2\sigma\sqrt{2\pi}} = d \cdot \frac{h}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(\lambda-\hat{\lambda})^2}{2\sigma^2}\right) \tag{2.8}$$

$$\frac{1}{2} = \exp\left(-\frac{(\lambda-\hat{\lambda})^2}{2\sigma^2}\right) \tag{2.9}$$

$$\lambda_{1,2} = \hat{\lambda} \mp \sigma \cdot \sqrt{2 \cdot \ln(2)} \tag{2.10}$$

From the last equation it can be verified easily that $b = \lambda_2 - \lambda_1 = \sigma\sqrt{2 \cdot \ln(2)}$, hence

$$\sigma = \frac{b}{\sqrt{2 \cdot \ln(2)}} \tag{2.11}$$

and as the spline estimates b, σ is known too.

As for h we have already encountered the relative line depth ϵ and it is equal to the ratio of the absolute line depth δ to the continuum level d

$$\epsilon = \frac{\delta}{d} \tag{2.12}$$

The absolute line depth is approximated via the LSE of a spectral line scaled by a factor γ around 1, which yields $\epsilon \approx \gamma \cdot \mathrm{LSE}_l / d$, so we arrive at

$$\frac{h}{\sigma\sqrt{2\pi}} = \epsilon = \frac{\delta}{d} \approx \frac{\gamma \mathrm{LSE}_l}{d} \tag{2.13}$$

$$h \approx \frac{\sigma\sqrt{2\pi}\gamma \mathrm{LSE}_l}{d} \tag{2.14}$$

Since we know already σ, d and the LSE, we are able to estimate h. Finally we have reasonably well determined all four initial parameters and plug them into `optim` to launch

the optimization process according to equation 2.4.

2.3 Astronomical magnitudes

An astronomical magnitude m is a logarithmic, dimensionless value derived from the observed total photometric flux f in Wm^{-2} with respect to the observed total photometric flux of a reference source f° over a specific wavelength domain D.

$$m_D = -2.5 \cdot \log\left(\frac{f_D}{f_D^{\circ}}\right) \tag{2.15}$$

The minus sign is inserted for historical reasons. A magnitude of zero means that the specific object is just as bright as the reference object in the given wavelength domain. The photometric flux depends on the wavelength domain ranging from λ_1 to λ_2 and has to be integrated over its photometric flux per unit wavelength f_λ (which is a spectral irradiance in units of $Wm^{-2}nm^{-1}$) times the response function of a standard system $R_D(\lambda)$,

$$f_D = \int_{\lambda_{D,1}}^{\lambda_{D,2}} f_\lambda(\lambda) R_D(\lambda) d\lambda \tag{2.16}$$

which is obtained according to Bessel (2005):

> *"The passband or response function of a standard system is normally obtained by multiplying together the filter transmission, the reflectivity of the telescope mirror, the transmission of the camera optics and the quantum efficiency of the detector used".* - p. 10

Due to the previous calibration of the spectra by Sebastian et al. (2012), plus the fact that in this Master's Thesis I use relative counts and perform Δa synthetic photometry, which belongs to differential photometry, what remains of $R_D(\lambda)$ is the filter transmission curve, $T_D(\lambda)$ (see section 1.3). Since most of today's detectors count photons or photon fluxes ϕ_λ instead of their energies or energy fluxes f_λ (photons or energy both

per unit time, area and unit wavelength) one has to redefine equation 2.16

$$\phi_D = \int_{\lambda_{D,1}}^{\lambda_{D,2}} \phi_\lambda(\lambda) R_D(\lambda) d\lambda \tag{2.17}$$

with $f_\lambda(\lambda) = e \cdot \phi_\lambda(\lambda)$, $e = \frac{h \cdot c}{\lambda}$ as energy of a photon, c as speed of light and h as Plank's constant and therefore we end up with

$$m_D = -2.5 \cdot \log\left(\frac{\phi_D}{\phi_D^\circ}\right) \tag{2.18}$$

If we plug eq. 2.17 into eq. 2.18 we obtain

$$m_D = -2.5 \cdot \log\left(\frac{\int_{\lambda_{D,1}}^{\lambda_{D,2}} \phi_\lambda(\lambda) R_D(\lambda) d\lambda}{\int_{\lambda_{D,1}}^{\lambda_{D,2}} \phi_\lambda^\circ(\lambda) R_D(\lambda) d\lambda}\right) \tag{2.19}$$

or in terms of energy fluxes

$$m_D = -2.5 \cdot \log\left(\frac{\int_{\lambda_{D,1}}^{\lambda_{D,2}} \lambda f_\lambda(\lambda) R_D(\lambda) d\lambda}{\int_{\lambda_{D,1}}^{\lambda_{D,2}} \lambda f_\lambda^\circ(\lambda) R_D(\lambda) d\lambda}\right) \tag{2.20}$$

In this Master's Thesis I compute all ϕ_D and afterwards I define the subset M_D as the set of ϕ_Ds as reference sources per sample (with N_{M_D} as amount of reference sources within M_D) to calculate ϕ_D° as the mean flux per filter.

$$\phi_D^\circ = \frac{1}{N_{M_D}} \cdot \sum_{x \in M_D} \phi_D^x \tag{2.21}$$

2.4 Fluxes & synthetic photometry

In synthetic photometry astronomical fluxes are computed by convolving a spectrum of an astronomical object with the response function of the corresponding filter-instrument-camera system and converted into magnitudes. Actually this is not a convolution by mathematical terms, it is a multiplication of the response function with the stellar spectrum and an integration over a wavelength domain, see equation 2.16 or 2.17. As I have

already pointed out in this Master's Thesis the response function is equal to the filter transmission curve, $R_D(\lambda) = T_D(\lambda)$ equation 2.17 simplifies to

$$\phi_D = \int_{\lambda_{D,1}}^{\lambda_{D,2}} \phi_\lambda(\lambda) T_D(\lambda) d\lambda \tag{2.22}$$

The spectra of the stars are given, the filter transmission curves are taken from Paunzen et al. (2005), only the wavelength range $[\lambda_{D,1}, \lambda_{D,2}]$ has not been defined yet. Theoretically it should range from zero to infinity, because the astronomical object will always provide at least some contribution to the total photometric flux at any wavelength, however for practical reasons most filter transmission curves show one or several peaks along the wavelength axis and at a certain point they drop down to zero or come close to it far away from their central wavelengths. Thus is also a matter of computational time, if the wavelength domain is much larger magnitudes than the actual significant domain of a filter. To deal with the matter of choosing $[\lambda_{D,1}, \lambda_{D,2}]$ one has to take a closer look at the specifics of the filter curves: The wavelength grid points of the Δa filter curves are not equally spaced. Over a broad range the wavelength step sizes of the a filter curves correspond to 1 Å, while it amounts to 0.5 Å for the a'-filter curves, but at the far end of the grid points the step sizes range from several Ångstrom to hundreds of Ångstrom. $\lambda_{D,1}$ and $\lambda_{D,2}$ of all filters are selected such that they are either at the same wavelengths, where the filter transmission curves reach zero or close to them. To shed light on this principle, see table 2.1. The filter transmission curve boundaries are listed in table 2.2.

For each spectrum all six fluxes $(\phi_{g_1}, \phi_{g_2}, \phi_y, \phi_{g_1'}, \phi_{g_2'}, \phi_{y'})$ are computed according to equation 2.22. This is done by evaluating a spectrum via a natural spline s_D on a higher resolved wavelength grid $\lambda_{s,D}$ (current step size of $\approx$ 1 Å divided by 100) in the range of $\lambda_{min} = \min(\lambda_{D,1}) - 10$ and $\lambda_{max} = \max(\lambda_{D,2}) + 10$. The ± 10 is introduced to shift boundary effects of the spline interpolation out of the interval $[\min(\lambda_{D,1}), \max(\lambda_{D,2})]$. However, most spectra range only to 5750 Å so the y and y' filters are cut down anyway. s_D is multiplied by $\phi_\lambda(\lambda_{s,D})$ to approximate the integrand of equation 2.22.

$$\phi_\lambda(\lambda) T_D(\lambda) \approx \phi_\lambda(\lambda_{s,D}) \cdot s_D(\lambda_{s,D}) = p_D \tag{2.23}$$

$\lambda\,[A]$	transmission [%]	boundaries	$\lambda\,[A]$	transmission [%]	boundaries
3000	0.00000		4871	0.00000	
4810	0.00005	$\lambda_{g_1,1}$	...	...	
4855	0.00015		4876	0.00000	$\lambda_{g_1',1}$
4870	0.00026		4877	0.00020	
4880	0.00036		4878	0.00060	
...	...		...	...	
5240	0.00033		5139	0.00076	
5260	0.00021		5140	0.00038	
5305	0.00010	$\lambda_{g_1,2}$	5141	0.00000	$\lambda_{g_1',2}$
7990	0.00000		...	...	
12000	0.00010		5146	0.00000	

Table 2.1: The original data points of the filter curves for the filters g_1 (left) and g_1' (right). As for g_1 the first entry with zero flux lies at 3000 Å, so I pick the entry after this one as left boundary $\lambda_{g_1,1}$, which lies at 4810 Å. $\lambda_{g_1,2}$ is set to 5305 Å, because the next larger entry in the table is about 2600 Å further away. As for the g_1' filter the last entry before the first non zero flux entry is chosen, $\lambda_{g_1',1} = 4876$ Å, because it lies close to the first non zero flux entry. $\lambda_{g_1',2}$ is set to 5141 Å, because it is the first zero flux entry on the red side of the filter transmission curve. For the other filters it works the same way.

D	$\lambda_{D,1}$	$\lambda_{D,2}$
g_1	4810	5305
g_2	4945	5450
y	5145	5760
g_1'	4876	5141
g_2'	5090	5340
y′	5150	5850

Table 2.2: The blue and red boundaries of the filter curves.

All flux values of equation 2.23, whose corresponding wavelengths are smaller than $\lambda_{D,1}$ or larger than $\lambda_{D,2}$ are set to zero, because they don't lie in $[\lambda_{D,1}, \lambda_{D,2}]$. In the next step the total area $A(p_D)$ of p_D is approximated by diving the total area into small trapezes with edges $(\lambda^k_{s,D}, 0), (\lambda^{k+1}_{s,D}, 0), (\lambda^{k+1}_{s,D}, p^{k+1}_D), (\lambda^k_{s,D}, p^k_D)$ for $k = 1, ..., n_D - 1$ with n_D as number of trapezes, splitting them up into rectangles and triangles and summing up their respective areas.

$$A(p_D) = \sum_{k=1}^{n_D-1} \left(\left(\lambda^{k+1}_{s,D} - \lambda^k_{s,D}\right) \cdot p^k_D + \left(\lambda^{k+1}_{s,D} - \lambda^k_{s,D}\right) \cdot \frac{p^{k+1}_D - p^k_D}{2} \right) \tag{2.24}$$

$$= \frac{\Delta\lambda_{s,D}}{2} \cdot \sum_{k=1}^{n_D-1} \left(p^{k+1}_D + p^k_D\right) \tag{2.25}$$

Here $\Delta\lambda_{s,D} = \lambda^{k+1}_{s,D} - \lambda^k_{s,D} = \text{const.}$, because the $\lambda_{s,D}$ is a sequence of equidistant grid points and it can be taken out of the sum. Finally ϕ_D is approximated by $A(p_D)$

$$\phi_D = \int_{\lambda_{D,1}}^{\lambda_{D,2}} \phi_\lambda(\lambda) T_D(\lambda) d\lambda \approx A(p_D) = \frac{\Delta\lambda_{s,D}}{2} \cdot \sum_{k=1}^{n_D-1} \left(p^{k+1}_D + p^k_D\right) \tag{2.26}$$

To derive magnitudes from the fluxes, first one has to compute ϕ°_D, the reference flux. Here it is equal to the mean flux per filter transmission curve of a predefined set of spectra per CoRoT sample, so there are 6 mean fluxes per CoRoT sample. These spectra have to be A0V spectra with a `MKCLASS` quality flag of at least `vgood`, contain no chemical peculiarities, provide sufficient signal (high c_{norm}) and low noise, have small radial velocity errors (compared to the velocity errors of the corresponding sample, typically $< 20\,km/s$) and have to be checked manually for any bad counts. Of course such strict criteria cut the set of reference spectra down to a small number of spectra: 12 (IRa1), 4 (LRa1), 9 (LRa2). Once ϕ°_D is known, finally we plug it in eq. 2.18 and derive magnitudes.

3 Data reduction

3.1 Origin and structure of the data

The data have been obtained by Sebastian et al. (2012) at the Anglo-Australian Telescope (AAT) using the AAOmega multi-object-spectrograph and they consist, amongst others, of 14467 spectra, ranging from the blue to the visible wavelength regime, distributed over the three samples as the following: 3999 spectra from members of the CoRoT field IRa1, 1659 spectra from LRa1 members and 8809 spectra from LRa2 members. Furthermore, the combination of a CoRoT field name plus the term "sample" is referred to as the spectra of stars within the corresponding CoRoT field, e.g. the "IRa1 sample". These three samples of data are provided as unnormalized, wavelength calibrated, radial velocity uncorrected spectra containing counts along equally spaced wavelength data points and they are stored as 1D image FITS files (Flexible Image Transport System, Pence et al. 2010). The resolution of the spectra accounts to approximately 1 Å/pixel. Each file consists of a primary HDU and a table of dimension 2048×1 (IRa1 sample) or 1847×1 (LRa1 & LRa2 samples) of counts detected during a given exposure. The exposure times typically lie in the range of $300\,\text{s} \leq t_{exp} \leq 900\,\text{s}$.

3.2 Computing wavelength data points

The wavelength data points λ^k are computed by extracting the stored values of the header keys `CRVAL1`, `CDELT1`, `CRPIX1`, `CD1_1` and applying them, according to FITS standards, on to the pixel data points k for each sample, with k as

$$k \in \mathbb{N} \quad , \quad k = \begin{cases} 1, ..., 2048 & \text{for} \quad \text{IRa1} \\ 1, ..., 1847 & \text{for} \quad \text{LRa1, LRa2} \end{cases} \tag{3.1}$$

The coordinate transformation from the vector of pixel coordinate elements p_j to the vector of intermediate world coordinate elements x_i and further on to the vector of world

coordinate elements w_i is done according to Pence et al. (2010) by

$$x_i = \sum_{j=1}^{N} \left(\texttt{CDi\_j} \cdot (p_j - \texttt{CRPIXj}) \right) \tag{3.2}$$

$$w_i = \texttt{CRVALi} + x_i \tag{3.3}$$

with i as index variable for world coordinate axes, j as index variable for pixel coordinate axes, N as number of world coordinate axes. Since the spectra contain only one world coordinate $N = 1$ and therefore $i, j = 1$, thus $p_1 = 1$ (here the pixel coordinate starts at 1, which is the first element of k) the sum can be neglected and these equations simplify to

$$w_1 = \texttt{CRVAL1} + \texttt{CD1\_1} \cdot (1 - \texttt{CRPIX1}) \tag{3.4}$$

Now we have obtained the starting point for the wavelength axis w_1 in units of Angström and compute the wavelength data points λ^k, with k from equation 3.1 once per sample, because within a sample the wavelength grid points are always the same.

$$\lambda^k = w_1 + \texttt{CDELT1} \cdot (k - 1) \tag{3.5}$$

The IRa1 sample spectra have been calibrated by a given wavelength dependent function, displayed in figure 3.1, so I divided all IRa1 sample spectra by the same set of data points to recalibrate them.

$$y^k = \frac{y^k_{cal}}{\text{calibration}^k} \tag{3.6}$$

All spectra, which either do not contain a full CoRoT identification number (i. e. corot102540069), any other non-CoRoT designation (corotsky, corotparked, etc.) or whose sum of absolute counts over the given wavelength domain is smaller than 100 (i. e. spectra with zero counts) are discarded. This reduces the sizes of the samples down to 3104 (-895, IRa1 sample), 1658 (-1, LRa1 sample) and 8765 (-44, LRa2 sample) spectra.

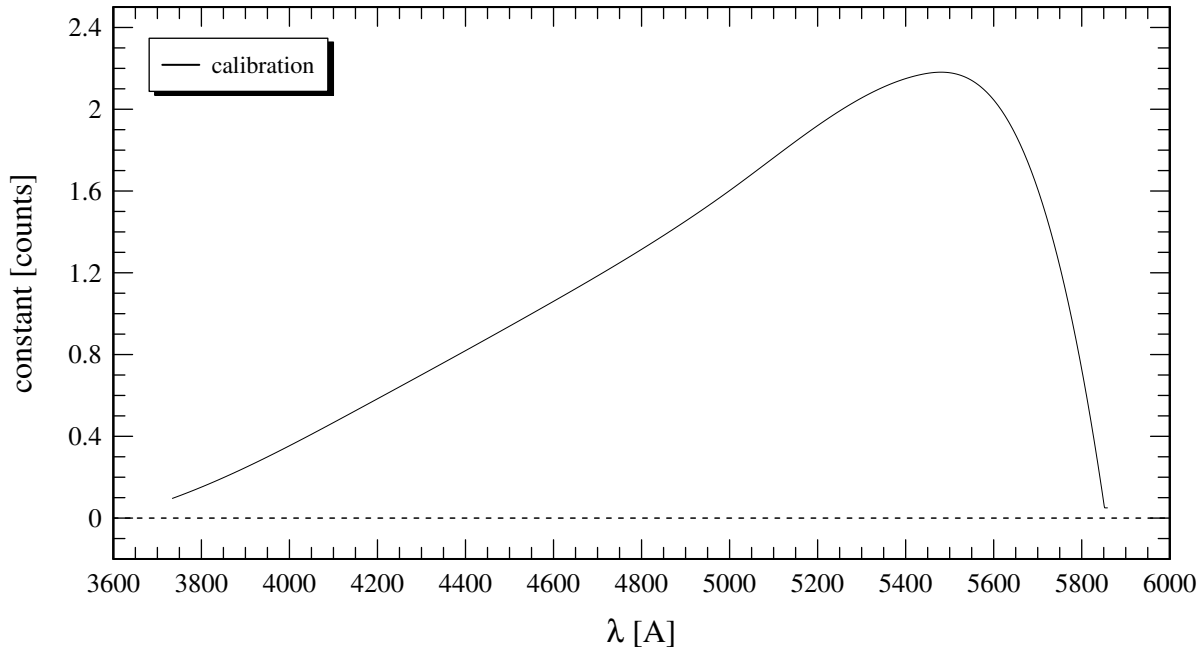

Figure 3.1: The calibration function for IRa1 spectra. For wavelengths close to the right boundary it will induce a step in the spectra, so counts with $\lambda \geq 5700$ will be lowered too much and should taken with caution.

3.3 Low signal spectra

To obtain an accurate set of spectra, I removed low signal-to-noise ratio (S/N) spectra to maintain decent correction for radial velocities and avoid spectral misclassification (i.e. misdetecting luminosity classes and chemical peculiarities) or erroneous Δa photometric results. In my thesis I did not compute S/N ratios, I calculated three counts levels for certain wavelength domains, combined them and checked if those combinations lie above or below empirically determined levels, so it is a matter of detecting "low-counts-spectra" instead of low S/N spectra. The first one, c_{norm}, is derived from a linear interpolation based on two medians of counts in the wavelength intervals $4900\,\text{Å} \leq \lambda \leq 5100\,\text{Å}$ and $5300\,\text{Å} \leq \lambda \leq 5500\,\text{Å}$ and evaluated at 5200 Å. This spectral domain has been chosen, because it is one of the regions where Δa effects occur and where the essential steps to determine Δa indices are taken, hence producing more homogeneous and comparable results. For the hotter stars this number provides a useful representation of an average counts level, because due to Plank's law there are not many lower counts at higher

wavelength domains left, since the spectra range to a maximum of 5750 Å. The other two previously mentioned counts levels are the 5% and 95% quantiles of counts in the range "abs" 3948.47 Å $\leq \lambda_{abs} \leq$ 3988.47 Å ($q_{0.05}(y_{abs})$, $q_{0.95}(y_{abs})$). This interval is chosen to include the strong Ca II at 3968.47 Å and H_ϵ at 3970.08 Å absorption lines. The spectral types of interest in this thesis range from early B-type to late F-type stars and on this sequence of spectral types either the H_ϵ or the Ca II line is going to be a dominant spectral feature. $q_{0.05}(y_{abs})$ provides an estimate of the counts level close to the core of one of the two absorption lines, while the $q_{0.95}$ estimates the counts level of the continuum. The quantiles have been chosen over maxima or minima, simply because they allow statistical outliers such as noise, flawed pixels or counts.

Four criteria are created from these three numbers to detect low counts spectra:

- c_{norm} has to be smaller than an empirically chosen counts level (IRa1: 1000 counts, LRa1: 600 counts, LRa2: 1000 counts).
- $q_{0.05}(y_{abs}) + q_{0.95}(y_{abs}) < 0$.
  As long as $q_{0.095}(y_{abs})$ is large enough, the core of the absorption line can be negative and it still will not be considered as low counts spectrum.
- $q_{0.05}(y_{abs})/c_{norm} < -0.05 \quad \wedge \quad (q_{0.95}(y_{abs}) - q_{0.05}(y_{abs}))/c_{norm} > 0.15$
  The first condition is met if the core of the spectral line is on a negative relative counts level, the second condition excludes low temperature spectra (later than $\sim$ K0), as the interquantile range will become smaller with decreasing effective temperature, so those spectra are not affected by the second part of the third criterion.

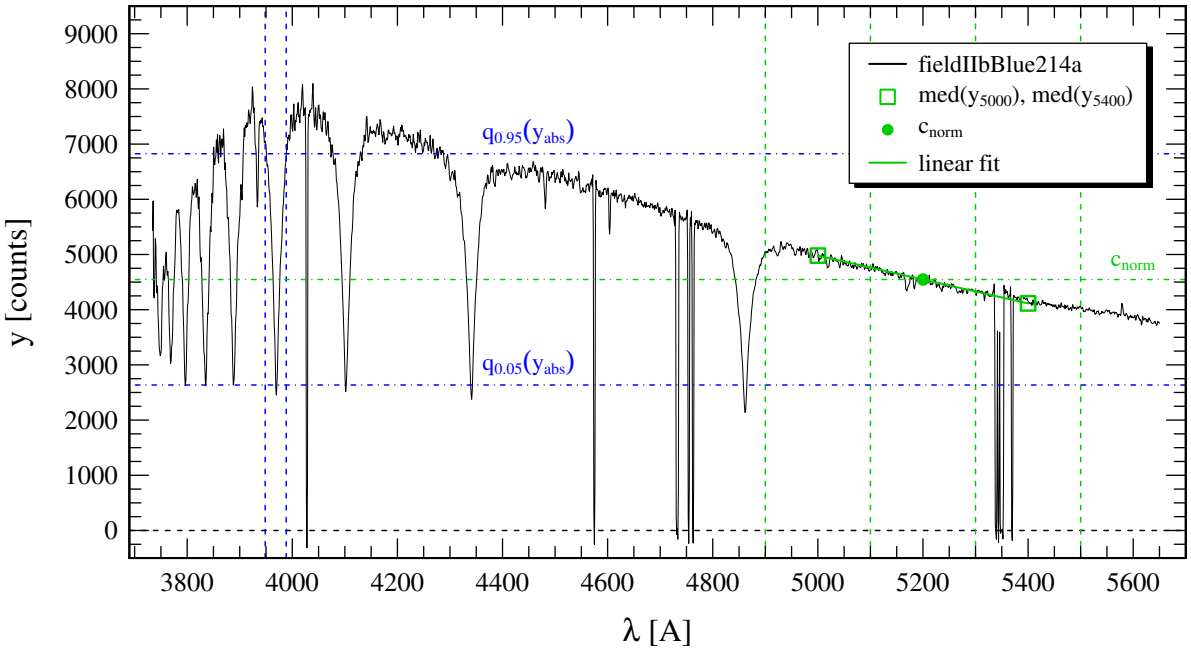

Figure 3.2: The spectrum of the star corot102540130 (fieldIIbBlue214a). Vertical lines refer to the ranges of data points used to compute the quantities $q_{0.05}(y_{abs}), q_{0.95}(y_{abs}), c_{norm}$.

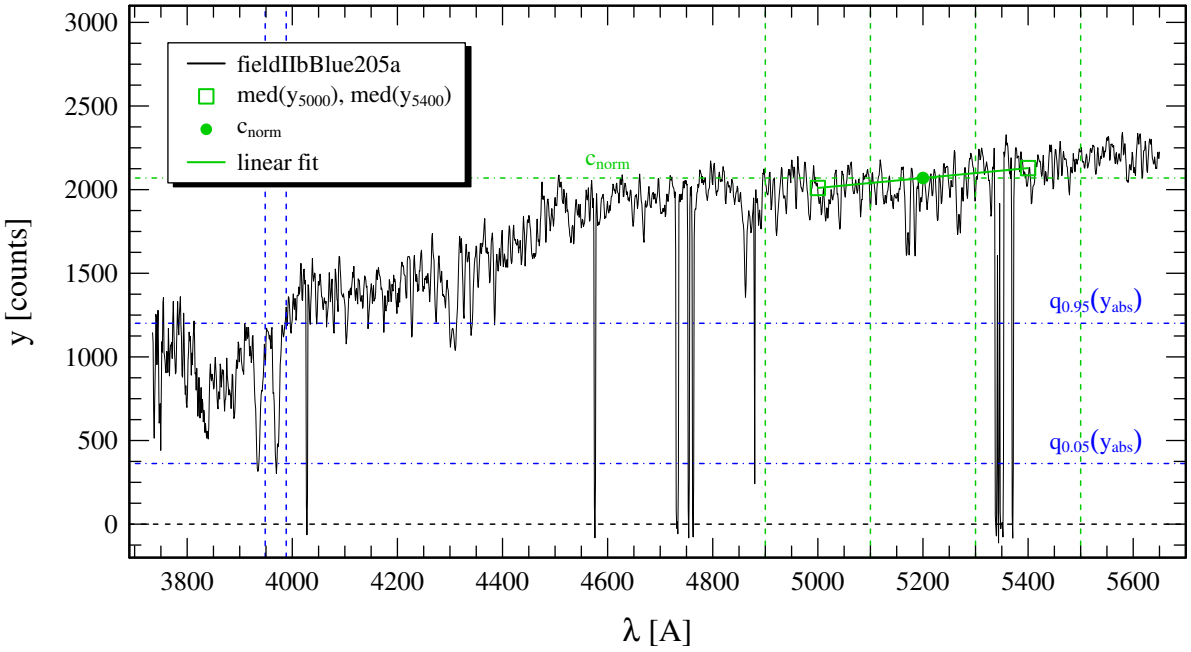

Figure 3.3: The spectrum of the star corot102544096 (fieldIIbBlue205a), see figure 3.2 for the description.

3.4 Negative counts spectra

There are spectra in the CoRoT samples, which contain a significant amount of negative or zero counts for $\lambda < 4000\,\text{Å}$ (and in a few cases even for $\lambda > 4000\,\text{Å}$), not resulting from bad pixels, but rather probably introduced by a previous calibration or data reduction analysis. To keep and not exclude such spectra they are shifted in the positive counts domain by adding an individual counts number y_{shift}. Computing a shift based on continuum estimates might be misleading, because on average the continuum counts level can be positive, while the minima counts of spectral features are not. Therefore once again I use the 5% and 95% quantiles of the absorption lines Ca II and/or H_ϵ. The shift is chosen in such a way that the shifted 95% quantile is equal to three times the shifted 5% quantile counts level or to put it in other words: The shifted 5% quantile has to be 1/3 of the shifted 95% quantile

$$q_{0.95}(y_{abs} + y_{shift}) \stackrel{!}{=} 3 \cdot [q_{0.05}(y_{abs} + y_{shift})] \tag{3.7}$$

$$q_{0.95}(y_{abs}) + y_{shift} = 3 \cdot [q_{0.05}(y_{abs}) + y_{shift}] \tag{3.8}$$

$$y_{shift} = [q_{0.95}(y_{abs}) - 3 \cdot q_{0.05}(y_{abs})]/2 \tag{3.9}$$

Here we have used the linearity of the quantile function with x as a vector or a sequence of values and a as a scalar: $q_p(x+a) = q_p(x) + q_p(a) = q_p(x) + a$. In case of $y_{shift} < 0$ the spectrum would be dragged down, but instead of doing so y_{shift} is set to 0. For $y_{shift} > 0$ the counts of a spectrum will be lifted, but only if there are at least 3 pixels close to 0 in the wavelength domain $3800\,\text{Å} \leq \lambda \leq 4000\,\text{Å}$ to avoid unnecessary shifts, as they may belong to noise or statistical outliers.

The formula in eq. 3.9 however does not imply any limit on the amount of the lifting. That means in combination with a subsequent scaling of absolute counts levels to the absolute counts levels at a certain wavelength, $(y)_{5200} := y(\lambda = 5200\,\text{Å})$, spectral features will appear weaker in relative values $(y + y_{shift})/(y + y_{shift})_{5200}$ compared to the same spectrum without a constant shift to higher counts $y/(y)_{5200}$. Therefore the limit for the magnitude of the shift is set to $0.5 \cdot c_{norm}$. Spectra with higher shifts are excluded from

the data set. The ones that fall below $0.5 \cdot c_{norm}$ are shifted to

$$y^* = y + y_{shift} + 0.1 \cdot c_{norm} \tag{3.10}$$

The last summand covers the effect of statistical outliers and the amount of shift is stored in the header of each file (either 0 or $y_{shift} + 0.1 \cdot c_{norm}$). In case of $y_{shift} > 0$, c_{norm} is recomputed for the new count levels.

3.5 Rectangular-like patterns

Several hundreds of LRa2 spectra show a superposition of the spectrum itself and randomly repeated rectangular-like functions, who either sharpen spectral lines or extinguish them completely and therefore are of no use for the further analysis. Such rectangular patterns appear to come in many different shapes, sizes and amplitudes, but almost all of them do show some kind of upper and lower "pseudo continuum". That means that most data points of a spectrum with a rectangular-like patter lie either on (or close to) an upper counts level or on (or close to) a lower counts level, but only very few counts are located in between.

This observational fact is used to distinguish between spectra with rectangular-like patter and spectra without: The aim is to normalize the spectrum to the upper pseudo continuum and to perform a basic statistical analysis on the distribution of normalised counts in their histogram. Here the upper instead of the lower pseudo continuum is used, simply because for the unpolluted spectra only an astrophysical continuum exists. In the first step the counts of a spectrum within the wavelength interval "bin" of $5015\,\text{Å} \leq \lambda_{bin} \leq 5275\,\text{Å}$ y_{bin} with a range of 260 Å are divided into eight 32.5 Å wide bins. I choose this specific interval to avoid the influence of Balmer absorption lines or wavelength domains, which are extremely affected by flawed pixels, see section 3.7.1 and to still retain a rather large wavelength range for the further analysis. The 95%-quantile of counts per bin, $y_{bin,max}$, as well as the corresponding bin central wavelengths, $\lambda_{bin,max}$, are computed for all bins.

These 8 pair of data points $(\lambda_{bin,max}, y_{bin,max})$ are used to fit a polynomial of second or-

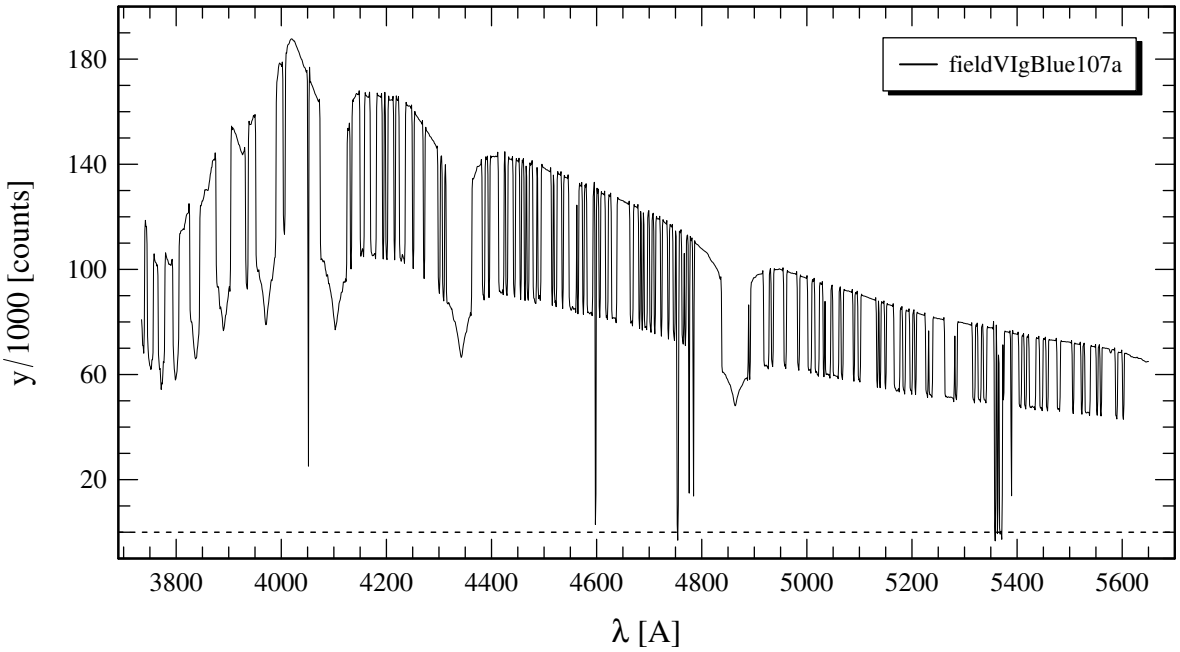

Figure 3.4: A spectrum of corot110667815 (fieldVIgBlue107a). One can clearly see the rectangular pattern overlaid onto the spectrum on the whole wavelength range. The upper and lower pseudo continua are clearly visible.

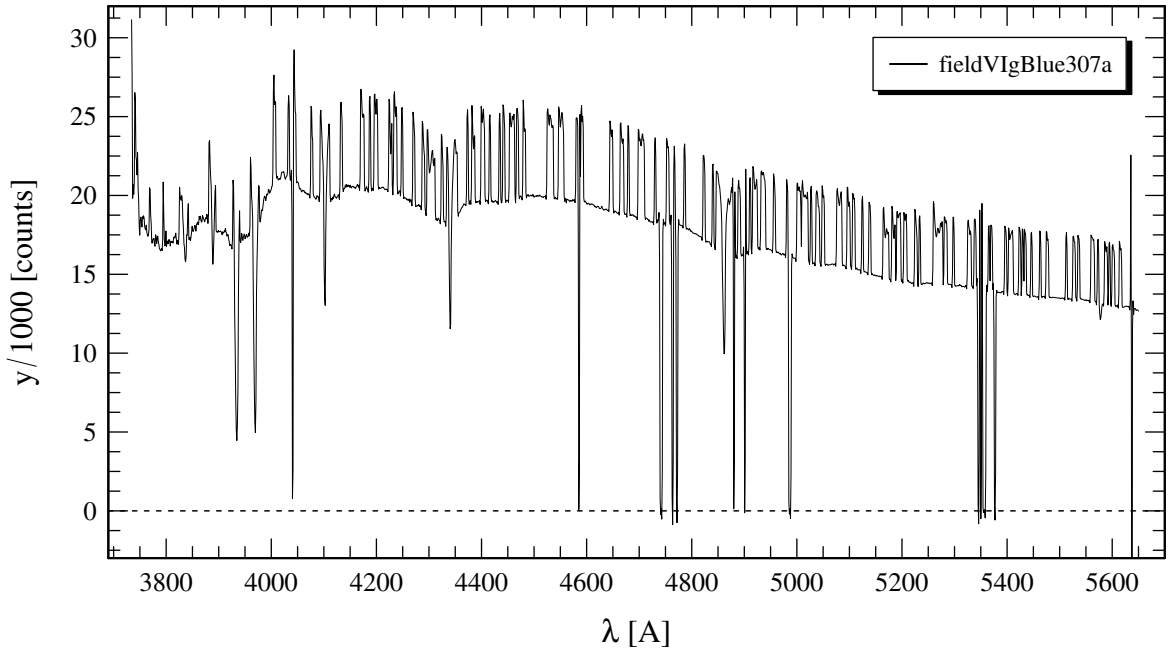

Figure 3.5: A spectrum of corot110675161 (fieldVIgBlue307a). Same as in fig. 3.4 , but here only a strong lower pseudo continuum is visible.

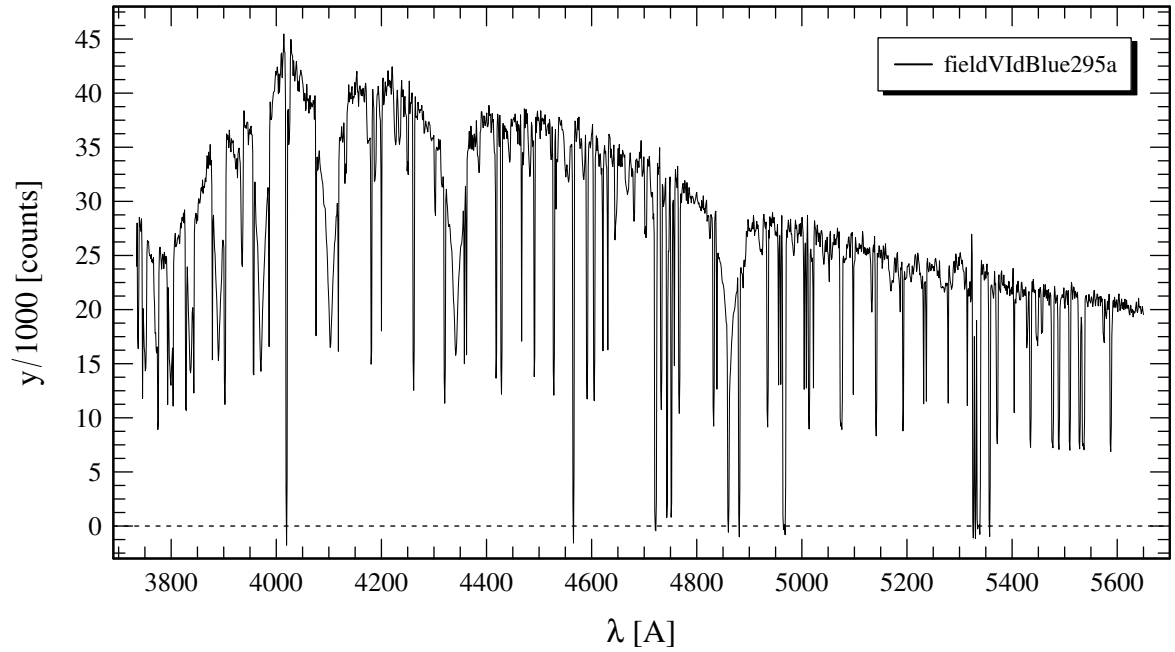

Figure 3.6: A spectrum of corot110670892 (fieldVIdBlue295a). Same as in fig. 3.4, but here only a strong upper pseudo continuum is visible.

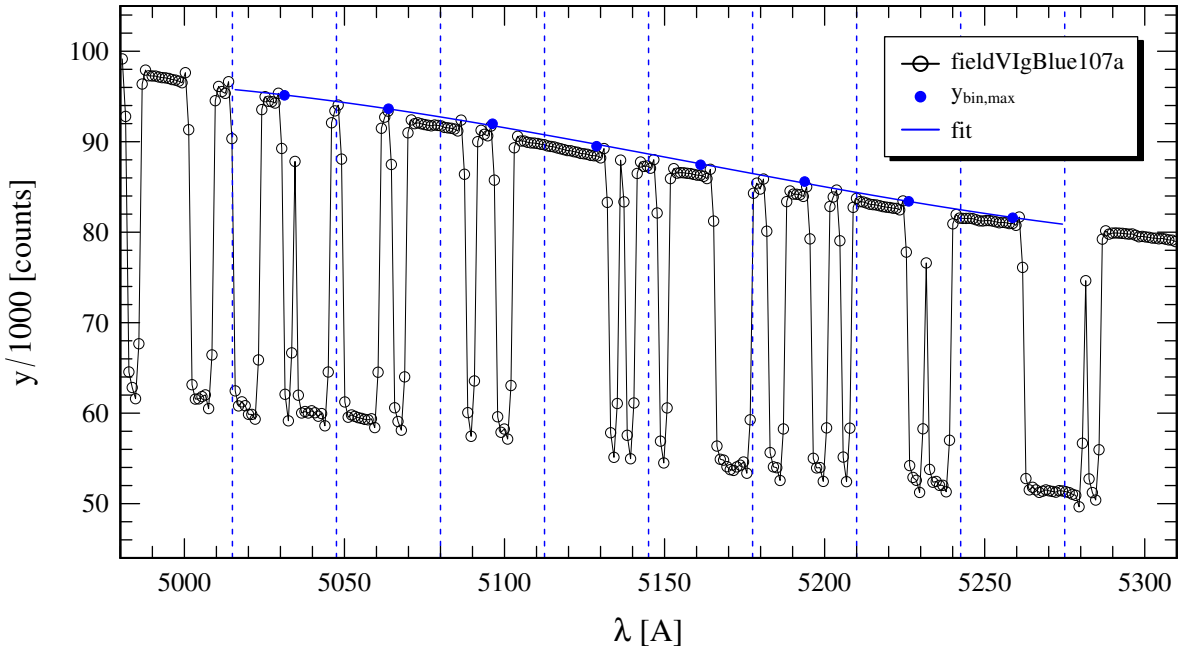

Figure 3.7: A zoom-in onto a spectrum of corot110667815, see fig. 3.4. Open circles are data points in counts/1000 from the spectrum. Vertical lines represent the boundaries of the 8 bins and the bin maxima $y_{bin,max}$ are evaluated at the center of the bins.

der (see "fit" in fig. 3.7) using a linear model via the `R` built-in routine `lm` in the provided wavelength domain. Since some parts of a rectangular-like pattern may be wider than 32.5 Å, the maximum within a bin might fall down into the valley (i.e. the lower pseudo continuum) and as a consequence the computed bin data points will not represent the upper pseudo continuum accurately (red line in fig. 3.8). This behaviour is taken into account by applying a 3σ-clipping onto the 8 maximal counts $y_{bin,max}$ twice: On the absolute values $y_{bin,max}$ and on the relative counts normalised to the polynomial $(y_{bin,max} - \mathrm{fit}_{bin,max})/\mathrm{fit}_{bin,max}$. If 1 to 4 of 8 bin maxima values are out of the 3σ range they will be linearly interpolated and afterwards a new polynomial is fitted to the new set of data points. The resulting polynomial (or the previous one, if no interpolation has taken place) is evaluated at the wavelength data points of the spectrum in the given interval, fit_{bin}, (fig. 3.7). The same domain of the spectrum is normalized with respect to

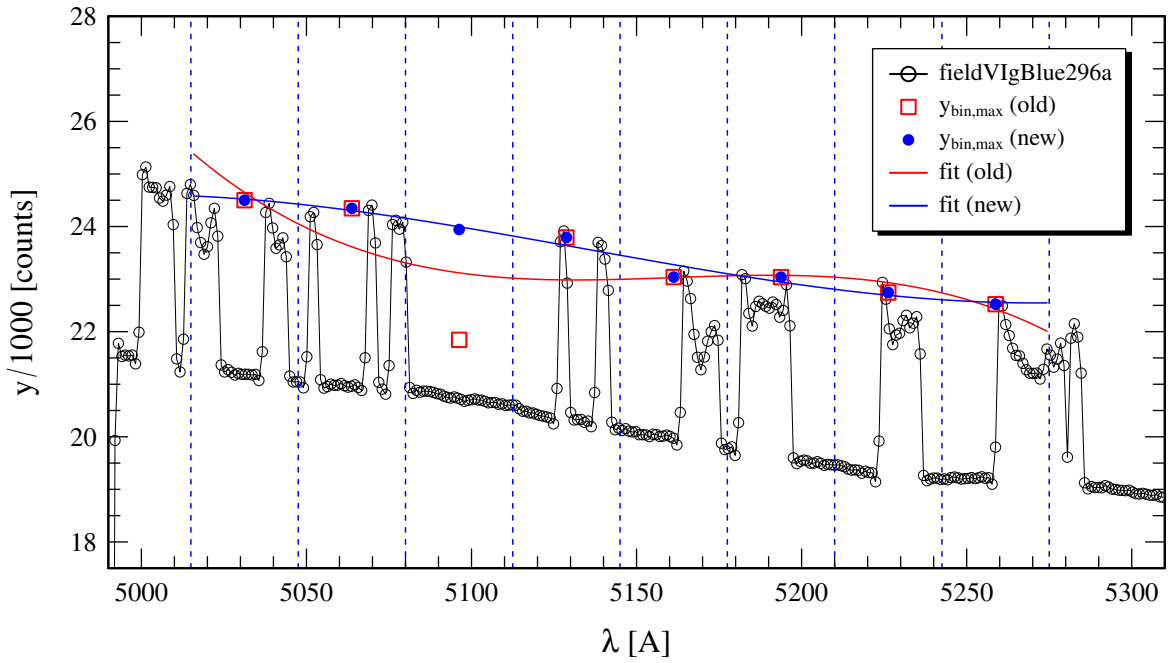

Figure 3.8: A zoom-in onto the spectrum of the star corot110755949. In the third bin all data points apart from one lie at the lower pseudo continuum, so the old $y^3_{bin,max}$ (red square) is dragged down towards the overwhelming majority of data points, resulting in a bad approximation of the upper pseudo continuum. After applying the 3σ clipping it is linearly interpolated and the new $y^3_{bin,max}$ (blue filled circle) resembles the maximum in that bin more accurate, hence the fit approximates the upper pseudo continuum much better.

this fit

$$y_{bin,norm} = y_{bin} / \mathrm{fit}_{bin} \tag{3.11}$$

The values of $y_{bin,norm}$ are divided into 5 horizontal groups/bars $N_{bin,norm}$, but to take care of outliers, the lowest 20 and highest 20 of about in total 260 values are not taken into account when computing the range and equally spaced intervals of the distribution of $y_{bin,norm}$ (see fig. 3.9).

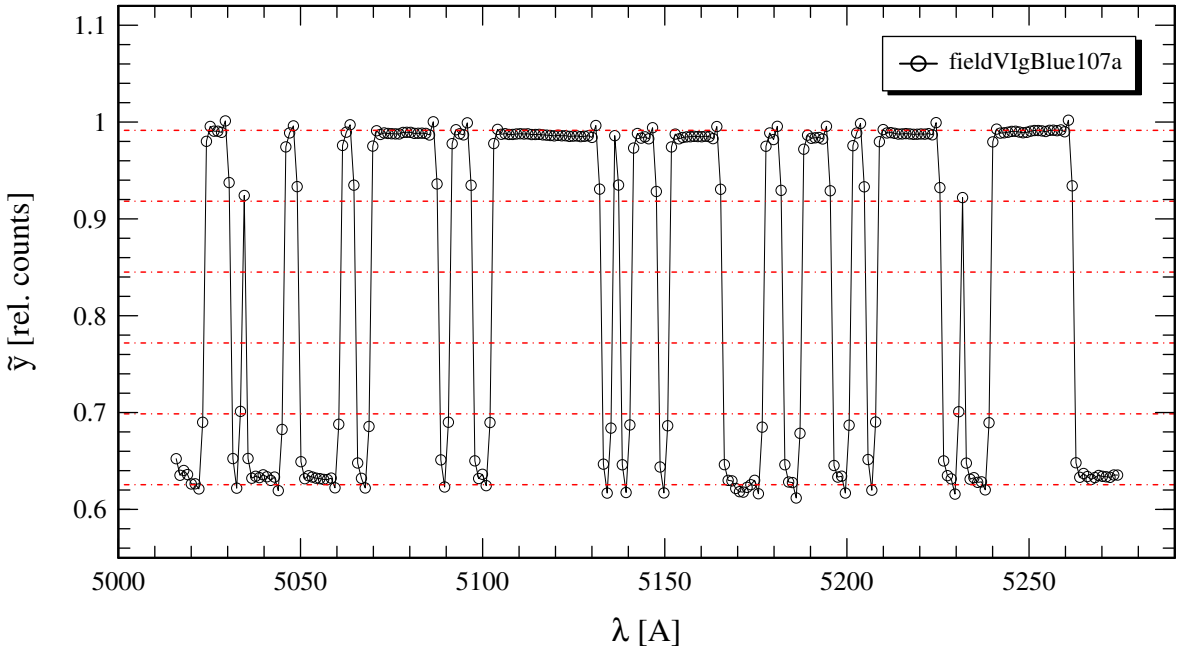

Figure 3.9: A zoom-in onto the spectrum of corot110667815. Counts have been normalized to the fit of the upper pseudo continuum from fig. 3.4. Horizontal lines represent the boundaries of the 5 groups or bars. They exclude the 20th lowest and 20th highest normalized count values to take into account statistical outliers and are equally spaced. One can see that most data points (black circles) are located in the lowest ($N^1_{bin,norm}$) or highest bar ($N^5_{bin,norm}$), while in the three bars between there are almost no data points.

Each bar holds the number of values of $y_{bin,norm}$ within its boundaries. Spectra containing rectangular-like patterns are expected to show an excess in the 1st (containing the lower pseudo continuum) and 5th bar (containing the upper pseudo continuum), while there

should be almost no values of $y_{bin,norm}$ in the three bars in between. Spectra without those patterns may follow several different rather smooth and/or uniform distributions, depending on spectral type, luminosity class, chemical abundances, etc., but they are not expected to be distributed bimodally. The spectrum of corot110667815 is compared to two spectra without rectangular-like patterns in fig. 3.10. It is essential to point out

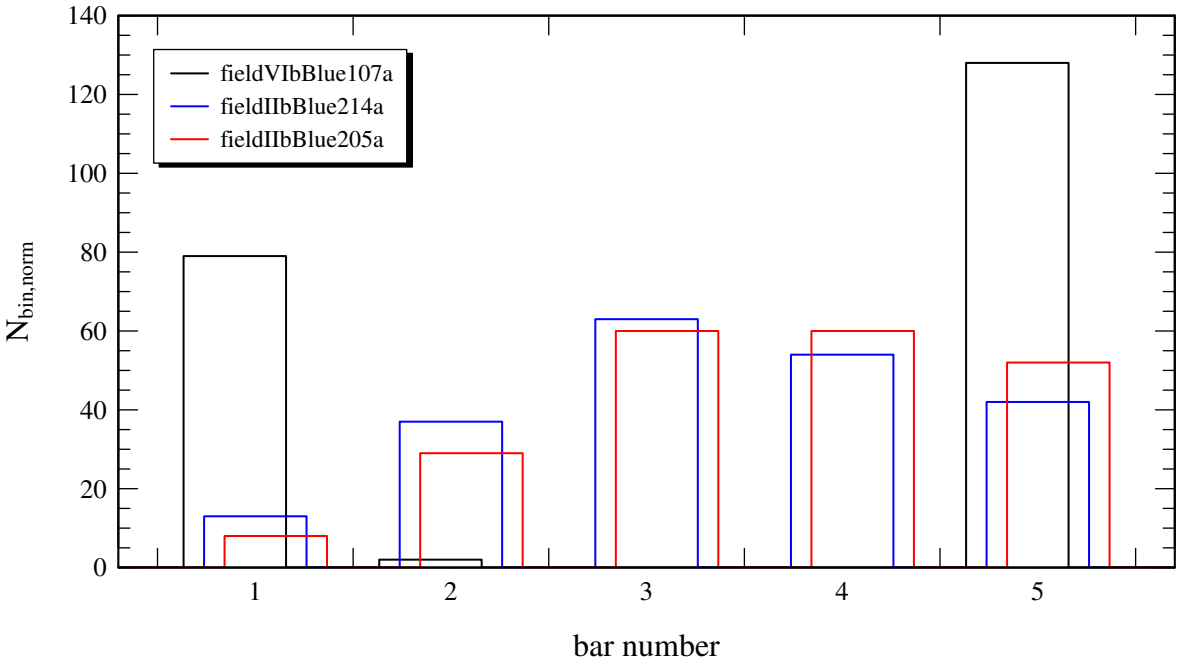

Figure 3.10: The distributions of normalized counts between 5015 Å and 5275 Å of three spectra. The black bars represent the number of data points within each red interval in fig. 3.9. The spectra of the other two distributions are plotted in fig. 3.2 and fig. 3.3, they belong to A- and G-type stars.

that this technique is not used to deduce physical parameters from the spectra. The purpose of use is to detect just one specific non-astrophysical distribution (bimodal) out of a variety of astrophysical distributions. Two simple empirical criteria are assumed to differentiate between the spectra: If the sum of number of values of the 1st and 5th bar is larger than the sum of numbers of the 3 bars in between or the number of values in the 3rd bar is larger than the sum of numbers of all other bars, then the spectrum is said to

interval name	$r_{\sigma,1}$	$r_{\sigma,2}$	$r_{\sigma,3}$	$r_{\sigma,4}$	$r_{\sigma,5}$	$r_{\sigma,6}$
λ_1	3990	4535	4690	4885	5275	5570
λ_2	4060	4610	4810	5015	5395	5650

Table 3.1: Boundaries of wavelength intervals in Å(λ_1, λ_2), where bad counts type b occur. Most of these are located near zero counts, except for $r_{\sigma,6}$, where very frequently one finds non physical peaks and troughs.

contain rectangular-like patters.

$$N^1_{bin,norm} + N^5_{bin,norm} > \sum_{i=2}^{4} N^i_{bin,norm} \quad \vee \quad N^3_{bin,norm} > \sum_{i=1, i\neq 3}^{5} N^i_{bin,norm} \tag{3.12}$$

All detections of the three samples are checked manually to verify positive detections (and discard these spectra) or to remove false detections (and keep these spectra), because the above conditions are not only sensitive to rectangular-like patterns, but also to other effects like low flux, noisy or late type spectra.

3.6 Bad counts

In the following the counts of the spectra are scaled to c_{norm} or c^*_{norm} from section 3.3, depending on whether they have been shifted according to section 3.4 and are referred to as relative counts $\tilde{y} = y/c_{norm}$ or $\tilde{y} = y^*/c^*_{norm}$. In the sample spectra two kinds of bad counts occur: (a) bad counts randomly distributed over the whole wavelength domain of the spectra, ranging from small to high negative or positive counts and (b) bad counts only within specific wavelength boundaries, which are listed table in 3.1. Either they are equal to zero (IRa1) or spread around zero (LRa1, LRa2). In case of (b) from few percent up to $\sim 50\%$ of counts in each of the ranges can be flawed.

In this part of the analysis the objective is to detect and remove bad counts within the spectra, which is done in four steps following different approaches:

1. Amount of absolute deviation of two neighbouring counts;

2. Low relative counts on a level close to zero ($\tilde{y} < 0.01$) on the whole spectrum;

3. Low relative counts on a higher level, $\tilde{y}_{cor}$, for $\lambda > 4535$ Å;

4. $\kappa\sigma$-clipping on the relative counts in the intervals listed in table 3.1;

All detected bad counts in $\tilde{y}$ will be set to `NA` ("not available") in a second variable $\tilde{y}_1$ and for the steps 1 - 3 also all next two neighbouring relative counts will be set to `NA`. Steps 2. and 4. are rather self-explanatory. Step 1. refers to the function `rm.bad.counts`, which cycles through intervals of $\tilde{y}$ listed in table 3.2 and outputs its indices, where the differences to the respective previous relative counts are considered as too large (flagged indices).

In more detail flagging indices is triggered by the condition (see (b) in fig. 3.11)

$$\left|\tilde{y}^k - \tilde{y}^{k-1}\right| > \kappa \cdot \tilde{y}_{mad} \tag{3.13}$$

with k as index of $\tilde{y}$, $\tilde{y}_{mad}$ as median absolute deviation ($\tilde{y}_{mad} = \text{mad}(\tilde{y}^{[k_5-30,k_5-1]})$ of the previous 30 relative counts, but it will be computed only every 5th k to save computational time (see (a) in fig. 3.11) and κ as input parameter, typically ranging from 6 to 9. Every time the above condition is fulfilled, the index flagging will be switched on (if it has been off) or off (if it has been on). If it is switched on, it saves the local quantile of the previous 30 relative counts, which also will be computed every 5th k step, $\tilde{y}_{save} = \tilde{y}_{med} = q_p(\tilde{y}^{[k_5-30,k_5-1]})$ and starts flagging indices until it will be switched off again. Scanning over bad counts triggers the flagging twice: When it hits a bad count and when it returns from a bad count. However it might occur that the flagging misses bad counts and will be triggered not at all or too late, which is labeled as false triggering. Two conditions prevent false triggering, the first one (see (c) in fig. 3.11) is $\left|\tilde{y}^k - \tilde{y}_{med}\right| \leq \kappa \cdot \tilde{y}_{mad}$ to check if the current relative counts level is rather close to the previous 30 ones and in case so the flagging will be switched off. The second condition implies a counter for the number of flagged indices in a row, see (d) in fig. 3.11. If it is larger than 5, flagging will be switched off unless $\left|\tilde{y}^k - \tilde{y}_{save}\right| > \kappa \cdot \tilde{y}_{mad}$, then flagging will continue for further 5 indices, as the current $\tilde{y}^k$ has not returned from the "valley" or the "plateau" of a bad count. The intervals are chosen such that $\lambda \geq 4000$ Å and no

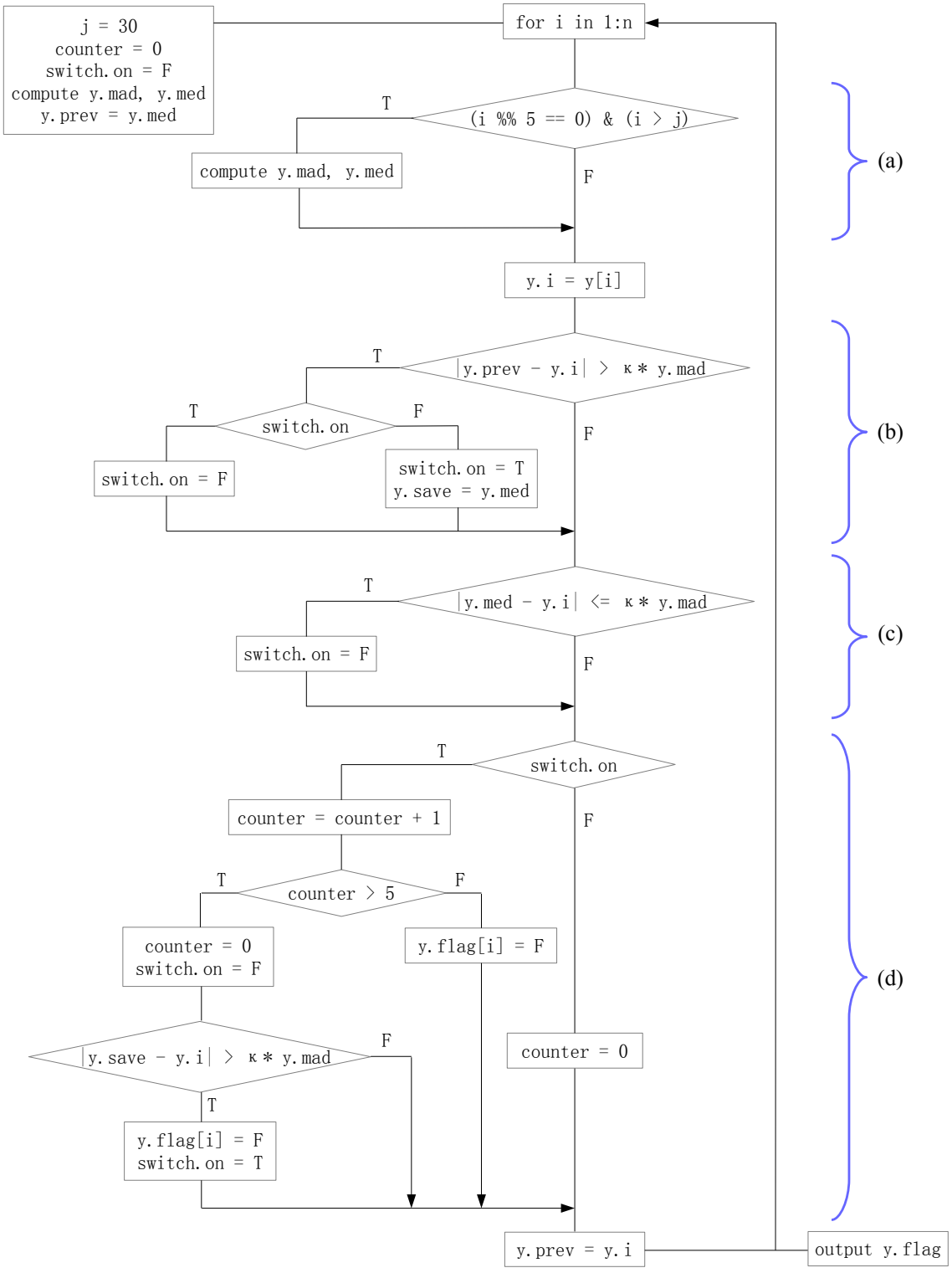

Figure 3.11: The flowchart of the core of the function `rm.bad.counts`, which is described in the text. Here the index i is used instead of k to maintain the nomenclature of the code in the flowchart.

interval name	$r_{bc,1}$	$r_{bc,2}$	$r_{bc,3}$	$r_{bc,4}$	$r_{bc,5}$
λ_1	4000.00	4142.68	4382.05	4904.32	5275.00
λ_2	4060.79	4298.89	4818.38	5275.00	5650.00
κ	6	8	7	9	7
p	0.75	0.75	0.75	0.75	0.50

Table 3.2: The first two rows show the boundaries of the wavelength intervals in Å, where `rm.bad.counts` will cycle through. In the third row the values of κ are listed and the larger this quantity is, the more outliers are tolerated. In case of hot stars, the spectra contain well defined continua with rather weak absorption features in $r_{bc,4}$ apart from strong He absorption lines. That means $\tilde{y}_{mad}$ will be rather small, hence κ has to be large or those He lines are flagged and removed from the spectra. The fourth row contains the p-value for the quantile function, if it is equal to 0.5 it the quantile becomes the median. This is necessary, because in that domain outliers in both directions, above and below the continua, occur.

H absorption lines interfere (the cut out range amounts to $\phi = \pm 30$ Å plus an additional Doppler shift corresponding to $v_{rad} = \pm 800\, km/s$ on both sides of the H line centers, $(\lambda^{\circ}_{H_{\beta,\gamma,\delta}} + \phi)/(v_{rad}/v_c + 1)$, see eq. 3.27). This function should not be applied on the $H_{\beta,\gamma,\delta}$ absorption lines, as $\tilde{y}_{mad}$ will become too large, so distinction between good or bad counts is impossible, hence there is step 3.

After 1. & 2. not all, but most of the bad pixels type (b) are known and based on their counts levels a threshold is derived to find the ones, who still remain unknown. With the indices of the current `NA`s in $\tilde{y}_1$ a sequence $\tilde{y}_{\mathtt{NA}}$ is generated, its 33%-quantile is computed and $\tilde{y}_{cor}$ is derived.

$$\tilde{y}_{cor} = \max(q_{0.33}(\tilde{y}_{\mathtt{NA}}) + 0.1, 0.2). \tag{3.14}$$

Due to the maximum function this quantity never falls below 0.2, provides an excellent cutoff limit for bad counts for $\lambda > 4535$ Å and also allows the case, where all counts (including bad counts) have been shifted to a higher level in section 3.4.

After working through all four steps the `NA`s in $\tilde{y}_1$ are counted overall and within the range of Δa photometry, $4900\,\text{Å} \leq \lambda_{\Delta a} \leq 5600\,\text{Å}$. If either the total amount of `NA`s is

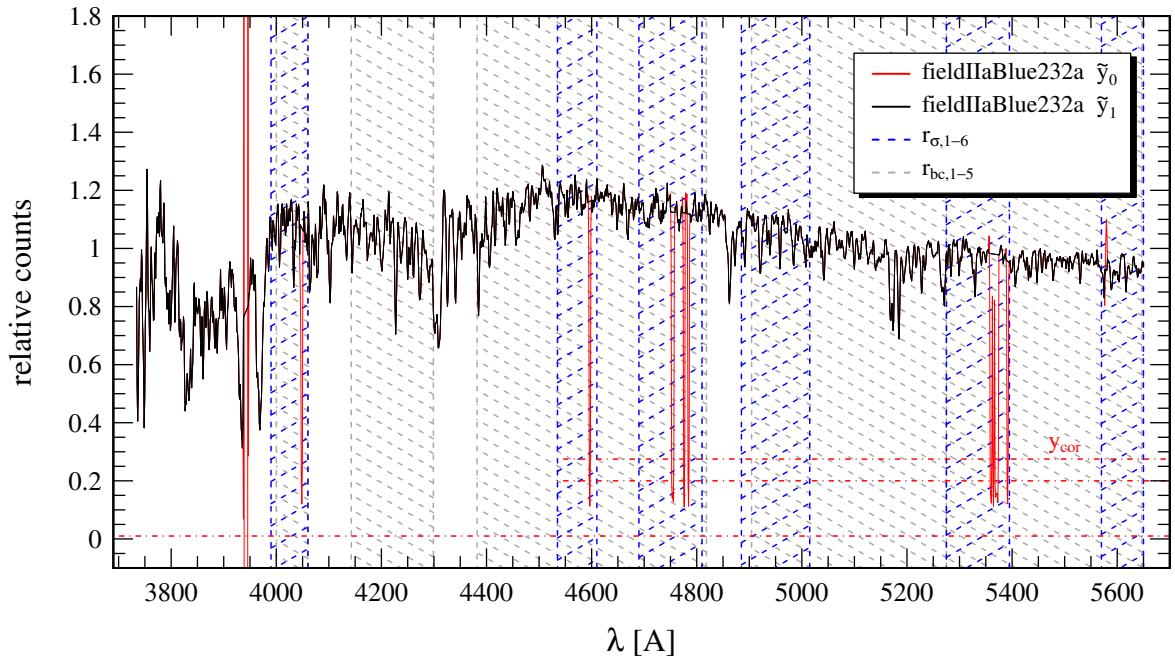

Figure 3.12: The spectrum of corot102546033 (fieldIIaBlue232a), here in relative counts, scaled to c_{norm}. The red spectrum is the uncorrected spectrum, while the black one is the corrected spectrum, where all bad counts have been removed and interpolated. The lined areas display the wavelength domains where `rm.bad.counts` or a $\kappa\sigma$-clipping will be applied. The horizontal red lines correspond to the 0.01, 0.2 and $\tilde{y}_{cor}$ limits (see text). Here $\tilde{y}_{cor}$ is larger than 0.2 due to the shifting from section 3.4, which is not the case for most spectra.

larger than 10% of the total amount of counts or the amount of NAs in $\lambda_{\Delta a}$ is larger than 10% of the amount of counts within $\lambda_{\Delta a}$, the spectrum is considered as "poor quality spectrum", because too many data points have been removed (and would have been interpolated later on).

3.7 Interpolation of bad counts

Performing a robust interpolation of NAs in the spectra is quite easy, if none of them lie in the interval $r_{\sigma,4}$ (see table 3.2), but if they do, then it gets much more complicated, depending on which parts of the H_β absorption line are affected from NAs. In the best case the far end of a line wing contains some NAs (see fig. 3.14) and even a linear interpolation will do, but as for the worst case the line center (see fig. 3.15) and sometimes the wings

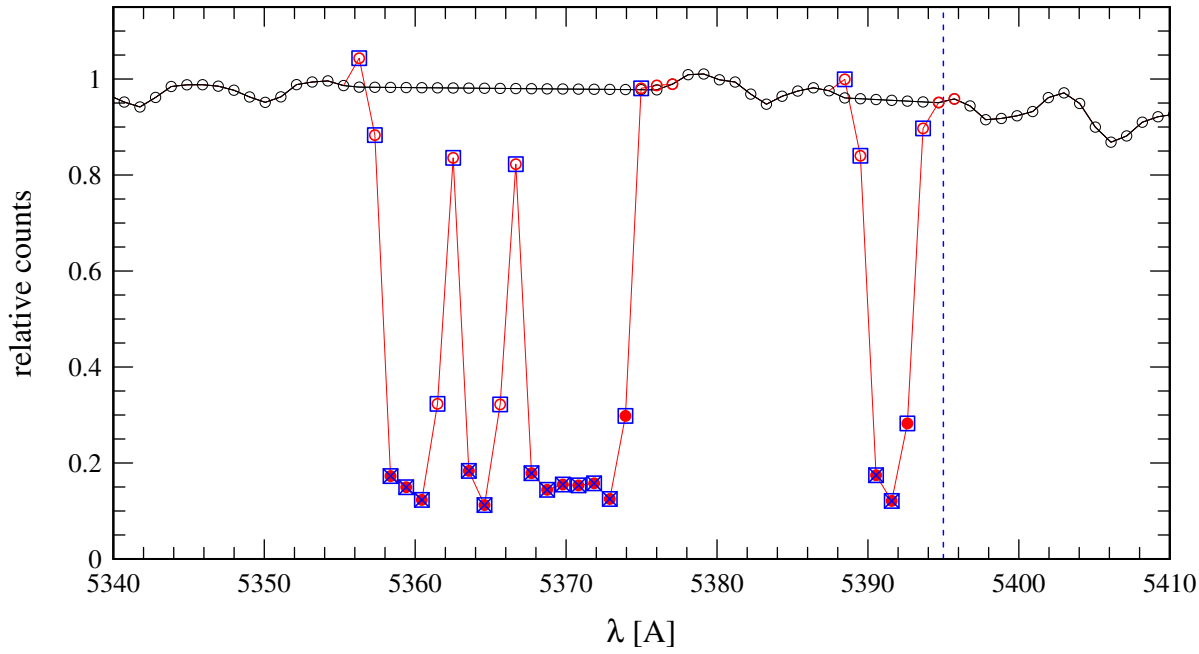

Figure 3.13: A zoom-in onto the spectrum corot102546033. Black circles are corrected the output of the whole routine. Red dots are data points cut out due to the function `rm.bad.counts` (step 1), red circles their neighbors. Blue crosses are removed data points during step 3 and blue squares indicate their neighbores. The reduction from the $\kappa\sigma$-clipping (step 4) is not shown here.

too (see fig. 3.16) can be heavily polluted by bad counts and therefore all counts are set to `NA`s. While the pollution of H_β hardly influences Δa results directly as the transmission of the g_1 and g_1' filter curves in that wavelength domain is about the order of 0.1%, it may have an impact on the accuracy of the radial velocity computation and the spectral classification process. At least the latter plays an important role for Δa results. Therefore I try to reconstruct the H_β absorption line if necessary (`NA`s in the range discovered) and if possible (range contains not too many `NA`s and reconstruction is accurate). Afterwards I apply a smooth linear interpolation on the remaining parts of the spectrum.

3.7.1 H_β absorption line interpolation

First of all it is essential to point out that the aim of reconstructing H_β is not to obtain precise line strengths, but to determine only the position of the line minima and prevent misclassification of the stellar spectral type, so the function f from equation 2.2 will do,

but the range of λ_{H_β} has to be adapted: It consists of relative counts evaluated on a rather small interval, based on the LSE_{Mg_b} covering the line center, $(\lambda^c_{H_\beta}, \tilde{y}^c_{H_\beta})$ and four more pairs of data points from the far end of the line wings. The later are 90%-quantiles computed from four intervals

$$\lambda_0 - 80\,\text{Å} \leq \lambda \leq \lambda_0 - 60\,\text{Å} \tag{3.15}$$

$$\lambda_0 - 60\,\text{Å} \leq \lambda \leq \lambda_0 - 40\,\text{Å} \tag{3.16}$$

$$\lambda_0 + 40\,\text{Å} \leq \lambda \leq \lambda_0 + 60\,\text{Å} \tag{3.17}$$

$$\lambda_0 + 60\,\text{Å} \leq \lambda \leq \lambda_0 + 80\,\text{Å} \tag{3.18}$$

with $\lambda_0 = 4861.35$ and evaluated at their interval centers respectively. Merging $(\lambda^c_{H_\beta}, \tilde{y}^c_{H_\beta})$ with these 4 pairs of data points yields $(\lambda_{H_\beta}, \tilde{y}_{H_\beta})$.
The method itself is described shortly in the following:

1. Fit function f from eq. 2.2 to the H_β absorption line.
2. If the radial velocity derived from the fit's line minimum is not in good agreement with the expected radial velocity or the fit itself is not good enough (i.e. if largely differs from its initial parameters), a natural spline will be used to interpolate the missing values. Again f will be fitted to the new set of data points.
3. If the radial velocity derived from the new fit's line minimum is not in good agreement with the expected radial velocity or the new fit itself is not good enough (like in the previous step), a linear interpolation is used to fill up the `NA`s of the H_β line.

(1) Regarding spectral line fitting in section 2.2 the challenging part here is to find a set of four reasonable initial parameters, because due to the fact that there might several `NA`s in that range, only the fourth parameter, d, can be determined solely from the data points of the spectrum $(\lambda_{H_\beta}, \tilde{y}_{H_\beta})$, namely $d = 1.05 \cdot q_{0.95}(\tilde{y}_{H_\beta})$. Especially $\hat{\lambda}$, has to be derived from other spectral features (explained in section 3.10.2) and not from the data points within H_β. σ is deduced from the mean of the standard deviations of the H_γ and H_δ absorption lines,

$$\sigma = 1.25 \cdot (\sigma_{H_\gamma} + \sigma_{H_\delta})/2 \tag{3.19}$$

and a small scaling factor of 1.25 is added, because H_β used to be broader than the other two lines. To obtain the parameter k I use eq. 2.14 and the LSE of the H_γ and H_δ absorption lines. Once all four parameters $(\hat{\lambda}, \sigma, h, d)$ are determined (some lower and upper boundaries are considered) `optim` tries to find a minimum and returns the best-fit parameters $(\hat{\lambda}_{min}, \sigma_{min}, h_{min}, d_{min})$ to evaluate $y_{Gauss} = f(\lambda_{H_\beta}|\hat{\lambda}_{min}, \sigma_{min}, h_{min}, d_{min})$ and replace the `NA`s in $\tilde{y}_{H_\beta}$ by y_{Gauss}.

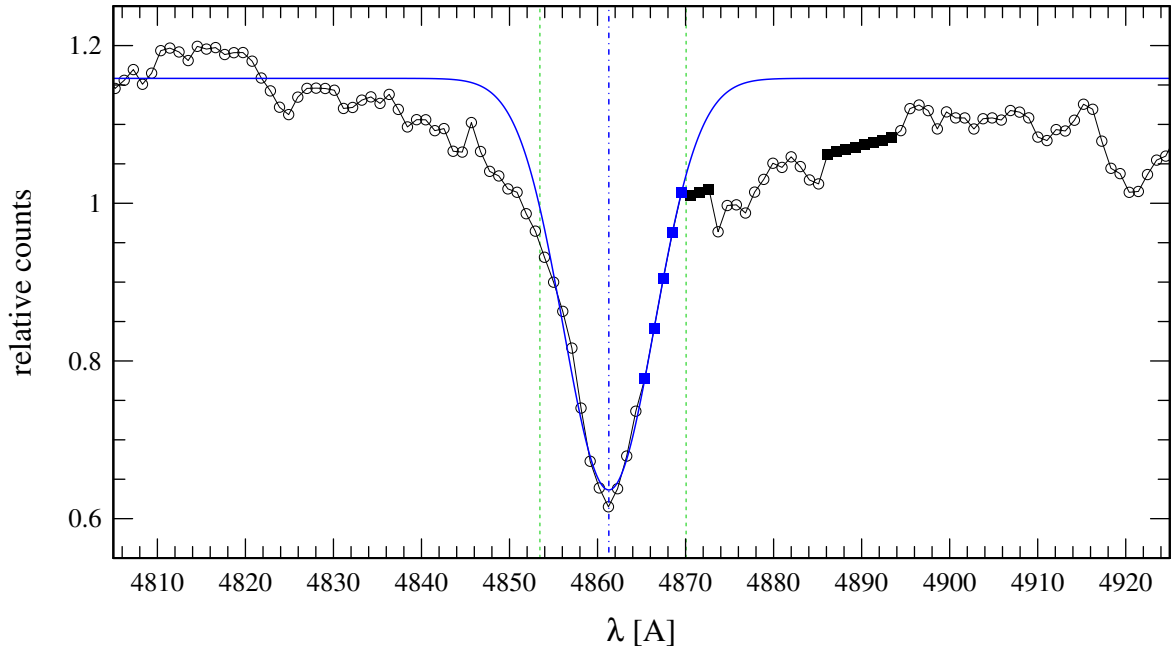

Figure 3.14: A zoom in onto the H_β absorption line of a spectrum. Black circles represent the data points $(\lambda_{H_\beta}, \tilde{y}_{H_\beta})$ at hand, black filled squares are interpolated data points after the H_β has been reconstructed according to section 3.7.2 and green vertical lines signal the range of data points of the line center $(\lambda^c_{H_\beta}, \tilde{y}^c_{H_\beta})$, on which the fit will be acting. f with the best-fit parameters is plotted as blue curve and blue squares are the points, where the corresponding curve has been evaluated. As one can see f does not resemble the shape of the H_β perfectly and there is a considerable deviation in the wings, but the position of the minimum can be fitted quite well.

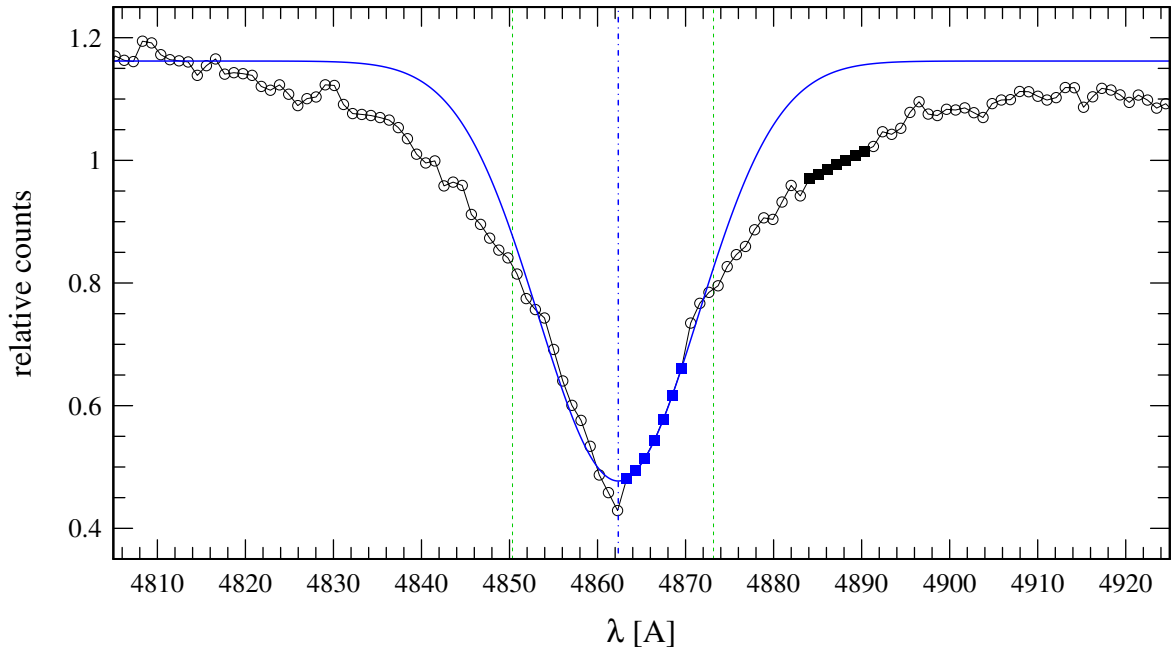

Figure 3.15: Same as in fig. 3.14 for a different star, but here the actual line minimum of $(\lambda_{H_\beta}, \tilde{y}_{H_\beta})$ might be missing or be affected by bad counts, since the lowest data point around 4861 Å is also the last before a sequence of missing data points occurs. The ordinary least squares minimzation forces the fit up, further away from this data point.

Now the question is what is the required accuracy for the output parameter? In first place a variety of criteria is employed to verify, if the fit generally is good enough:

- $\min(\tilde{y}_{H_\beta}) > y_{cor}$
  No absorption lines with data points that fall below y_{cor} are allowed, see section 3.7.

- $0.5 \cdot \Delta\lambda_s < \sigma_{min} < 10\,\text{Å}$
  This criterion sets the range for the standard deviation of the fit.

- $|\sigma - \sigma_{min}| < 2\,\text{Å}$
  The initial and the best-fit standard deviation shall not differ too much.

- $h_{min} > 0$
  Only absorption lines are considered, $k_{min} < 0$ would produce emission lines.

- $\frac{|d_{min} \cdot h_{min}|}{\sigma\sqrt{2\pi}} > 0.1$
  The fitted spectral line has to reach at least a certain minimum depth (see equation

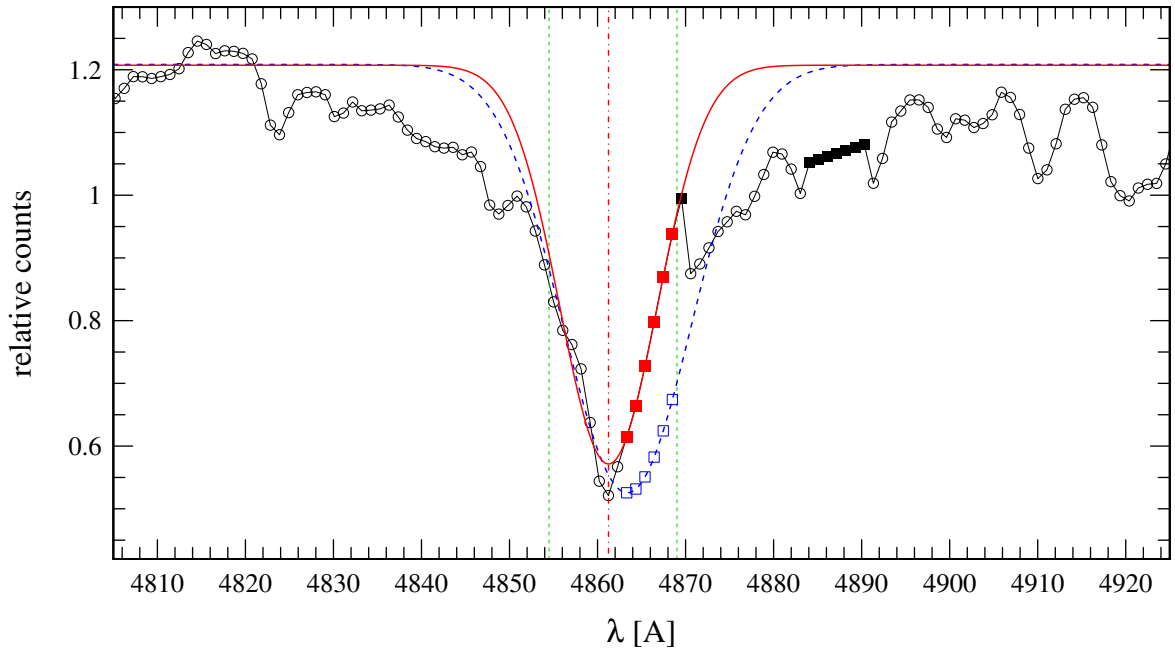

Figure 3.16: Same as in fig. 3.15, but this time the sequence of missing data points extends the right boundary, which means there will be a "loose end" for the fit on the right part of the H_β line. Luckily the line minimum appears not to be affected by bad counts, but still the original fit has difficulties at reconstructing this spectral line and it leads to the dashed blue curve. The wavelength position of this fit is too far away from the expected wavelength, so a new attempt of resembling the H_β line is done. Now the missing data points are interpolated by a spline, so the old set of $(\lambda_{H_\beta}, \tilde{y}_{H_\beta})$ plus the newly interpolated data points are the basis for the new fit leading to the red curve and red squares as results of the fitting.

2.6) and it ensures that no noisy feature is fitted.

If all criteria are fulfilled, in the next step $\hat{\lambda}$, the expected wavelength for the H_β line and $\hat{\lambda}_{min}$ are converted into radial velocities $(v_{rad}, v_{rad,min})$ via eq. 3.26. The separation of counts on the wavelength scale accounts to about 1 Å, which corresponds to a Doppler velocity of $\sim 60\,km/s$ for the H_β line. This is the limit of accuracy that should not be surpassed. For $|v_{rad,min} - v_{rad}| \leq 50\,km/s$ the fit f_0 is accepted (for $\leq 30\,km/s$ the quality flag will be "G1", for $> 30\,km/s$ it will be "G2"), (2) else the fit is rejected and a natural spline will be considered as next interpolation scheme for the `NA`s. If at least one of the above criteria is not fulfilled, then also a spline will be used for replacing `NA`s. Now once

again f and `optim` are used to find best-fit parameters in a similar way, but this time the original set of data points to fit is combined with the spline interpolation. Again the same set of criteria from the list above have to be fulfilled and if this is the case, a new $v_{rad,min}$ is computed. For $|v_{rad,min} - v_{rad}| \leq 100\,km/s$ the new fit f_1 is accepted (for $\leq 50\,km/s$ the quality flag will be "S1", for $50\,km/s < |v_{rad,min} - v| \leq 100\,km/s$ it will be "S2" and else "S?"). (3) If neither f_0 nor f_1 are accepted, a simple linear interpolation is performed (with flag "L?").

3.7.2 Smooth linear interpolation

Once the H_β absorption line interpolation is finished, all remaining `NA`s are interpolated by a robust linear interpolation. If $\tilde{y}^k$ is the last non-`NA`, before a sequence of m `NA`s occur and if $\tilde{y}^{k+m+1}$ is the first non-`NA` after the same sequence, then the sets of data points $(\lambda^{[(k-20):k]}, \tilde{y}^{[(k-20):k]})$ and $(\lambda^{[(k+m+1):(k+m+21)]}, \tilde{y}^{[(k+m+1):(k+m+21)]})$ are used to compute the left and right average counts interpolation points.

$$\lambda_{left} = \text{median}\left(\lambda^{[(k-20):k]}\right) \tag{3.20}$$

$$\lambda_{right} = \text{median}\left(\lambda^{[(k+m+1):(k+m+21)]}\right) \tag{3.21}$$

$$\tilde{y}_{left} = \text{q}_{0.75}\left(\tilde{y}^{[(k-20):k]}\right) \tag{3.22}$$

$$\tilde{y}_{right} = \text{q}_{0.75}\left(\tilde{y}^{[(k+m+1):(k+m+21)]}\right) \tag{3.23}$$

If there are `NA`s in the extended ranges $[(k-20) : k]$ or $[(k+m+1) : (k+m+21)]$, they will be removed from $\tilde{y}_{left}, \tilde{y}_{right}$, but not from $\lambda_{left}, \lambda_{right}$. Finally a linear interpolation of $(\lambda_{left}, \tilde{y}_{left})$ and $(\lambda_{right}, \tilde{y}_{right})$ is done and evaluated at the wavelengths with `NA`.

3.8 Wave-like patterns

Similar to the rectangular-like patters from section 3.5, a few hundreds of spectra, do contain wave-like patterns, ranging over significant parts of the whole wavelength domain. Sebastian et al. (2012) refer to it as "wave-like fringe pattern" and explain that it results from an instrumental problem:

> *"In AAOmega starlight enters through a small rectangular prism, which then reflects the light into the fiber [...]. This prism is normally glued to the fiber entrance, but in some cases, the prism is not perfectly glued to the fiber entrance. Then there is a little gap between the fiber and the prism, th resulting "Newtonian rings" then cause the wave-like patter in these spectra."* - p. 4

Not only that these pattern influence the spectral classification process, they also have a huge impact on Δa results, because - depending on the wavelength and phase of the pattern - they might reproduce Δa-like flux depressions (or flux excesses) around 5200 Å. Hence, spectra containing such wave-like patterns are detected and removed from the sample of spectra. Two examples of spectra with wave-like patterns are illustrated in fig. 3.17 and fig. 3.18. A Fourier analysis did not produce meaningful results, so after studying hundreds of such spectra I decided to fit the wave pattern via the heuristically determined function v

$$v(\lambda|A,\nu,\varphi,h,d,\alpha) = A \cdot \sin\left(\nu \cdot (\lambda-\varphi) \cdot \left(\frac{\lambda}{5000}\right)^{\alpha}\right) + h \cdot (\lambda-\varphi) + d \tag{3.24}$$

The right hand side consists of a wave part, a linear term and an offset. The linear term and the offset approximate the continuum, however any non-linearity in the continuum affects the wave part. Therefore the range of the spectra for this analysis has to be chosen such that no Balmer jump or strong absorption lines are included and it should not exceed an interval range of $\sim$ 1000 Å. The exact range *var* from 4600 Å to 5640 Å, but three subdomains are cut out: the H_β line (4800 Å $\leq \lambda \leq$ 4940 Å), the Mg triplet (5160 Å $\leq \lambda \leq$ 5200 Å) and another domain in (5255 Å $\leq \lambda \leq$ 5280 Å). Even though the whole range extends to 1040 Å, only a total domain of 840 Å, corresponding to a bit less than 840 data points $(\lambda_{var}, \tilde{y}_{var})$, is effectively used. The wave part has an amplitude A, an angular wavenumber ν, a phase φ and a wavelength variation α. The latter is necessary, because a small wavelength dependency of the wavenumber occurs and to maintain constant ν I introduce this potential law approach to model this wavelength or wavenumber dependency.

Similar to the fitting technique of section 2.2 and equation 2.4, an accurate set of initial parameters has to be found $(A,\nu,\varphi,h,d,\alpha)$ for v. The initial α is set to -1.2 as the majority

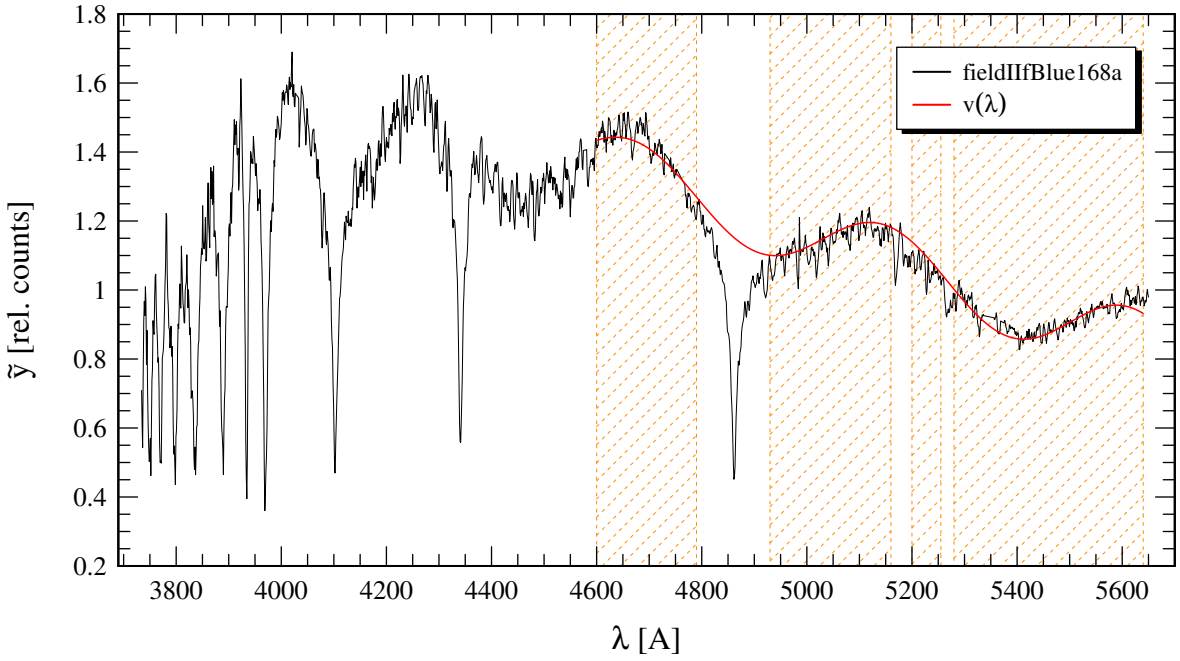

Figure 3.17: A spectrum of corot102565338. The vertical red lines define the domain of the fit, the red curve represents the fit of v with the following parameters: $A = 0.102$, $\nu = 0.01314$ or $\lambda_{pattern} = 478\,\text{Å}$, $\phi = 5030\,\text{Å}$, $h = -0.0005$, $d = 1.15$, $\alpha = 0.762$. Please note that the H_β absorption line and two more domains of the spectrum have been cut out for this analysis.

of the wave-like patterns decrease in wavenumber by increasing wavelength. h and d are determined by a linear regression analysis and are used to derive the initial amplitude by normalizing the spectrum to the continuum and computing

$$A = 1.5 \cdot (q_{0.75}(\tilde{y}_{var,norm}) - q_{0.25}(\tilde{y}_{var,norm}))/2 \tag{3.25}$$

This formula computes the mean of a "rather-peak" and a "rather-valley" of a wave and scales it by 1.5 to extend it to the peak and the valley. To obtain ν and φ I set up a mesh grid of two sequences of wavenumbers and phases $(\nu_{grid}, \varphi_{grid})$ ranging from 50 Å to 1040 Å with a step size of 10 Å for pattern wavelengths λ_{grid} (corresponding to grid wavenumber, $\nu_{grid} = 2\pi/\lambda_{grid}$, from 0.3 to 5 with non-equidistant interval points), and from 5000 Å to 5600 Å in step sizes of 20 Å for grid phases. Therefore the mesh grid is a 100×61 matrix and for each combination of ν_{grid} and φ_{grid} (as well as the other four already known initial parameters, which are held constant here) the root mean square

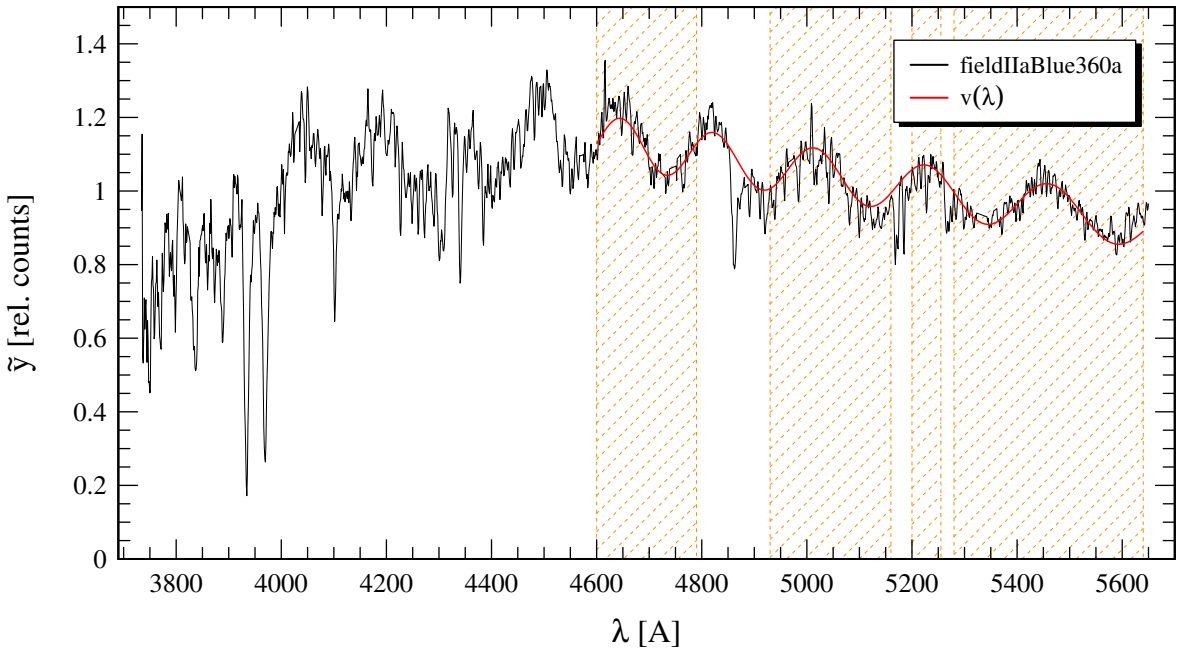

Figure 3.18: A spectrum of corot102757205. The vertical red lines define the domain of the fit, the red curve represents the fit of v with the following parameters: $A = 0.068$, $\nu = 0.02864$ or $\lambda_{pattern} = 219$ Å, $\phi = 5400$ Å, $h = -0.0002$, $d = 0.96$, $\alpha = -1.189$.

(rms) of $(\tilde{y}_{var}, v(\lambda_{var}|A, \nu_{grid}, \varphi_{grid}, h, d, \alpha))$ with the given set of parameters is computed. The minimum of all rms entries in the mesh grid provides a combination of the initial values of ν and φ. Finally this set of 6 initial parameters $(A, \nu, \varphi, h, d, \alpha)$ will be used for the optimization via `optim` to derive the best-fit parameters of v. This is done for all spectra per sample and the results are plotted in figures 3.19 to 3.21. The scope of this analysis is to detect wave-like patterns, not to analyse them further and therefore only 2 out of 6 parameters are essential: A and ν (or pattern wavelength instead of ν). φ, h, d and α are only used to provide accurate A and ν values. This is why only the A and ν are plotted and used as criteria for wave-like patterns (and also high absolute deviations between data and the fit).

It is worth noticing that there are several limitations to this algorithm:

- For hotter stars such patterns can be detected quite well, but for cooler stars it most likely fails due to strong molecular bands and those spectral features are mistaken for wave-like patterns. To avoid a false detection I set a boundary in terms of $\mathrm{LSE}_{Mg_b} < 0.25$ to exclude cooler stars from this analysis, so all B- and A-type stars

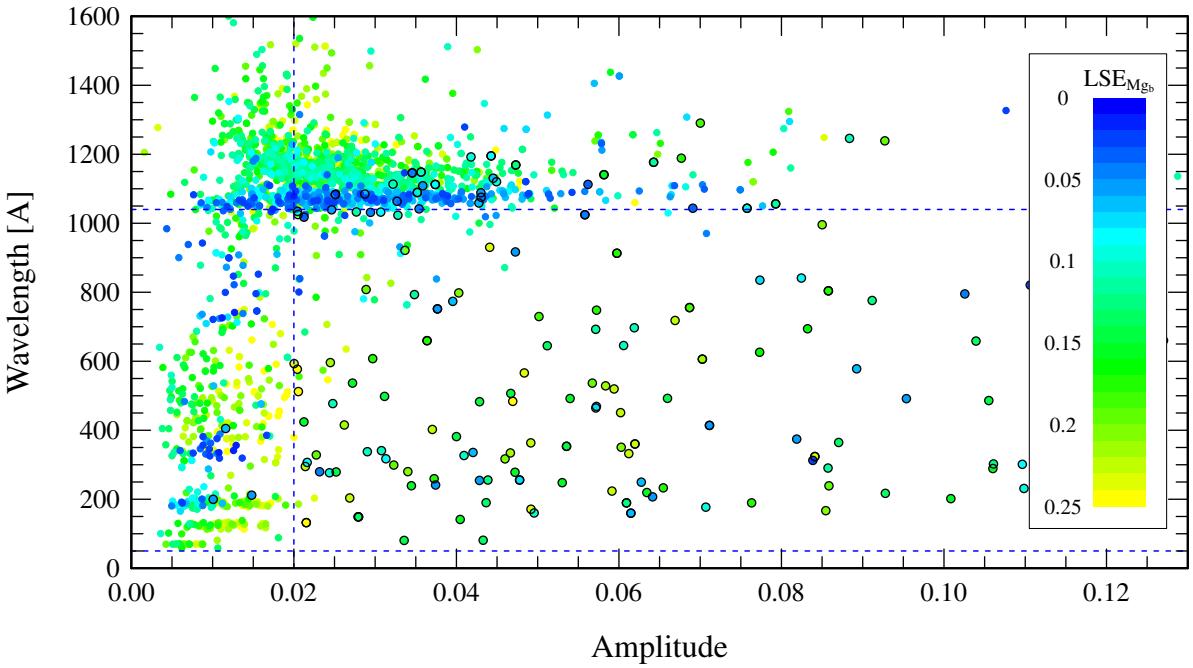

Figure 3.19: Best-fit amplitudes against best-fit pattern wavelengths of v for IRa1 spectra. A few points are located beyond 1600 Å in pattern wavelengths or 0.13 in amplitudes. The dashed lines represent the limits for the detection, black circles are actual detected spectra. The color scaling displays the LSE$_{Mg_b}$, a rough 1D spectral classification number (see section 2.1), with blue as hot stars (about early B-type stars) and yellow as cool stars (about late F). Above the upper limit for pattern wavelengths there is massive crowding due to the fact that the detection range is limited to 1040 Å. For wave patterns with pattern wavelengths larger than 1040 Å either the wave peak or the wave valley lie out of the interval, hence underestimating the initial amplitudes in equation 3.25. The optimization compensates for it and since it is no longer objected to fitting v to both peak and valley, the wave amplitude becomes an unbound free parameter and this leads to many more minima of the ordinary least squares minimization, hence the vast spread in wave amplitudes.

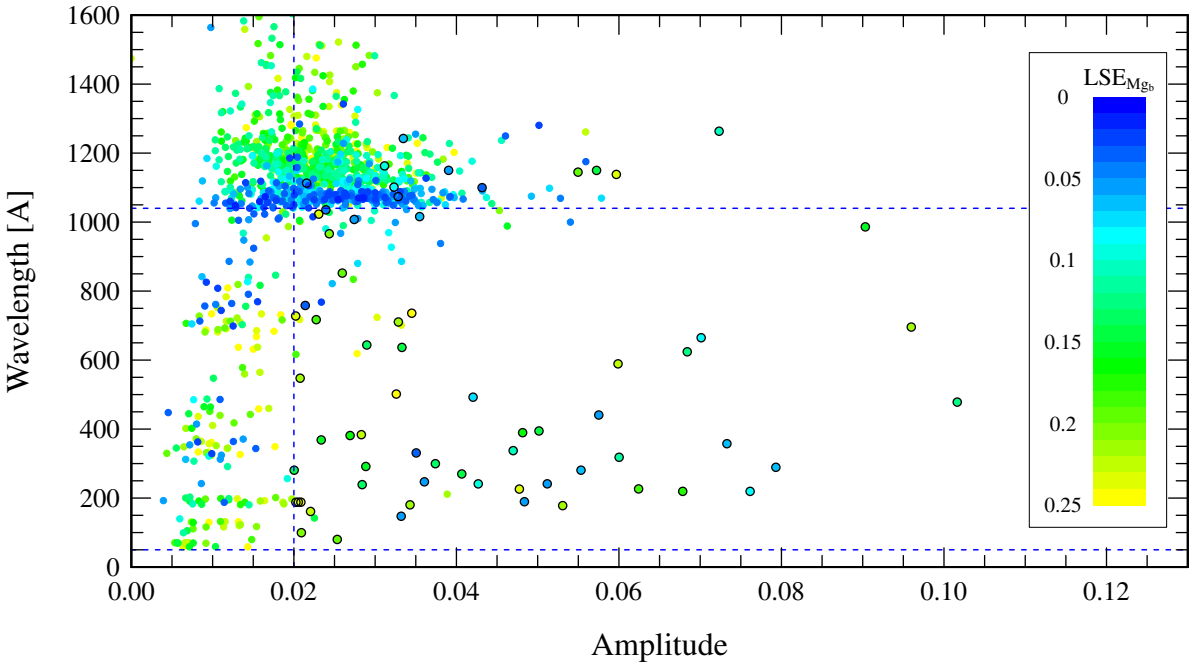

Figure 3.20: Best-fit amplitudes against best-fit wavelengths of v for LRa1 spectra, same as in fig. 3.19.

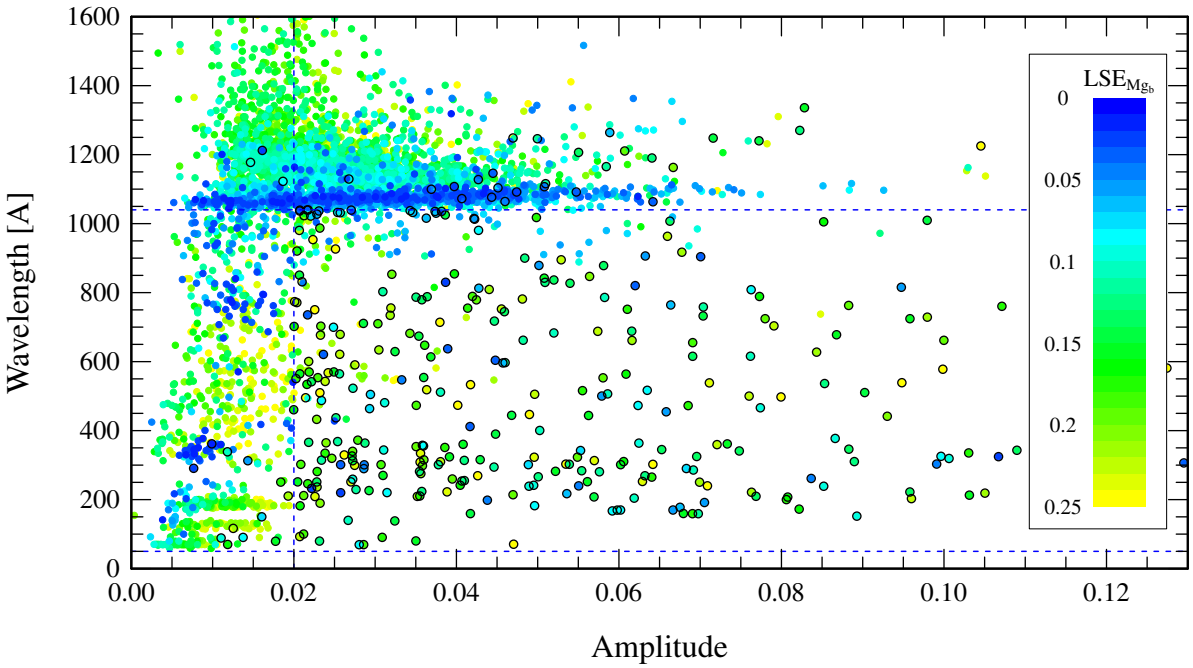

Figure 3.21: Best-fit amplitudes against best-fit wavelengths of v for LRa2 spectra, same as in fig. 3.19.

are analyzed, but some F-type stars might not.

- The use of this technique in terms of pattern wavelengths ($360/\nu$) is restricted: The upper pattern wavelength boundary is equal to the range of the interval, namely 1040 Å,. For higher pattern wavelengths it is almost impossible to distinguish between wave-like and non wave-like patterns (i.e. the slope of the continuum), see figure 3.19 for the explanation. As for the lower pattern wavelength boundary I choose 50 Å, simply because for lower wavelengths individual spectral features dominate the data points and might be interpretated as wave-like pattern.
- When it comes to amplitudes one also has to define at least a lower limit, because at some point the pattern will be suppressed by the line strengths of the spectral lines or by noise. Therefore I chose the lower amplitude limit as 0.02 in terms of relative counts for wavelength patterns smaller than 1040 Å and if the amplitude falls below, the spectrum will not be analyzed any further.

Because of these limitations some wave-like spectra will slip through, but they most likely will be detected by Δa photometry and can be removed afterwards easily. However they might influence the spread in Δa results and broaden the scatter around the normality line, depending on the amount of such spectra and the amplitudes of their wave-like patterns. Those spectra, where the algorithm reports to have detected wave-like patterns and many more spectra (from all spectra in case of LRa1 to several hundred as for LRa2), where no such patterns have been found, will be checked manually.

3.9 Multiple spectra per star

Sebastian et al. (2012) have spectroscopically observed hundreds of stars multiple times and multiple spectra are part of the samples too. Especially the LRa2 sample contains a large number of multiple spectra (almost half of the remaining sample) from 2 to even up to 7 spectra per star. After removing spectra of low quality from the sample and interpolating bad counts, the multiple spectra are now removed from the samples and obtain one spectrum per star. This is done by empirically setting a constant count levels of 13000 counts and picking the one spectrum out of all remaining multiple spectra per

star with c_{norm} closest to the corresponding level or - in case of all $c_{norm} > 13000$ - picking the spectrum with the lowest c_{norm}.

3.10 Radial velocities

3.10.1 Application of the Doppler effect

Radial velocities of stars are obtained by identifying spectral lines l in the spectra and measuring the wavelength shift of the observed lines λ_l to the corresponding rest frame line wavelengths λ_l°. Those wavelength shifts $\Delta\lambda = \lambda_l - \lambda_l^\circ$ are converted into radial velocities $v_{l,rad}$ via the (non relativistic) Doppler effect

$$\frac{\lambda_l - \lambda_{\circ,l}}{\lambda_l} = \frac{\Delta\lambda_l}{\lambda_l} = \frac{v_{l,rad}}{v_c} \tag{3.26}$$

$$\lambda_{\circ,l} = \lambda_l \cdot \left(1 - \frac{v_{l,rad}}{v_c}\right) \tag{3.27}$$

Radial velocities $v_{l,rad}$ are calculated for all lines l and a weighted mean radial velocity, v_{rad}, is computed. Finally v_{rad} is converted back into radial-velocity-corrected wavelengths $\lambda_\circ$ (see equation 3.27 without the index l due to the weigthed mean radial velocity) and used as new wavelength scale for the corresponding spectrum.

3.10.2 Spectral line list

The question arises, which spectral features can be picked for measuring the positions of line minima λ_l? Obviously this choice mainly depends on the spectral type, the luminosity class, the spectral resolution, instrumental limitations and observational effects.
As we have already discussed, LSE_{Mg_b} is a rough estimate of the stellar spectral type. It can be seen from figure 2.1 that each spectral line has a certain range of LSE_{Mg_b}, where it most likely is strong enough to allow a reasonable fit of f (see eq. 2.2) to the corresponding data points and to obtain the wavelength position of the spectral line minimum. Those ranges in terms of LSE_{Mg_b} for each spectral line are listed in table 3.4. The spectral resolution contributes to $R = 1300$ for all field samples, which is about today's common

low stellar classification resolution. While within the frame of this Master's thesis instrumental limitation is not an issue and signals are high enough for most of the spectra, observational effects do cause serious difficulties here: Bad counts type b from section 3.7 do not allow to pick spectral lines from certain wavelength domains, because it can not be determined whether such features are of astrophysical or instrumental origin. So the list of useful spectral lines is rather short, but once again, the aim of this Master's Thesis is not to derive precise radial velocities of the sample stars, but to perform numerical broad band photometry, so an accuracy of less than $50\,km/s$ is sufficient and this is what I am going to achieve.

3.10.3 Correcting for radial velocities

The overall procedure to correct for radial velocities is described below (see the flowchart in figures 3.22 and 3.23 with the corresponding steps labeled on the right side):

1. Pick few very strong absorption lines according to LSE_{Mg_b} and add them to the list of radial velocity estimators. This list contains lines which always should be dominant and easily measurable within the given range of LSE_{Mg_b}. Similar to table 3.4 these criteria are formulated in table 3.3.

2. Compute radial velocities for the lines on that list via the method described in 2.2, their weigthed estimated mean $v_{rad,est}$ and the standard deviation $\sigma_{v_{rad,est}}$. $v_{rad,est}$ serves as estimate of the true radial velocity and as initial radial velocity for the next steps. If the standard deviation is larger than $200\,km/s$ then no radial velocities are computed in the next steps. If the $|v_{rad,est}| < 150\,km/s$, the initial radial velocity will be set to $0\,km/s$. If it is larger, half of the $v_{rad,est}$ is assumed as initial radial velocity for the actual computation of radial velocities.

3. Pick absorption lines from table 3.4 depending on the LSE_{Mg_b} to form an individual spectral line list and check, if the H_β line can be added (the H_β flag has to be equal to "G1", "G2" or "S1", see section 3.7.1).

4. Launch an outer loop with three cycles, where either the initial radial velocity (1st cycle) or the one from the previous cycle (2nd or 3rd cycle) is used to compute

spectral line	$\lambda\,[A]$	LSE_{Mg_b} ($^{\text{from}}_{\text{to}}$)	0.00 0.10	0.10 0.20	0.20 0.40
H_ζ	3889.064		•		
H_ϵ	3970.075		•		
H_δ	4101.734		•	•	
H_γ	4340.472		•	•	
H_β	4861.35		(1)	(1)	
Ca	3933.66			•	
Ca	3968.47			•	•
Ca	4226.73				•
Fe	4383.545				•

Table 3.3: The complete spectral line list used for estimating $v_{rad,est}$ in dependence of LSE_{Mg_b} (see fig. 2.1). Black filled circles indicate for which LSE_{Mg_b} values a spectral line can be selected. For instance for a spectrum with $0.20 < \mathrm{LSE}_{Mg_b} < 0.40$ the last two Ca lines and the Fe line are considered. This list is chosen in such a way that their lines are so prominent and strong that they should be visible in all spectra within around respective LSE_{Mg_b} values.
(1): Because the H_β line might be affected by bad counts, the H_β-flag is checked and only lines with the following flags are taken into account: "$\sim\sim$", "G1", "G2", "S1". Rest frame wavelength positions for the spectral lines are taken from NIST, "Atomic Spectra Database Lines".

the expected line minimum wavelength for those lines, $\lambda_{l,new}$. Based on this quantity the wavelength range for each spectral line l is adapted according to eq. 3.26 $\lambda_{l,new} \pm \Delta\lambda = \lambda_{l,new} \cdot (1 \pm 500/v_c)$ for the current cycle.

5. Within this loop launch an inner loop, which cycles through the individual line list and restricts the line wavelength range for the current line l if necessary. The resulting data points $(\lambda_l^k, \tilde{y}_l^k)$ are fitted according to the spectral line fitting method (see section 2.2). Unlike as done in section 3.7.1, here the four initial parameters $(\hat{\lambda}, \sigma, h, d)$ are solely determined from the current spectral line itself and not from other lines.

6. After f has been fitted to the data points, again the parameters of the fit have to fulfill several criteria, which are the same as described in section 3.7.1, with just one exception: Since the spectra shouldn't contain any bad counts anymore, the y_{cor} criterion can be neglected. If the set of fitted parameters passes the criteria, a radial velocity v_l is computed following equation 3.26 for the specific line l.

7. Once all lines from the individual spectral line list are cycled through, the inner loop is finished and v_{rad} and $\sigma_{v_{rad}}$ are calculated based on all v_l as long as there are at least three results of v_l. If there are still cycles of the outer loop left, v_{rad} is used as initial radial velocity in step (4), otherwise it is used to convert the wavelength scale into a radial velocity corrected wavelength scale.

The histograms of v_{rad} and $\sigma_{v_{rad}}$ are plotted below, see figures 3.24 to 3.29.

spectral line	$\lambda\,[A]$	LSE_{Mg_b} (from / to)	0.00 / 0.05	0.05 / 0.10	0.10 / 0.15	0.15 / 0.20	0.20 / 0.25	0.25 / 0.30	0.30 / 0.40
H_ζ	3889.064		•	•	•				
H_ϵ	3970.075		•						
H_δ	4101.734		•	•	•	•	•		
H_γ	4340.472		•	•	•	•	•	•	
H_β	4861.35		(1)	(1)	(1)	(1)	(1)	(1)	
He	4143.761		(2)						
He	4387.930		(2)						
He	4471.480		(2)						
Ca	3933.66			•	•	•			
Ca	3968.47						•	•	
Ca	4226.73				•	•	•	•	•
Mg	4481.327		(2)	•	•	•	•	•	•
Mg	5183.604				•	•	•	•	•
Mg	5528.405								•
Fe	4143.868					•	•	•	•
Fe	4271.760					•	•	•	•
Fe	4325.762					•	•	•	•
Fe	4383.545					•	•	•	•
Fe	4404.75							•	•
Fe	4415.135								•
Fe	5227.15						•	•	•

Table 3.4: The complete list of spectral absorption lines. Rest frame wavelengths are taken from NIST, "Atomic Spectra Database Lines". The black filled circles indicate for which range of LSE_{Mg_b} a spectral line will be added to the individual lines list. However, even if it is on the list it will be reviewed separately (its LSE_{Mg_l} value has to surpass a threshold).
(1): Because the H_β line might be affected by bad counts, the H_β-flag is checked and only lines with the following flags are taken into account: "∼∼", "G1", "G2", "S1".
(2): The He line at 4471.480 Å and the Mg line at 4481.327 Å are rather close to each other and around $LSE_{Mg_b} \sim 0.05$ they are of the same order of magnitude. To prevent a mix-up, I estimate the LSE of the He line with $v_{rad,est}$ and eq. 3.27 to verify if it is He or Mg. Depending on the outcome, either the three He lines or the Mg line will be added to the individual line list.

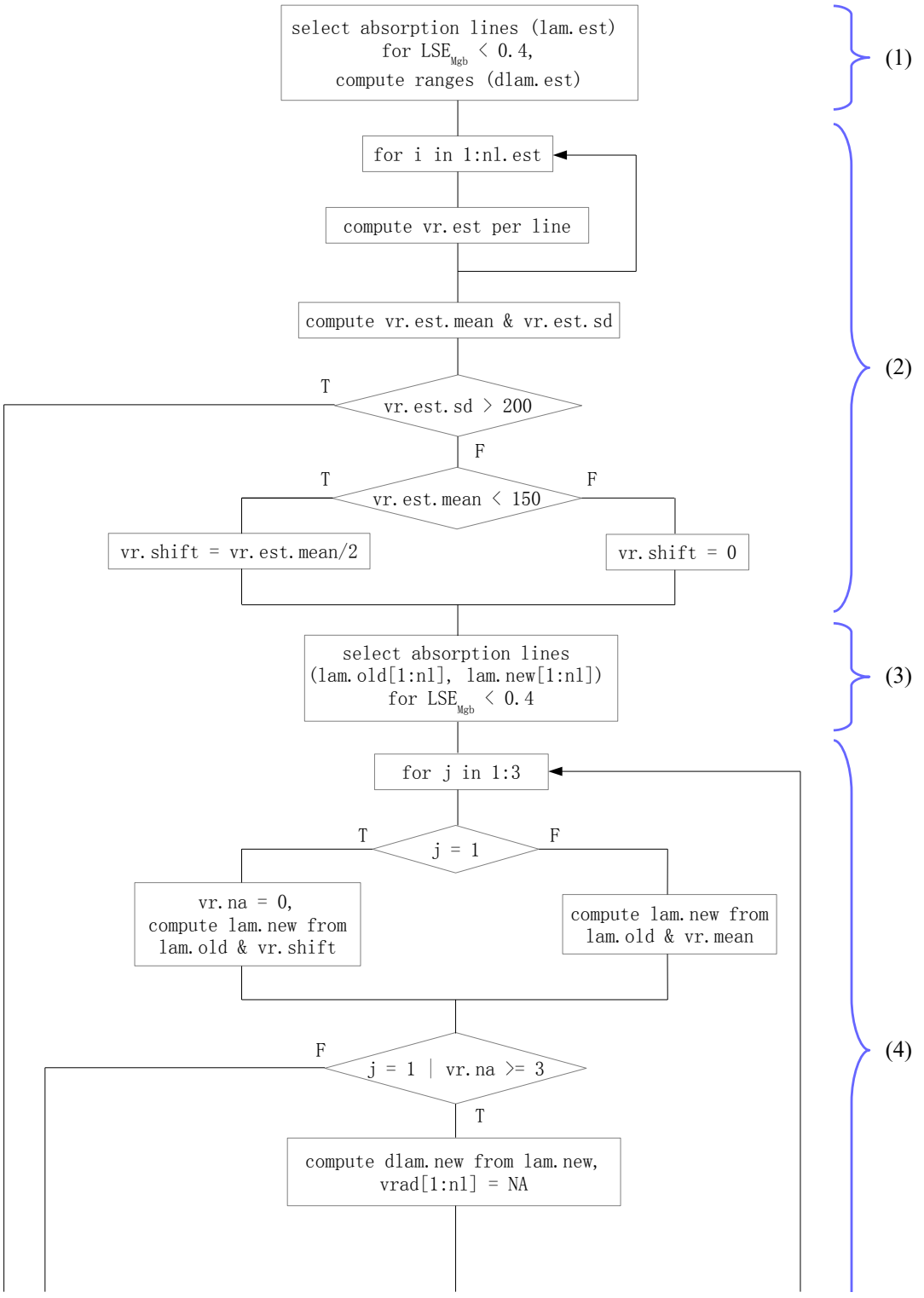

Figure 3.22: The first part of the flowchart to compute radial velocities. The individual steps from the text are labeled on the right side.

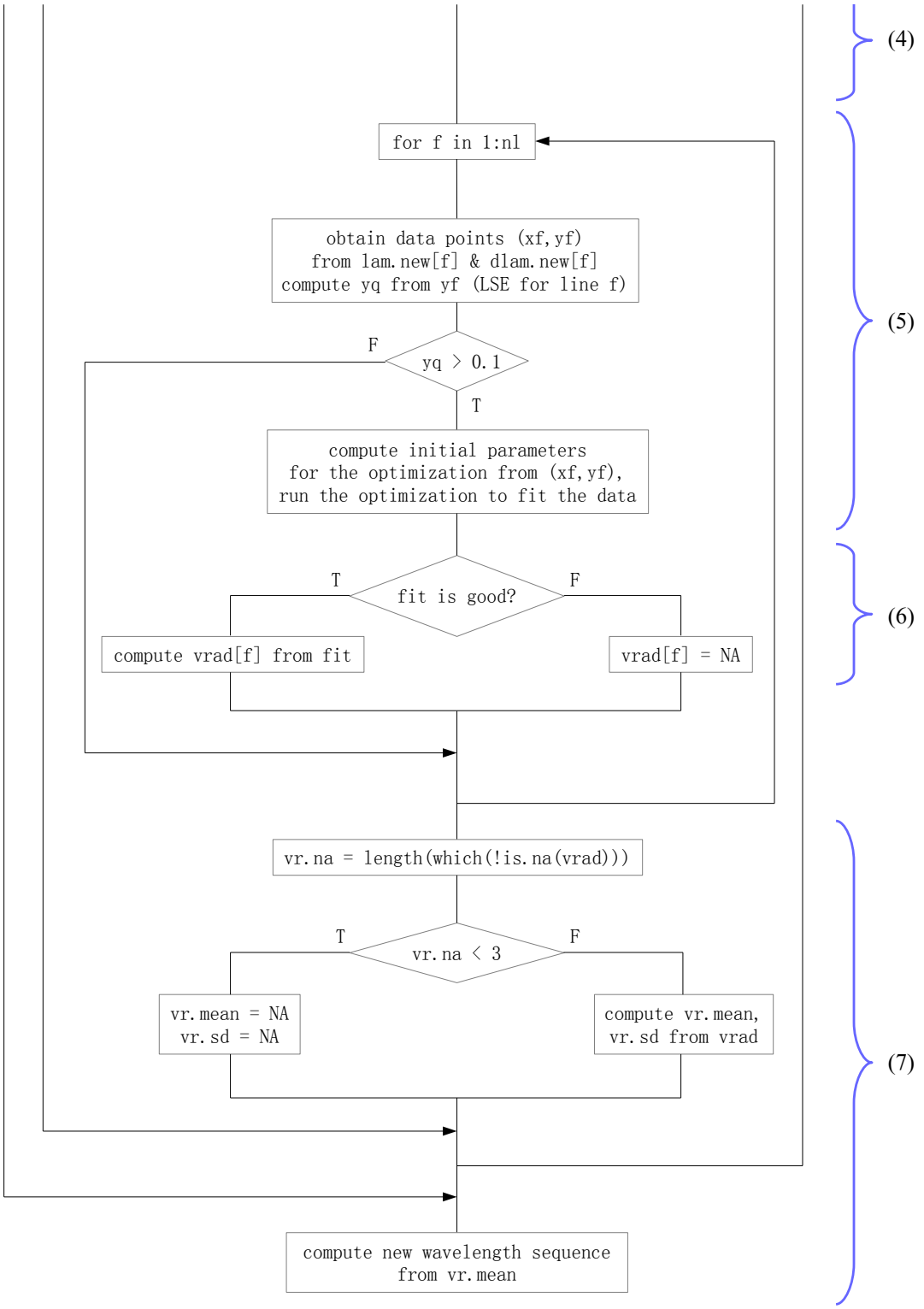

Figure 3.23: The second part of the flowchart to compute radial velocities. The individual steps from the text are labeled on the right side.

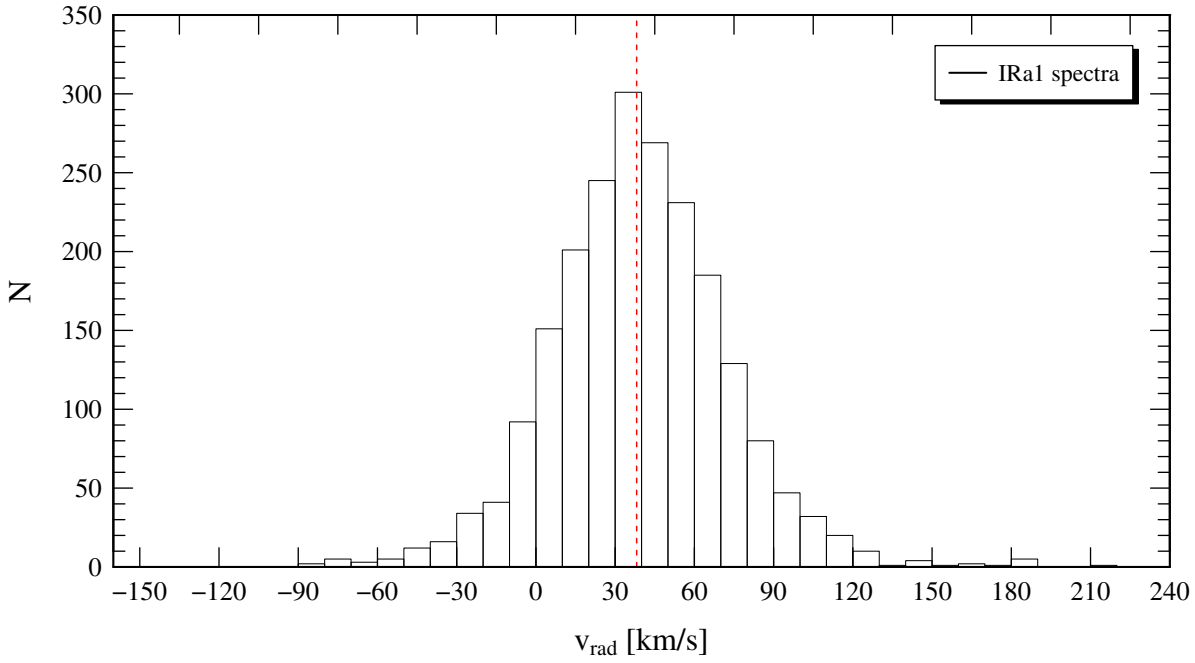

Figure 3.24: The distribution of computed radial velocities for the IRa1 sample. The bin width amounts to $10\,km/s$ and the vertical red line corresponds to the median radial velocity of the sample.

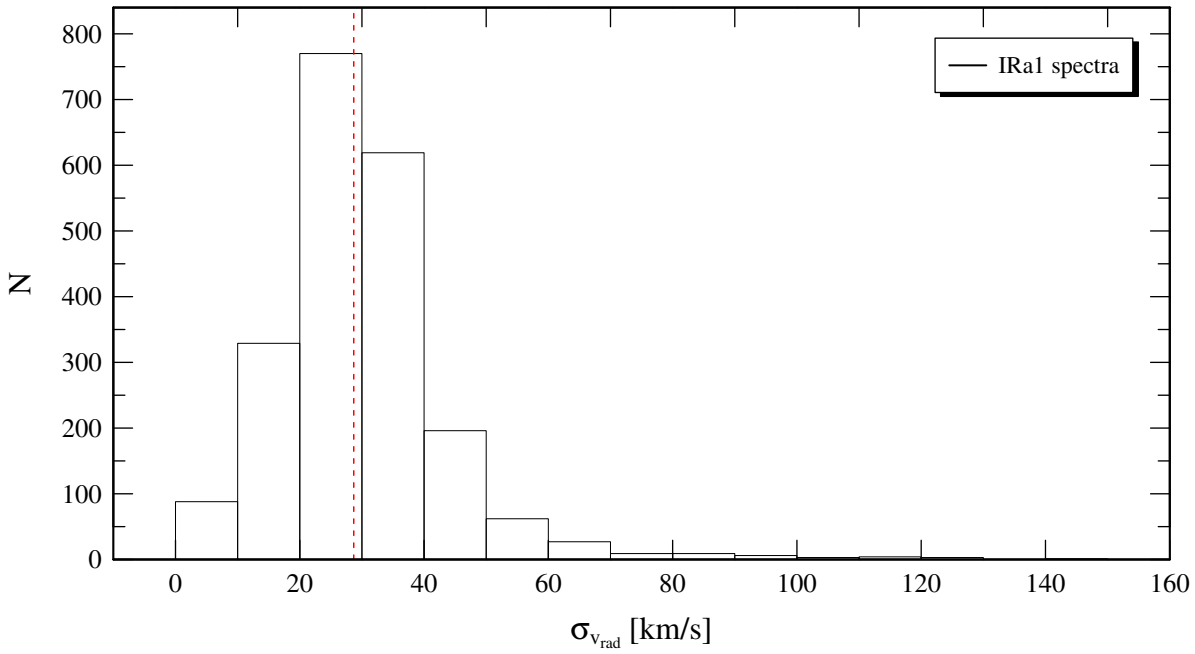

Figure 3.25: The distribution of computed standard deviations of radial velocities for the IRa1 sample. The bin width amounts to $10\,km/s$ and the vertical red line corresponds to the median standard deviation of the sample.

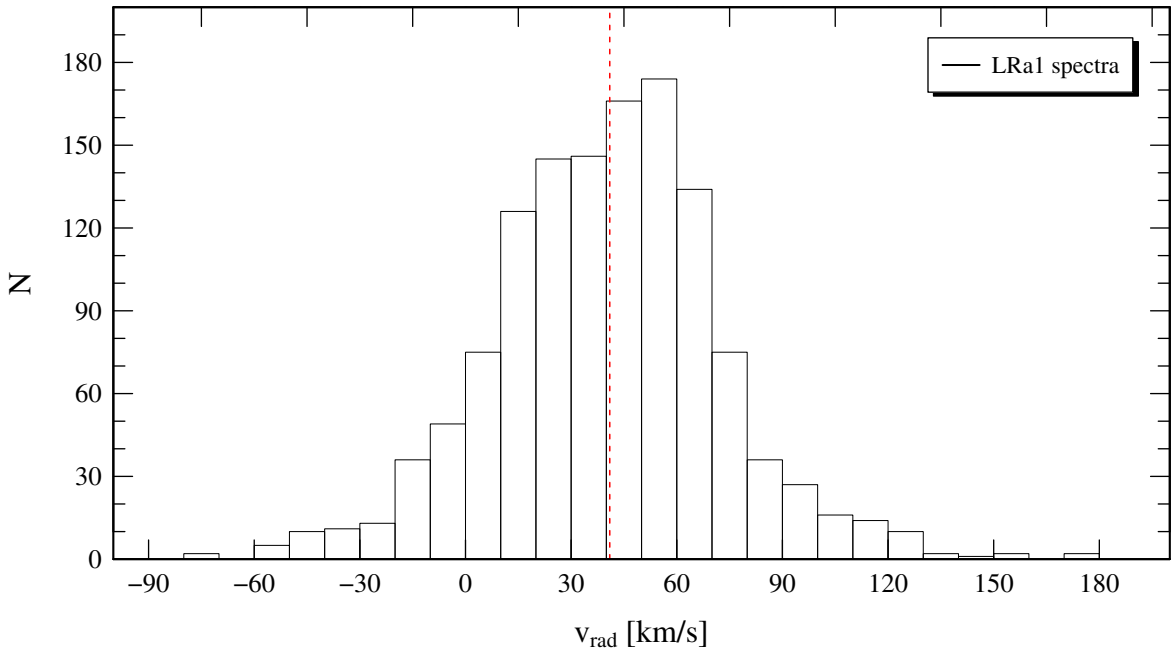

Figure 3.26: Same as in fig. 3.24 for the LRa1 sample.

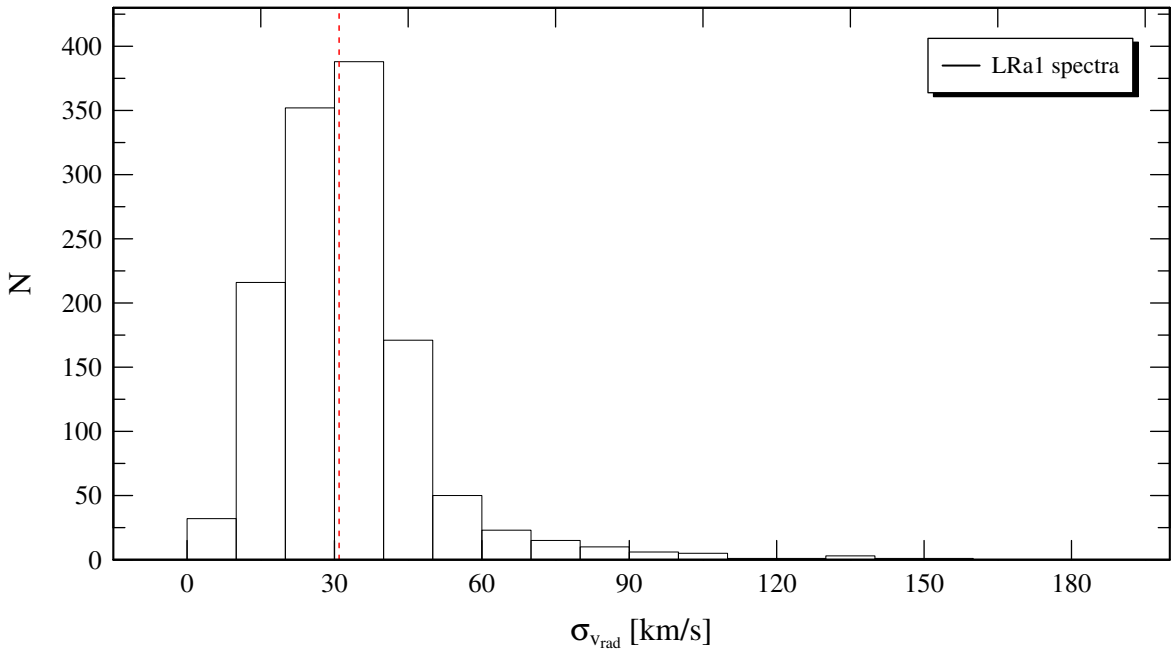

Figure 3.27: Same as in fig. 3.25 for the LRa1 sample.

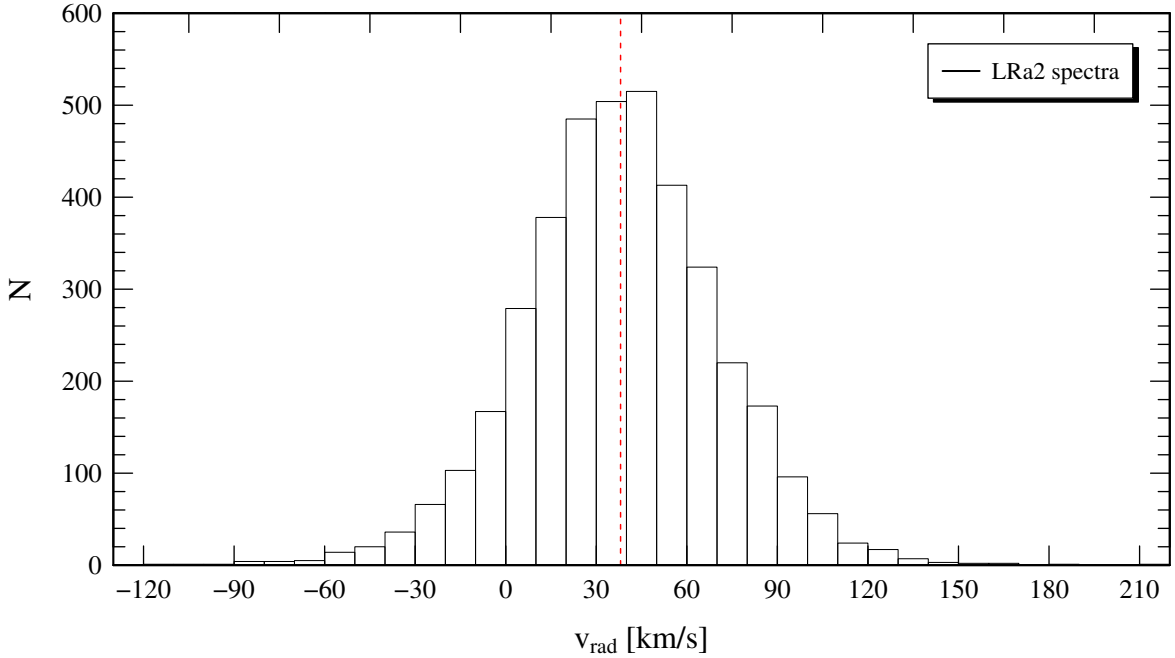

Figure 3.28: Same as in fig. 3.24 for the LRa2 sample.

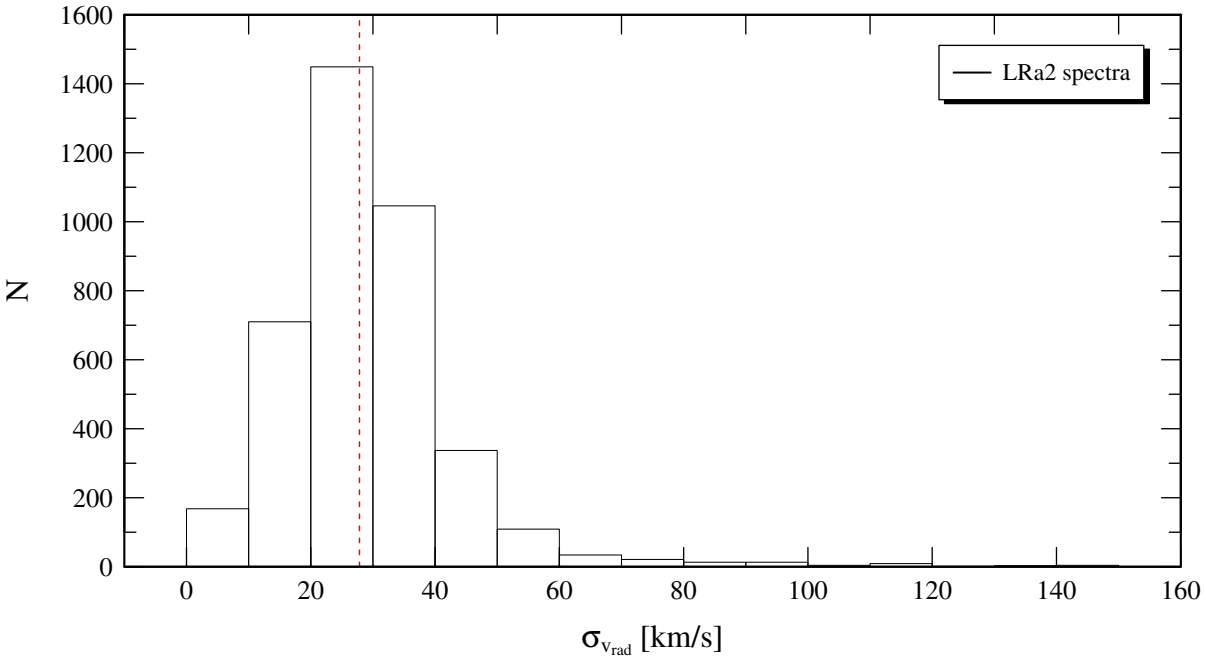

Figure 3.29: Same as in fig. 3.25 for the LRa2 sample.

4 Spectral types via MKCLASS

Spectral types and luminosity classes of the stars have already been published by Sebastian et al. (2012). They compared their observed spectra to template spectra from "The Indo-US Library of Coude Feed Stellar Spectra" (Valdes et al. 2004), computed the root mean square of the differences and minimized it by iteratively adjusting three parameters: radial velocity, interstellar extinction and the flux of the observed spectrum. The results are quite accurate for most spectral types, but it remains unclear, whether and how they handled the wave-like pattern problem and they explain that the determination of luminosity classes might be inaccurate. The accuracy of spectral types is not that important for Δa, since it does not affect the photometric fluxes, but only the assignment of the main spectral type B, A or F. However, obtaining precise luminosity classes is crucial, because until today Δa effects are known to take place only during the main sequence lifetime of a star. As a side effect a detailed stellar classification algorithm might provide at least hints on chemical peculiarities like peculiarities of CP subtypes.

For those reasons I applied an additional spectral type and luminosity class classification routine onto the sample spectra. I use the expert computer program MKCLASS[2] (Gray & Corbally, 2014), which automatically classifies stellar spectra according to the MK stellar classification scheme. The program comes with two libraries of stellar spectra (1.8 and 3.6 Å resolution) covering all spectral types from O- to M-type stars as well as almost all luminosity classes and two subroutines to prepare the input spectra for MKCLASS. The first one is srebin0, used for rebinning spectra and providing equally spaced wavelength data points and the second one is smooth2, used for smoothing spectra to the resolution of the corresponding library by convolving it with a Gaussian. The application of the two routines and MKCLASS is described in the pdf manual "Documentation for MKCLASS v1.03" in much more detail. The MKCLASS routine requires spectra containing fluxes, scaled to the flux at 4503 Å over a wavelength range from 3800 Å to 5600 Å (or smaller ranges for "essential" spectral features), which matches the underlying wavelength domain of the sample spectra perfectly fine. It offers two techniques, ROUGHTYPE 1 and ROUGHTYPE 2, for rather specific and faster (the first one) or more general, but slower classification of

[2] http://www.appstate.edu/~grayro/mkclass/

the spectra (the latter). In the end the short output consists of the name of the input file, a combination of standard or non standard spectral type(s), a luminosity class, possible chemical peculiarities and a quality flag (”excellent”, ”vgood”, ”good”, ”fair”, ”poor” or a blank space for no quality).

I converted the sample spectra that remained after the analysis in section 3.9, from (relative) counts into fluxes, rescaled them to the flux level at 4503 Å, applied the `srebin0` & `smooth2` routines and finally executed the `MKCLASS` routine. All irregular or unknown spectral types were neglected (”?”, ”??”, ”Unclassified spectrum”, white dwarves, carbon and emission line stars). Every outgoing spectral type is converted into a decimal number y_2 (spectral types from Sebastian et al. (2012) are also converted and stored in y_1), based on the spectral main class

$$\begin{bmatrix} O \\ B \\ A \\ F \\ G \\ K \\ M \end{bmatrix} \rightarrow \begin{bmatrix} 10 \\ 20 \\ 30 \\ 40 \\ 50 \\ 60 \\ 70 \end{bmatrix} \qquad \begin{bmatrix} I \\ I-II \\ II \\ II-III \\ III \\ III-IV \\ IV \\ IV-V \\ V \end{bmatrix} \rightarrow \begin{bmatrix} 1.0 \\ 1.5 \\ 2.0 \\ 2.5 \\ 3.0 \\ 3.5 \\ 4.0 \\ 4.5 \\ 5.0 \end{bmatrix}$$

plus the subclass, so for instance a B9.5 star corresponds to 29.5 and a G2 star to 52.0. The luminosity class is converted into a decimal number z_2 (again z_1 are luminosity classes from Sebastian et al. (2012)) neglecting the ”a”, ”b” and ”ab” luminosity suffixes for the luminosity class ”I”. If the output contained more than one spectral type (Am/CP1, λ-Bootis and other metal weak-type stars) an average spectral type is computed, but only if the maximum differences between the two or three spectral types are smaller than 10 subclasses. The comparisons of the distributions of spectral types from `MKCLASS` and Sebastian et al. (2012) are illustrated in figures 4.1 to 4.3. Considering the classification error, which are about a few subclasses, the results are overall in good agreement, but there are two significant deviations: While `MKCLASS` detects much more stars around spectral type

A0, Sebastian et al. (2012) have found many more around A5, but I have not investigated whether these deviations relate to each other. Since Sebastian et al. (2012) have used whole numbers for the luminosity classes and MKCLASS uses also half-numbers, a direct comparison of z_1 and z_2 is not carried out.

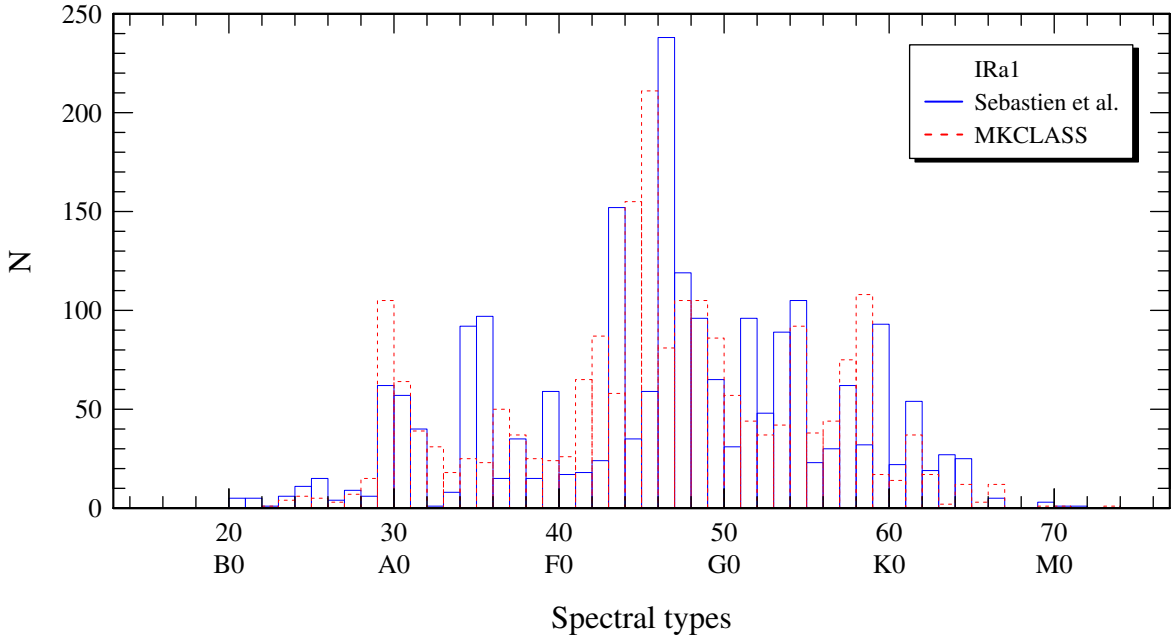

Figure 4.1: The distribution of spectral types of the IRa1 sample along the spectral sequence in direct comparison. Only spectra, where both methods provide spectral types, are taken into account.

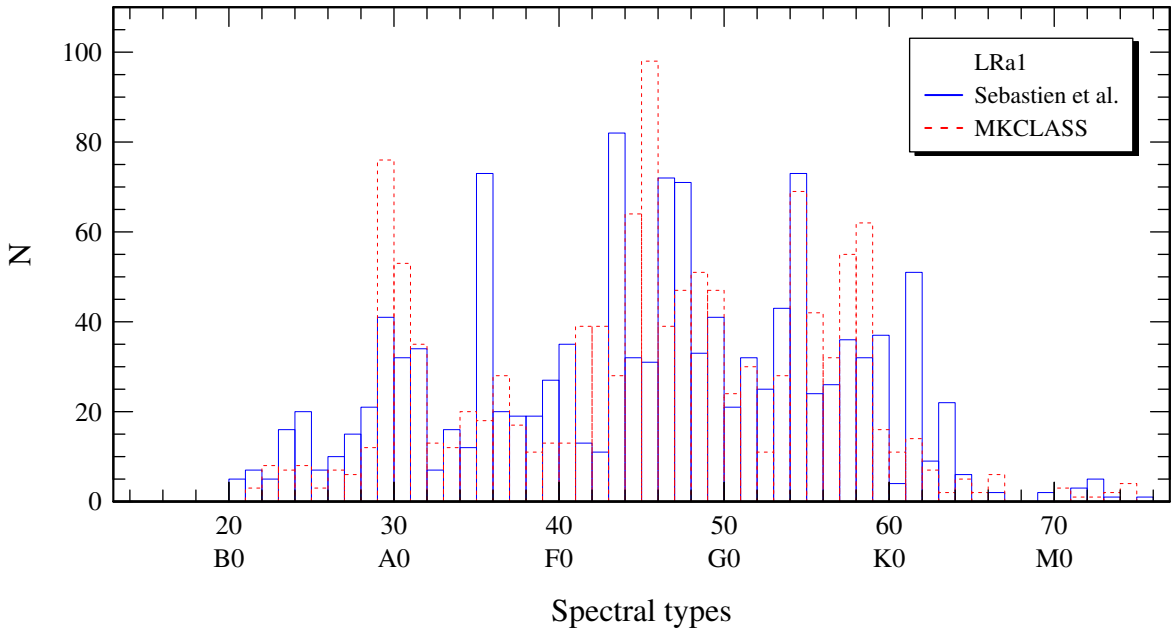

Figure 4.2: Same as in fig. 4.1 for the LRa1 sample.

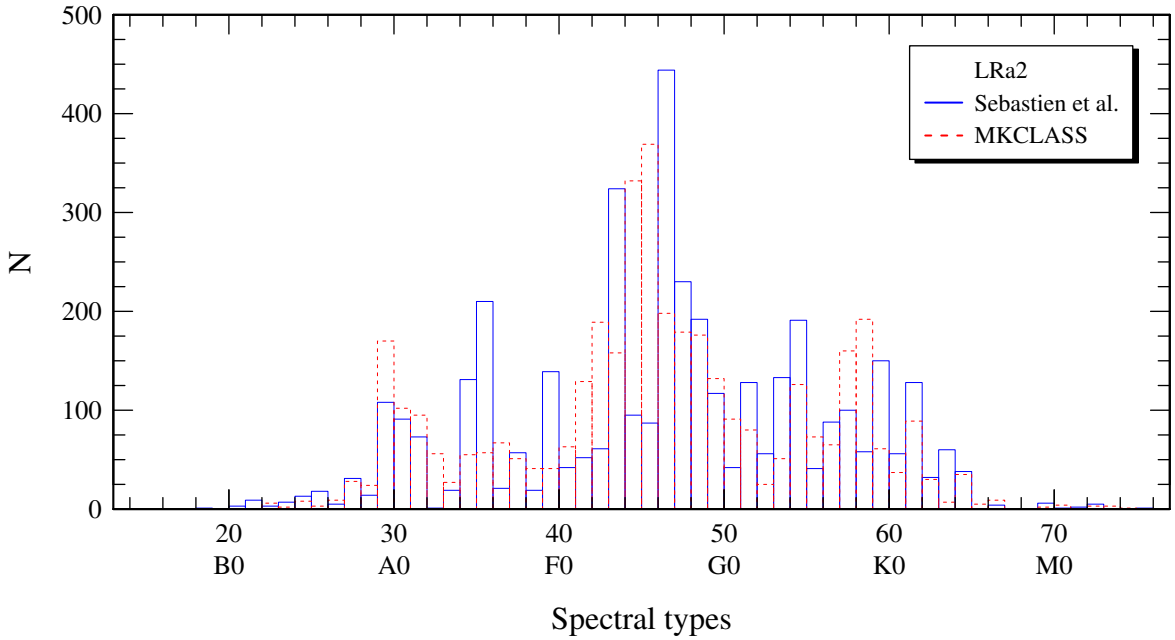

Figure 4.3: Same as in fig. 4.1 for the LRa2 sample.

The spectral types calculated from MKCLASS are automatically picked over the ones published by Sebastian et al. (2012), as long as the differences in spectral types and luminosity classes are small and/or fulfilled other criteria. If there is a significant deviation between the two methods, they are inspected manually. The criteria to inspect the spectra manually essentially restrict the sample of spectra to the following parameter space of y_1, y_2, z_1 and z_2:

- stars classified as λ-Bootis or "metal weak" stars,
- stars with $y_1 + y_2 \leq 99 \ \wedge \ |y_1 - y_2| \geq 5$, which allows huge differences of the two classification schemes, if one star is hot and the other one is cool or it allows small differences, if both stars are for instance late F-stars and they have to differ by at least 5 subclasses,
- $z_1 + z_2 \geq 6 \ \wedge \ |z_1 - z_2| \geq 1 \ \wedge \ z_1 > 4 \ \wedge \ z_2 < 4$, which excludes differences between giants and super giants, luminosity classes have to differ at least by one class, but only if the MKCLASS luminosity class is smaller than 4, because main sequence stars are detected quite accurately by MKCLASS.

In the end the final spectral type and luminosity class is determined hierachically:

1. manually (if neccessary)
2. from MKCLASS
3. from Sebastian et al. (2012)

5 Results

After computing a, a', magnitudes I selected reference spectra for the normality lines according to the criteria in section 1.3. Due to the accuracy of the MKCLASS quality flags, I also included stars with luminosity classes IV-V (the error is about 0.5 luminosity classes for the quality flag vgood). Usually a normality line is chosen such that it covers early B-type to late F-type stars, but within this Master's Thesis I obtain three normality lines (per filter system and per CoRoT sample), one for each main spectral type: B, A and F. This leads to less spectra per normality line, hence a_0 and a'_0 and therfore Δa and $\Delta a'$ are more sensitive to statistical outliers, but even the smallest sample of spectra (B-type stars in LRa1) consists of more than 40 spectra (the other samples contain several hundreds of spectra), hence Δa and $\Delta a'$ can be applied.

Still one major issue has to be dealt with, as some spectra do show all the spectral characteristics of a certain spectral type, but the overall slope of the SED (rather spectral counts distribution, SCD) does not correspond to this spectral type at all. This may be caused by two effects: (a) interstellar reddening, which is wavelength dependent and which has been neglected in this analysis and (b) previous wavelength dependent calibrations to the spectra or observational effects, which I refer to as "color bias". While the first effect only leads to reddening, the second one works in both directions, reddening or blueing a spectrum. In figure 5.1 four spectra of the same field and the same spectral type (F5 V) are plotted, but one can clearly make out the different slopes of the continua in all spectra, hence the different colors. While colors can be heavily biased, a and a' magnitudes are hardly affected by this bias, as they measure only the depth of a possible flux depression around 5200 Å. If such a spectrum is used to derive the normality line, it will provide an a magnitude, which lies within the expected range of its spectral type, but it will be shifted to much redder domains in the color-a plot and therefore influence Δa diagnostics in two ways: lowering the slope of the normality line and increasing the spread $\sigma_{\Delta a}$ for Δa. Hence it is crucial to exclude spectra from the a_0 selection list, by setting a limit on the most likely color domain for each subsample. Spectra, which are just slightly affected by this, will remain in these lists of course and this is why I am not just analyzing the $|\Delta a| > 3\sigma_{\Delta a}$, but also the $|\Delta a| > 2\sigma_{\Delta a}$ outliers. Once the a_0 lists are set, a

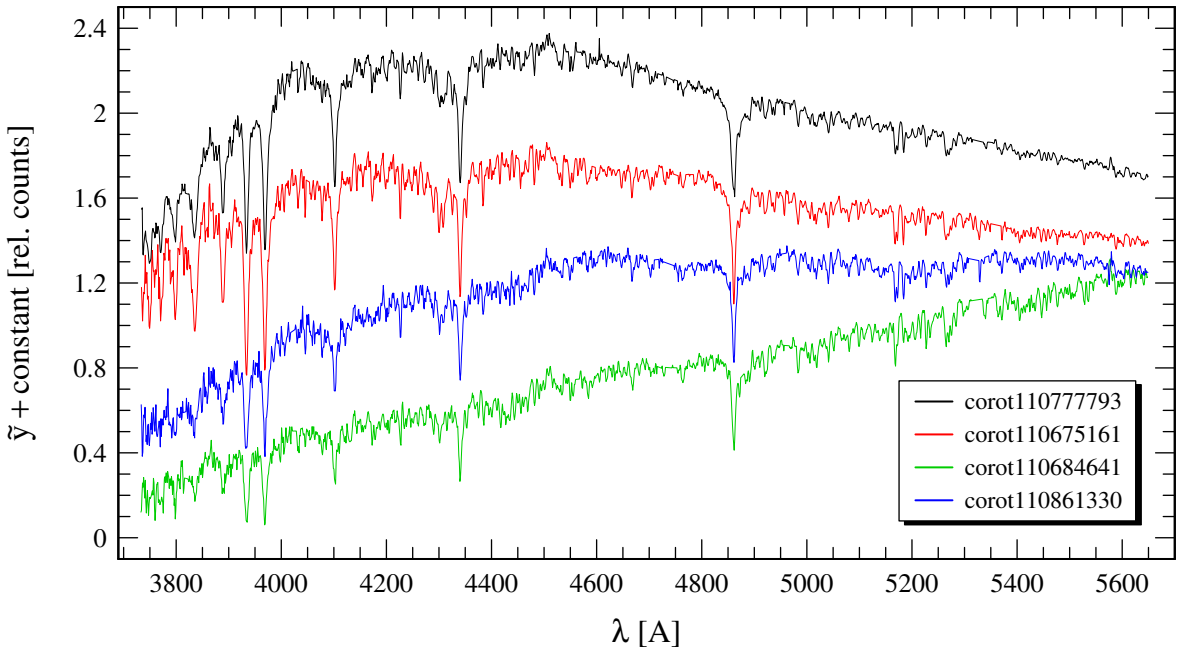

Figure 5.1: Four non peculiar spectra from the same CoRoT field and with the same spectral type according to MKCLASS (F5 V). Even though the slopes of the continua are very different, the individual spectral features are similar. Due to this effect the color $(g_1 - y)$ leads to a significant spread: The green spectrum has a much redder color than the black one, but their Δa will be comparable. Please note that the blue spectrum shows an undeteced faint wave-like pattern, which simulates a flux depression around 5200 Å.

normality line is computed via linear regression analysis and Δa indices are obtained for all stars with the corresponding main spectral type, luminosity classes between IV-V to V and all peculiar star designations of MKCLASS (λ Bootis, Am, metal weak) regardless of luminosity class, but again with the corresponding main spectral type.

The results are plotted in the following sections, but due to clarity only the results of the original Δa system and a few important spectra, which illustrate either the success or the flaws of the method, are plotted within the respective subsections. The results of all spectra are summed up in table 6.1.

5.1 IRa1 sample

5.1.1 B-type stars

In the IRa1 field the B-type stars show a small spread around the normality line in the *a* - color diagram, with one exception.

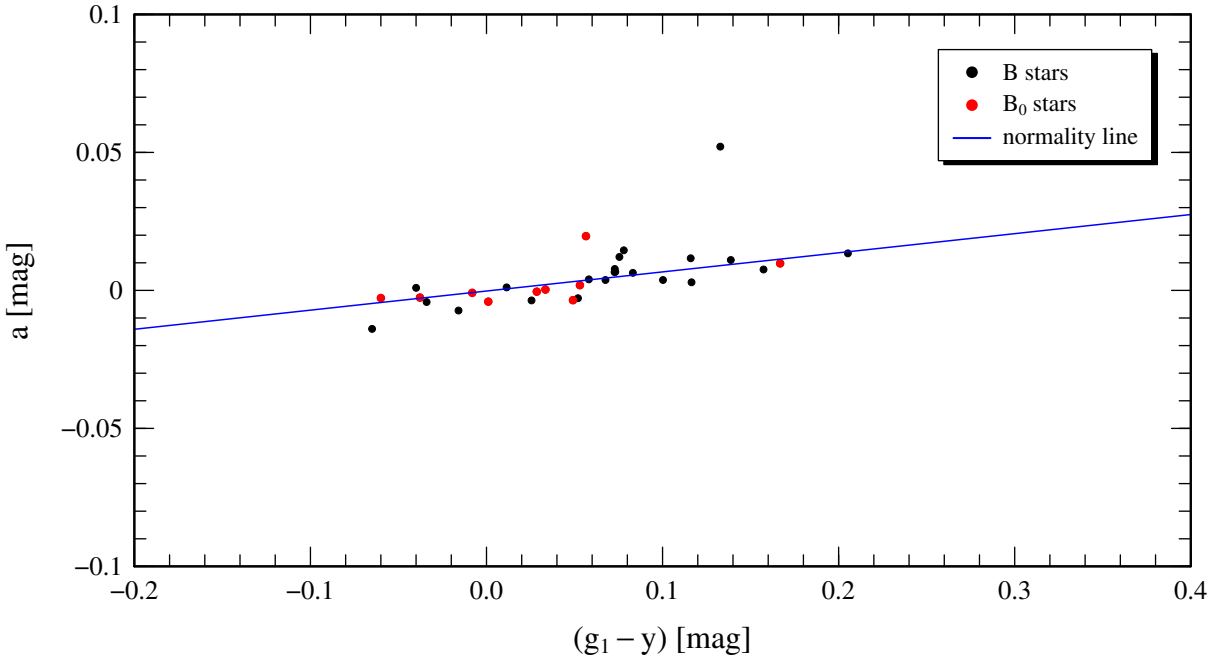

Figure 5.2: The *a* - color diagram of all IRa1 B-type stars with luminosity class of IV-V or V, plus all λ Bootis stars and spectra with "metal rich" or "metal weak" suffixes in the same spectral range.

The statistical outlier (1B) is a CP2 star according to Δa and also `MKCLASS` classifies it as B9.5 V Si star (CP2). Prominent CP2 spectral feature are visible in the spectrum (Si II $4128 - 4131$ Å, see fig. 5.4).

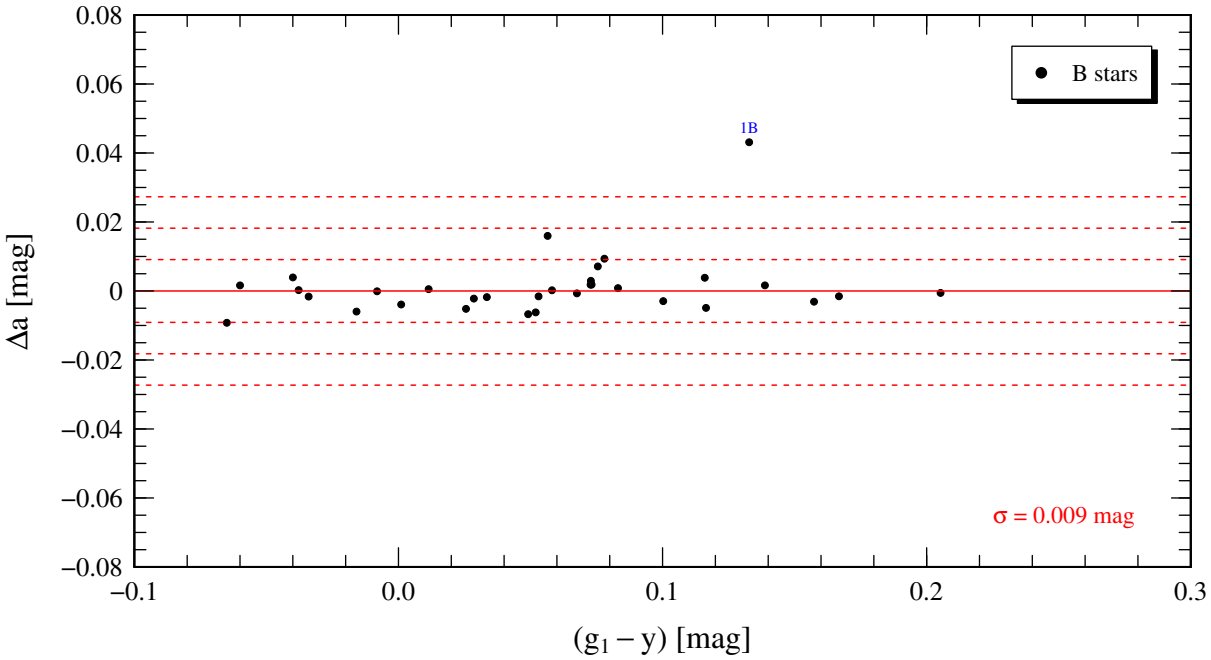

Figure 5.3: The Δa - color diagram of all IRa1 B-type stars with luminosity class of IV-V or V, plus all λ Bootis stars and spectra with "metal rich" or "metal weak" suffixes in the same spectral range. The horizontal lines represent ± 1, ± 2 and $\pm 3\sigma_{\Delta a}$ ranges of the underlying sample of spectra.

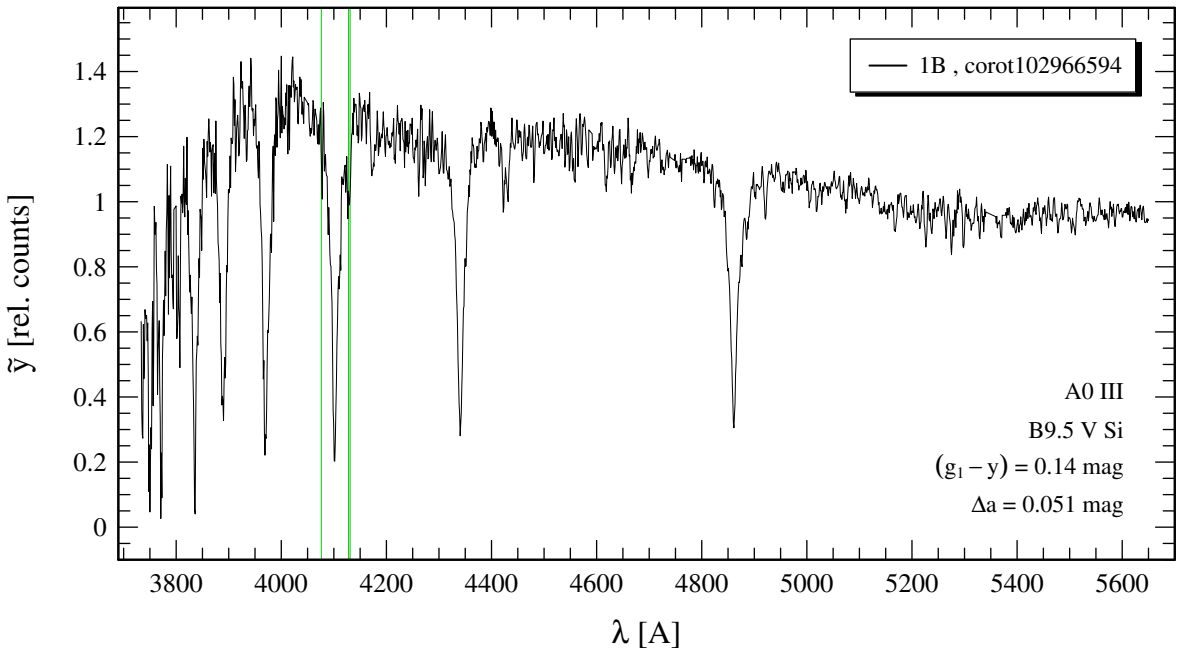

Figure 5.4: The spectrum of corot102966594. The Si II doublet at $4128 - 4131$ Å can clearly be seen.

5.1.2 A-type stars

As for the A-type stars Δa the largest statistical outlier lies between $7\sigma_{\Delta a}$ and $8\sigma_{\Delta a}$ (14A, not shown in the Δa - color diagram) and is most likely a remnant of spectra with wave-like patterns. Its spectrum does not contain any visible dominant features of chemical peculiarities and also MKCLASS does not classify it as chemically peculiar star, but a distinct wave patter can be recognized. The remaining possible CP star candidates differ from doubtful to distinct CP stars, but also a few spectra with wave-like patterns appear to be part of the ± 2 to $\pm 3\sigma_{\Delta a}$ domain.

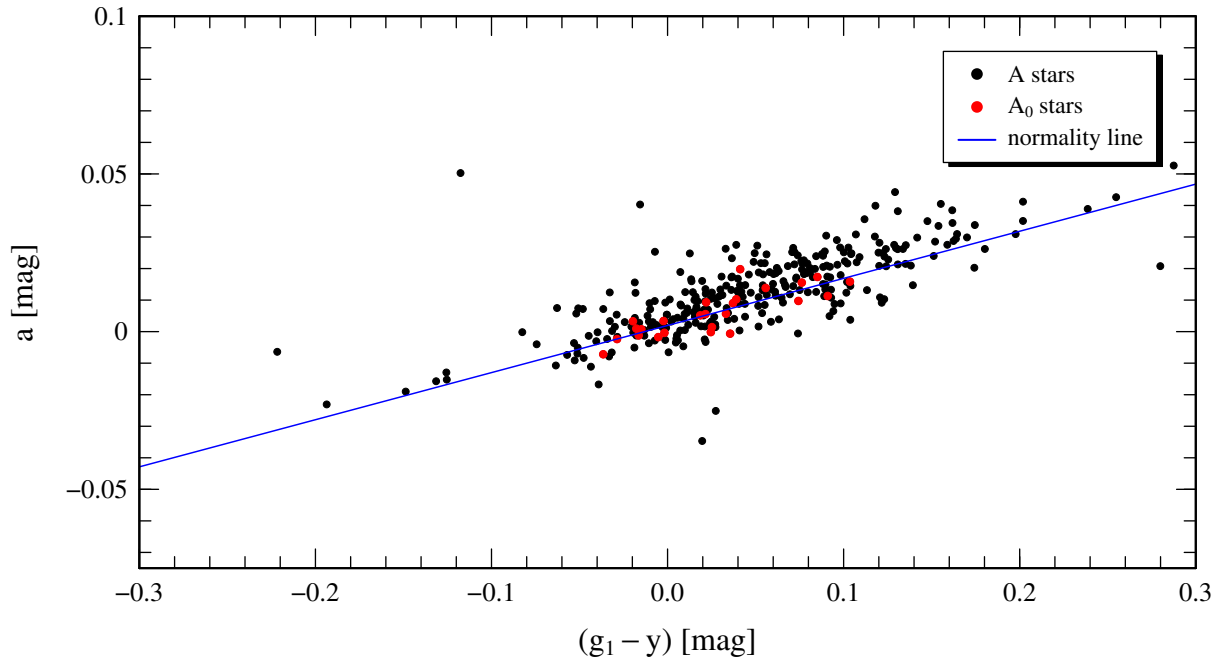

Figure 5.5: Same as in fig. 5.2 for A-type stars.

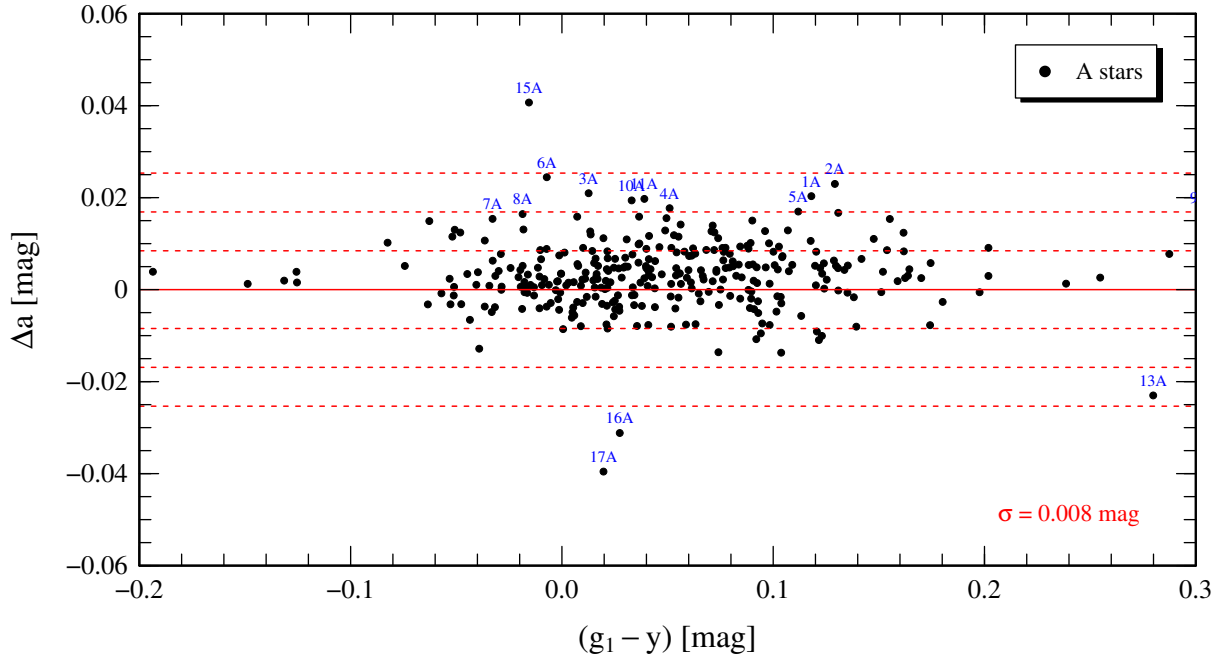

Figure 5.6: Same as in fig. 5.3 for A-type stars.

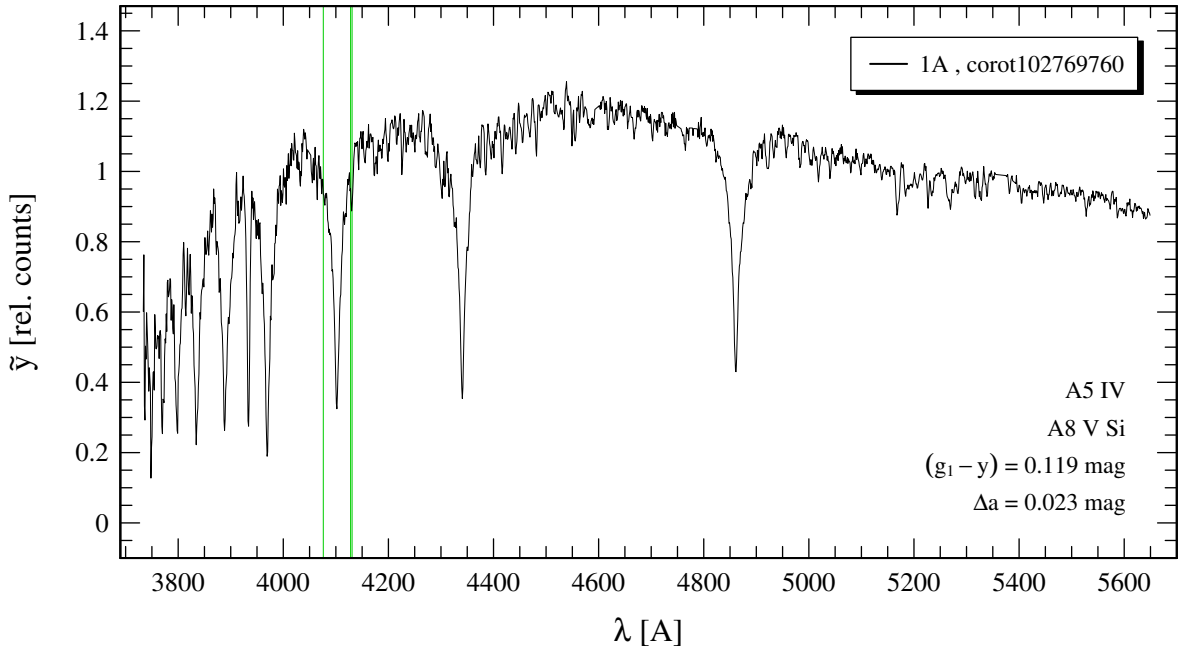

Figure 5.7: The spectrum of corot102769760. The Si II doublet at 4128 − 4131 Å can clearly be seen.

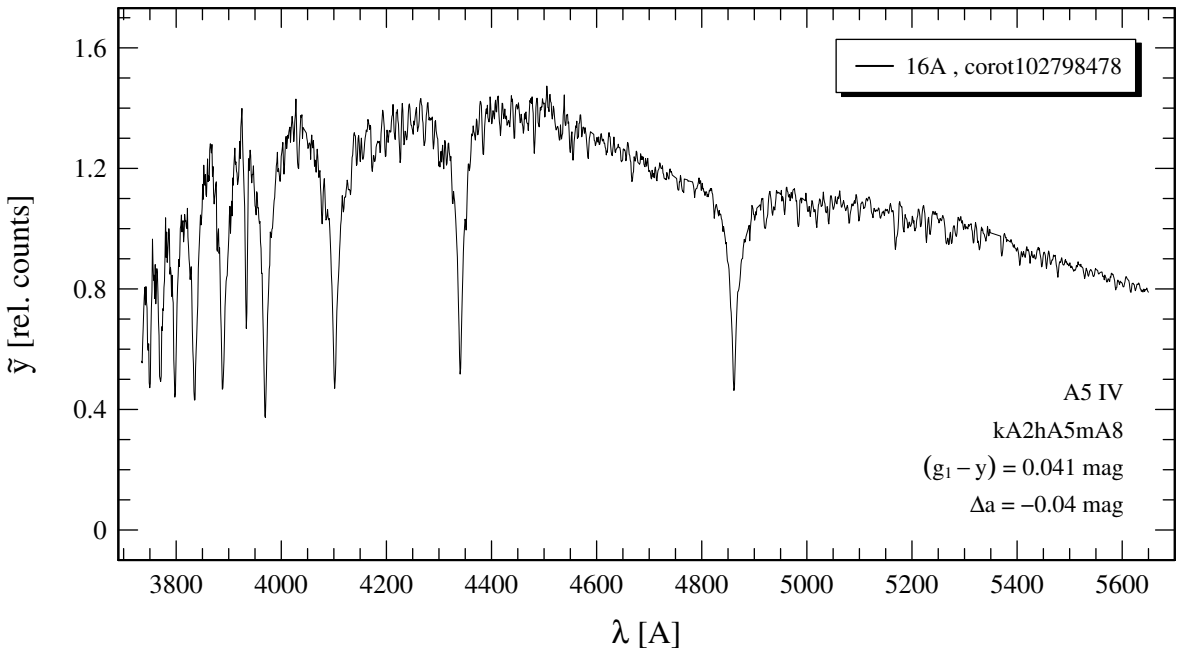

Figure 5.8: The spectrum of corot102798478. This star is classified as Am-type star, because while the calcium lines indicate an early A-type star, the hydrogen lines suggest an intermediate A-type star (compare it to fig. 5.7) and the metallic lines correspond to a late A-type star.

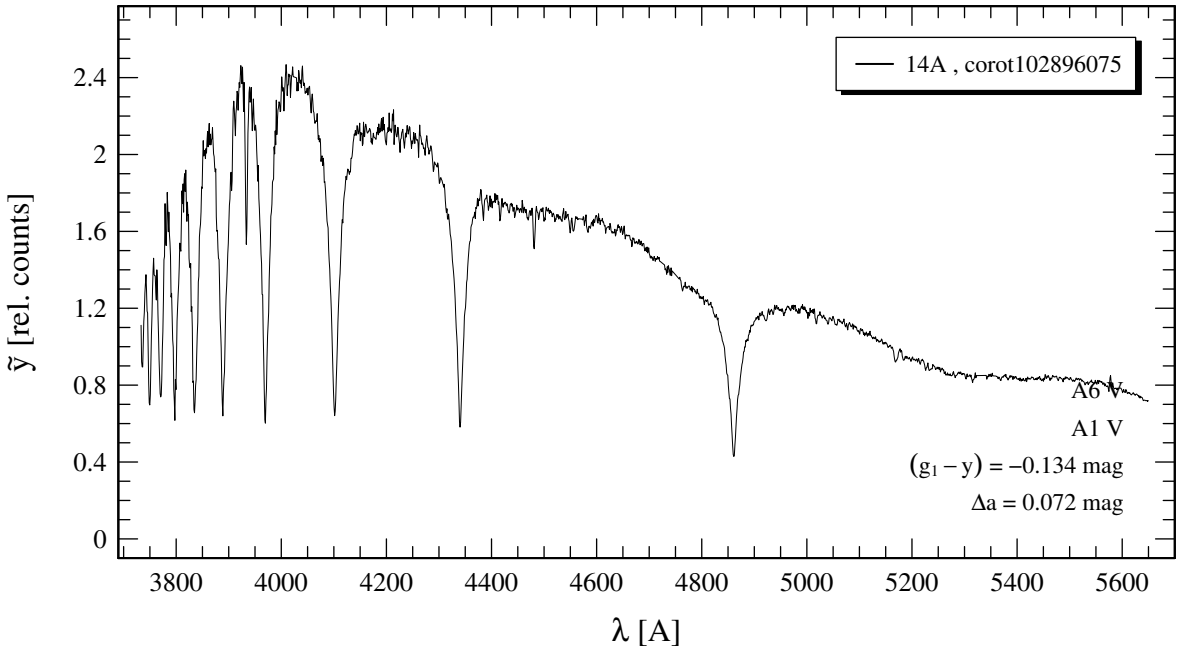

Figure 5.9: The spectrum of corot102896075. No chemical peculiarities are detected, the major Δa flux depression has a small offset to its usual wavelength domain around 5200 Å and there appears to be another bump around 4600 Å so it is very likely that this is not a Δa effect, but a wave-like pattern with a pattern wavelength of about 300 Å.

5.1.3 F-type stars

Unfortunately the detected F-type star spectra do not contain any visible chemical peculiarities, but also MKCLASS classified those spectra as non-peculiar. The main reason for this is that the wave-like pattern analysis from section 3.8 does not cover all F-type spectra and spectra with such patterns tend to stand out in Δa magnitudes more likely. Another reason is bad sampling of F_0 stars, as they do not exceed 0.16 mag in color, while the sample contains stars with colors above 0.3 to 0.4 mag. However, this effect is rather small compared to the deviations of certain spectra from the normality line and selecting F_0 spectra on a wider color range, hence making the normality line more robust, would only shift a few spectra from the low $2\sigma_{\Delta a} - 3\sigma_{\Delta a}$ domain to the high $1\sigma_{\Delta a} - 2\sigma_{\Delta a}$ domain in Δa - color diagram.

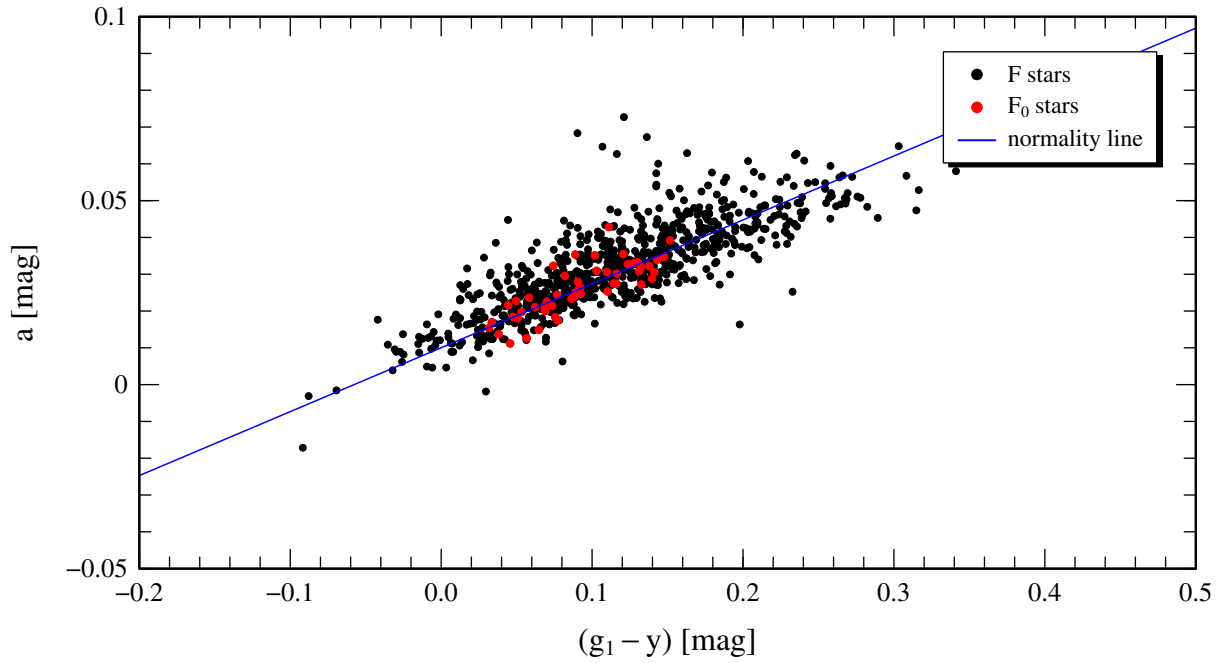

Figure 5.10: Same as in fig. 5.2 for F-type stars.

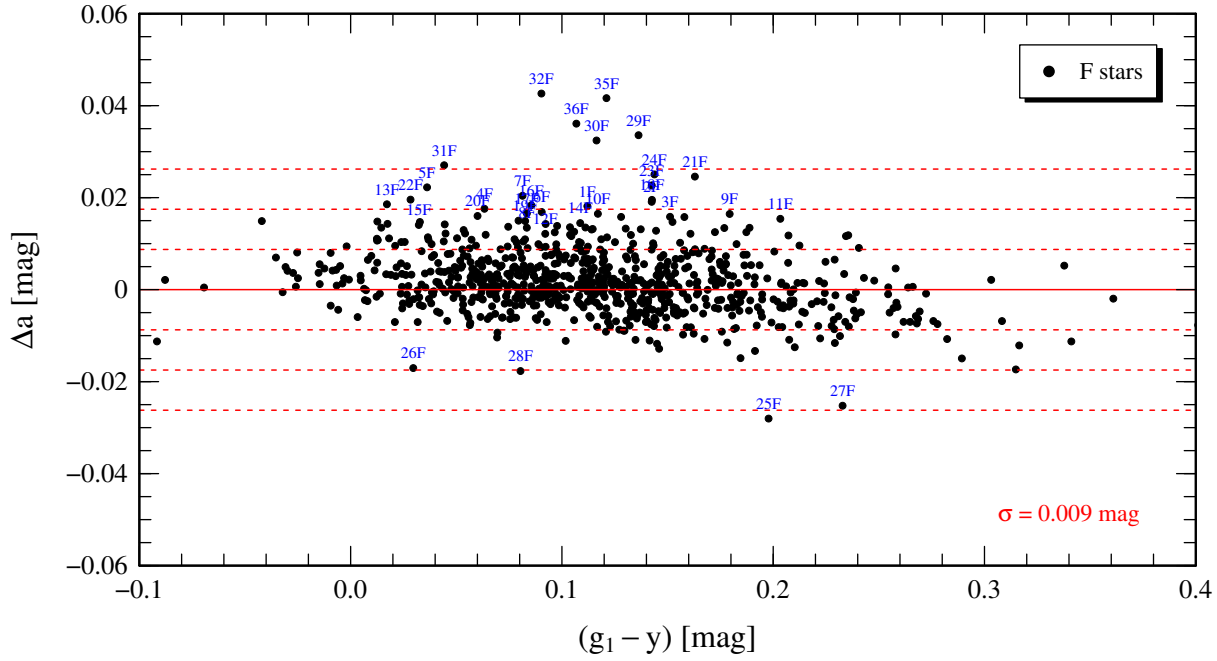

Figure 5.11: Same as in fig. 5.3 for F-type stars.

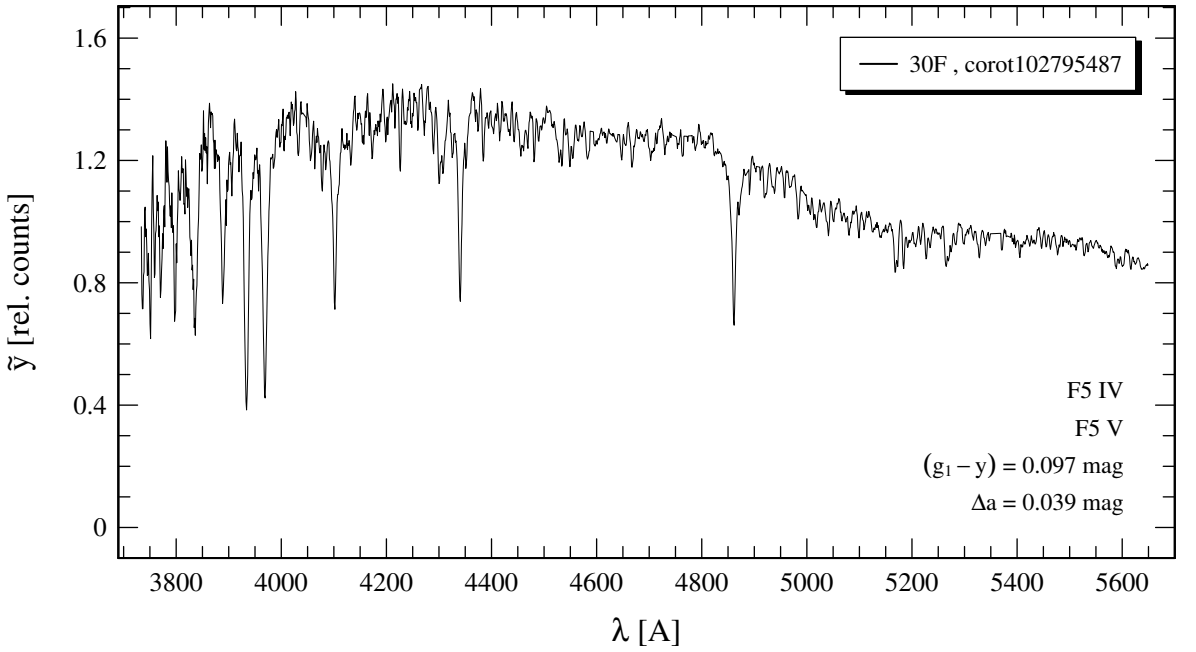

Figure 5.12: The spectrum of corot102795487. A huge flux depression around 5200 Å with unclear origin is visible, but `MKCLASS` does not indicate a chemical peculiarity.

5.2 LRa1 sample

5.2.1 B-type stars

The B-type stars of the LRa1 sample contain two distinct CP2 stars (2B, 3B) and unfortunately one misclassified F-type star with low S/N (1B).

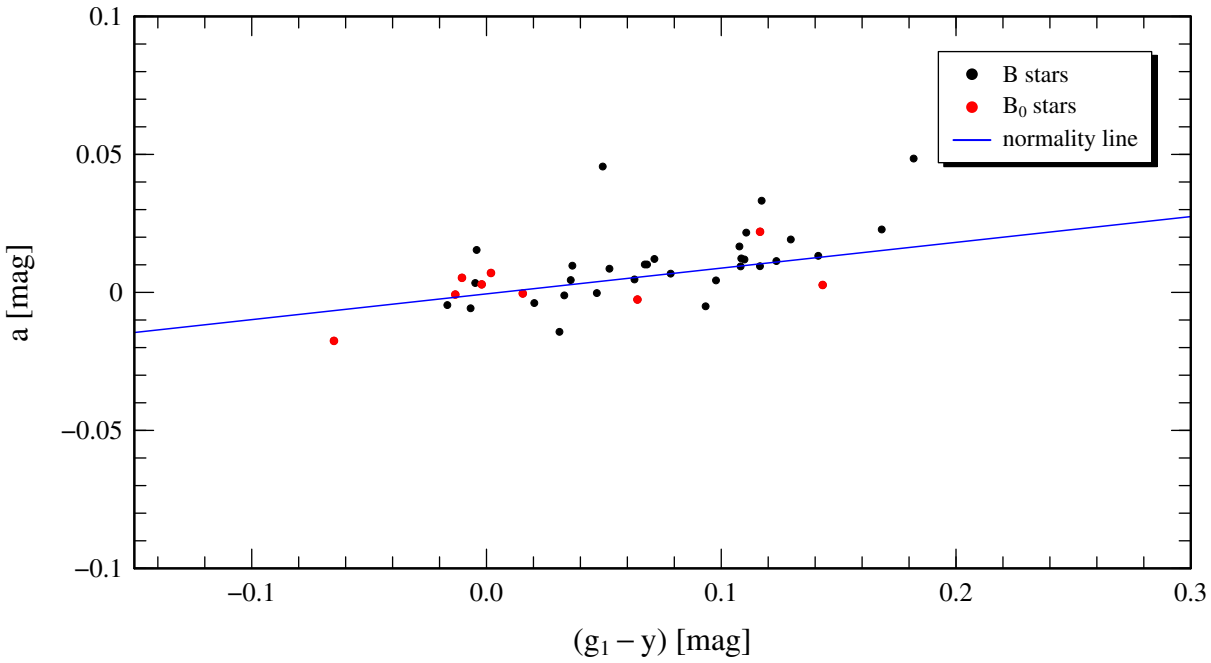

Figure 5.13: The *a* - color diagram of all LRa1 B-type stars with luminosity class of IV-V or V, plus all λ Bootis stars and spectra with "metal rich" or "metal weak" suffixes in the same spectral range.

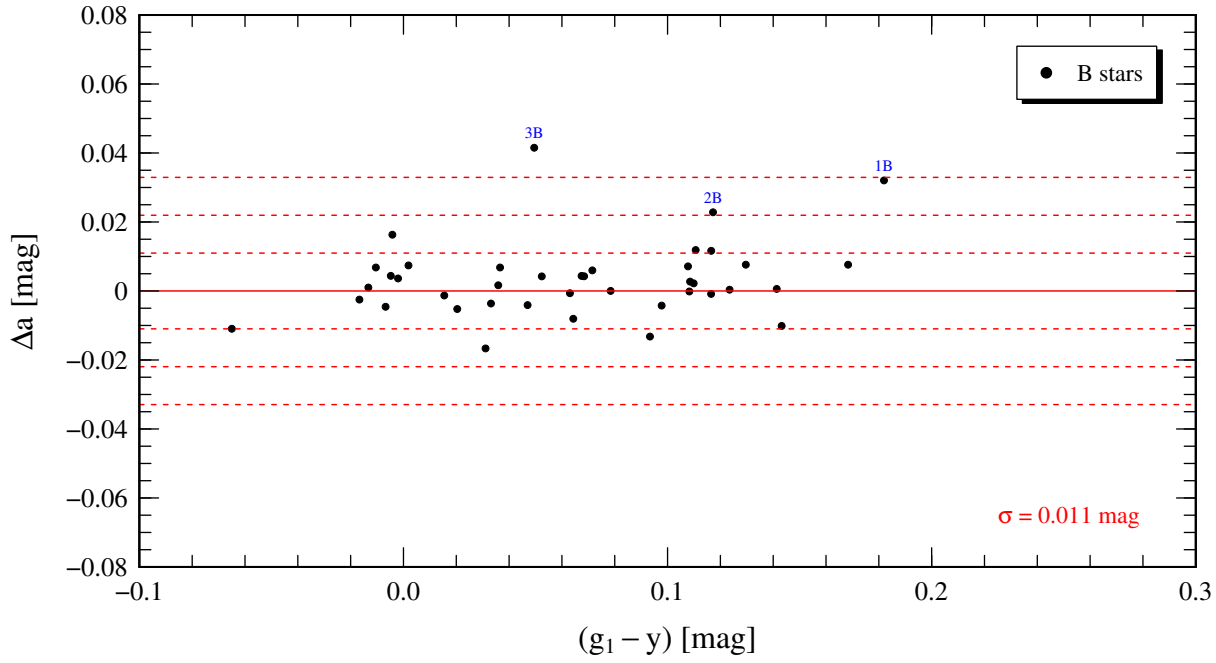

Figure 5.14: The Δa - color diagram of all LRa1 B-type stars with luminosity class of IV-V or V, plus all λ Bootis stars and spectra with "metal rich" or "metal weak" suffixes in the same spectral range. The horizontal lines represent ± 1, ± 2 and $\pm 3\sigma_{\Delta a}$ ranges of the underlying sample of spectra.

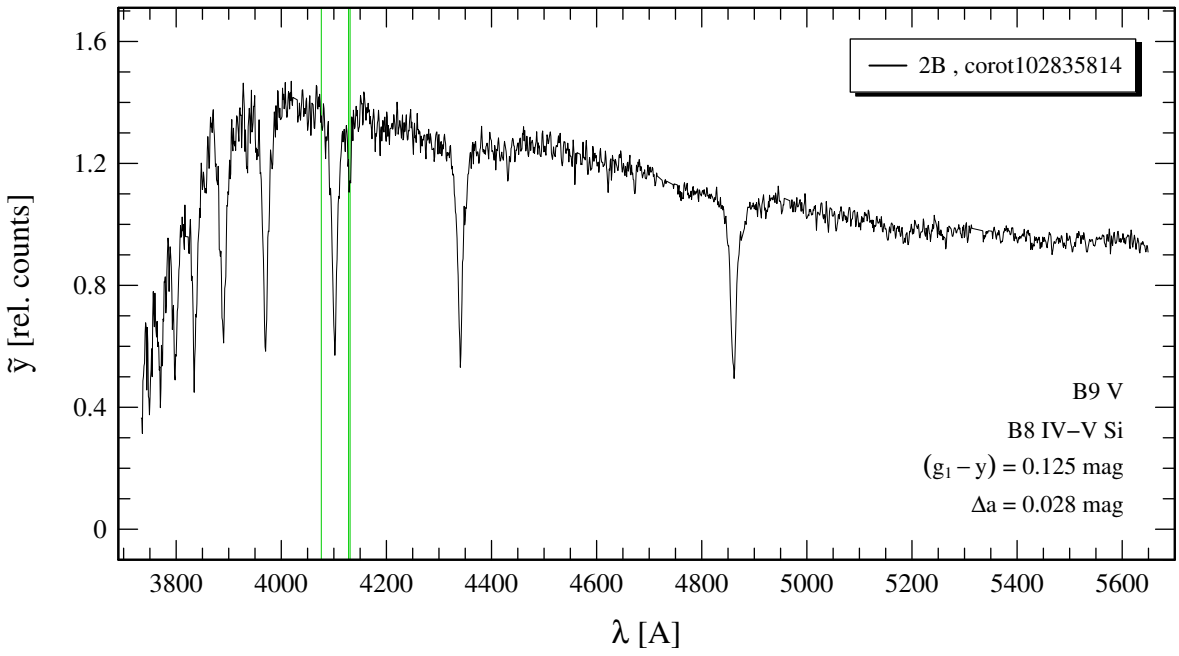

Figure 5.15: The spectrum of corot102835814. There is a slight flux depression around 5200 Å, but a clear Si II doublet at $4128 - 4131$ Å.

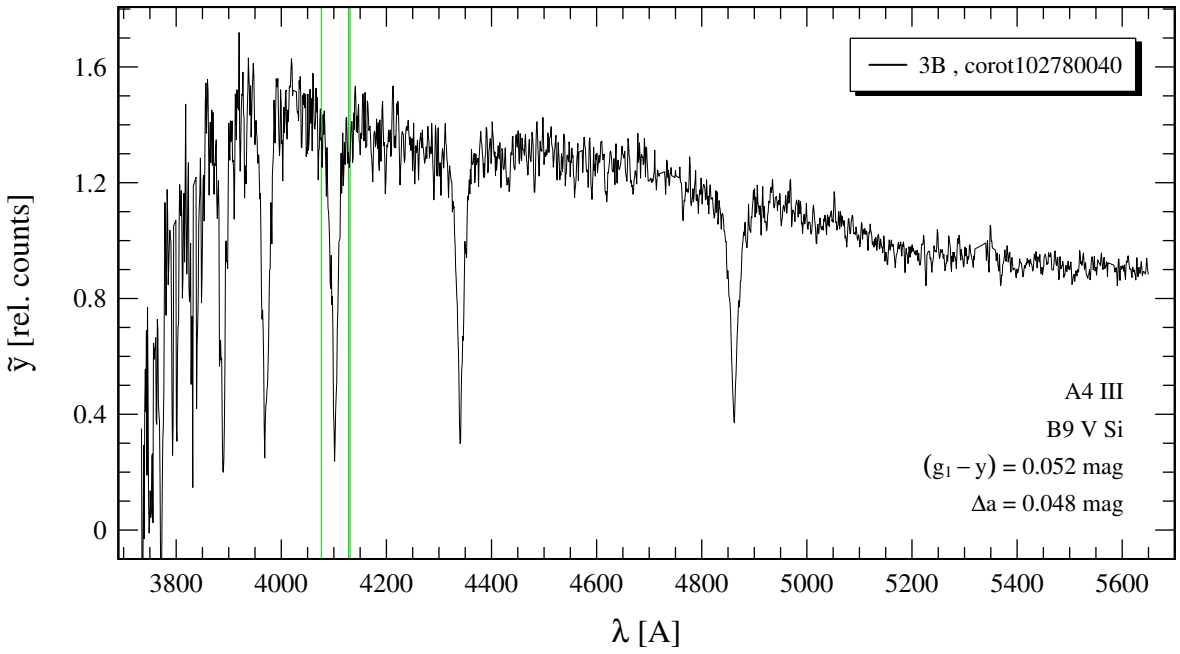

Figure 5.16: The spectrum of corot102780040. A decent flux depression around 5200 Å can be seen, though the Si II features are rather weak. This spectrum has a pretty low S/N, a lot of noise occurs.

5.2.2 A-type stars

As for the A-type stars, one spectrum with $-4\sigma_{\Delta a} < \Delta a < -3\sigma_{\Delta a}$ might contain a wave-like pattern (12A), but `MKCLASS` detects a possible Sr overabundance. Another one with $2\sigma_{\Delta a} < \Delta a < 3\sigma_{\Delta a}$ (3A) provides no hints on chemical peculiarities. Apart from these two spectra all other CP candidates are most likely either CP2, metal rich (CP1) or λ Bootis stars.

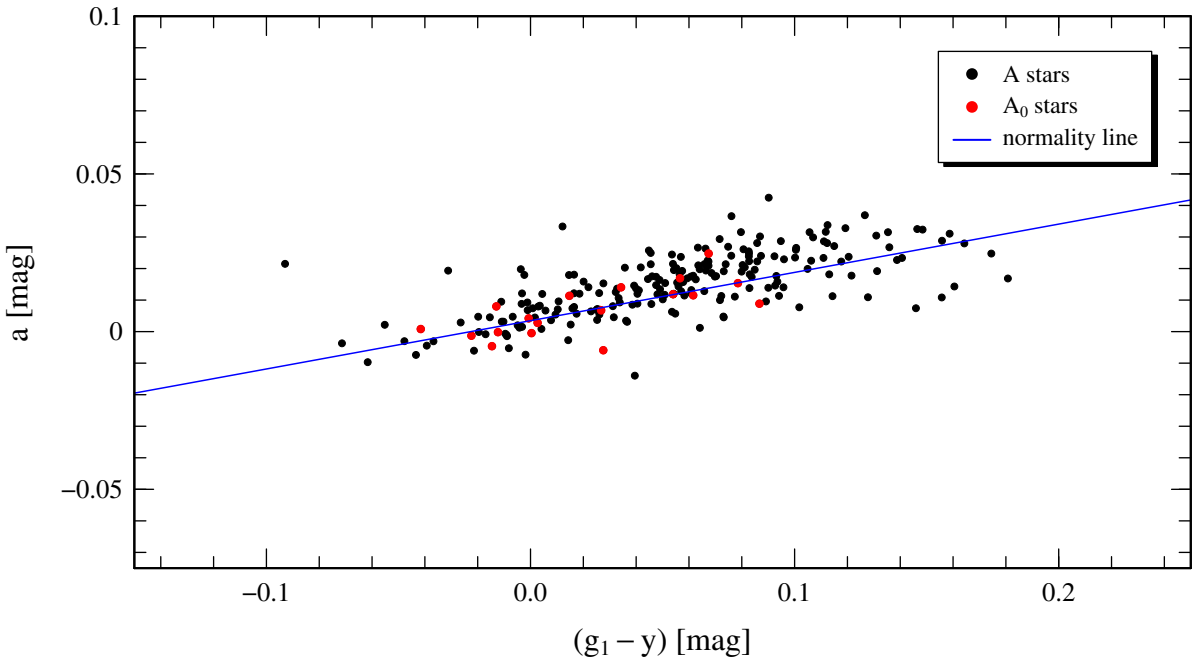

Figure 5.17: Same as in fig. 5.13 for A-type stars.

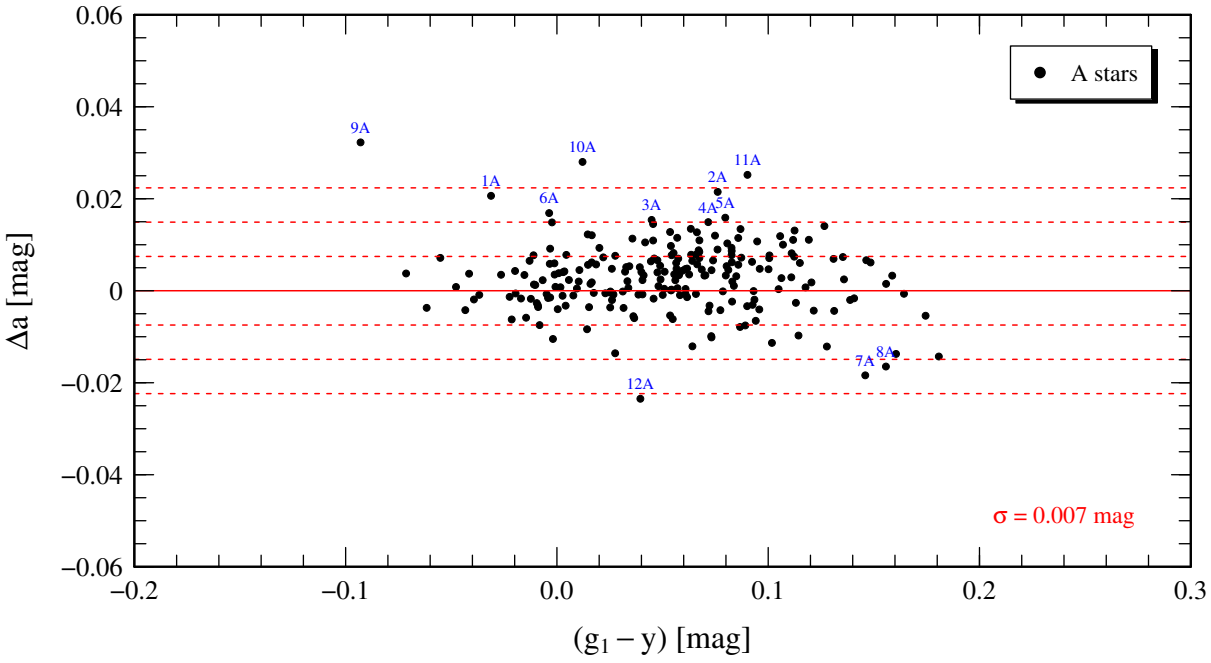

Figure 5.18: Same as in fig. 5.14 for A-type stars.

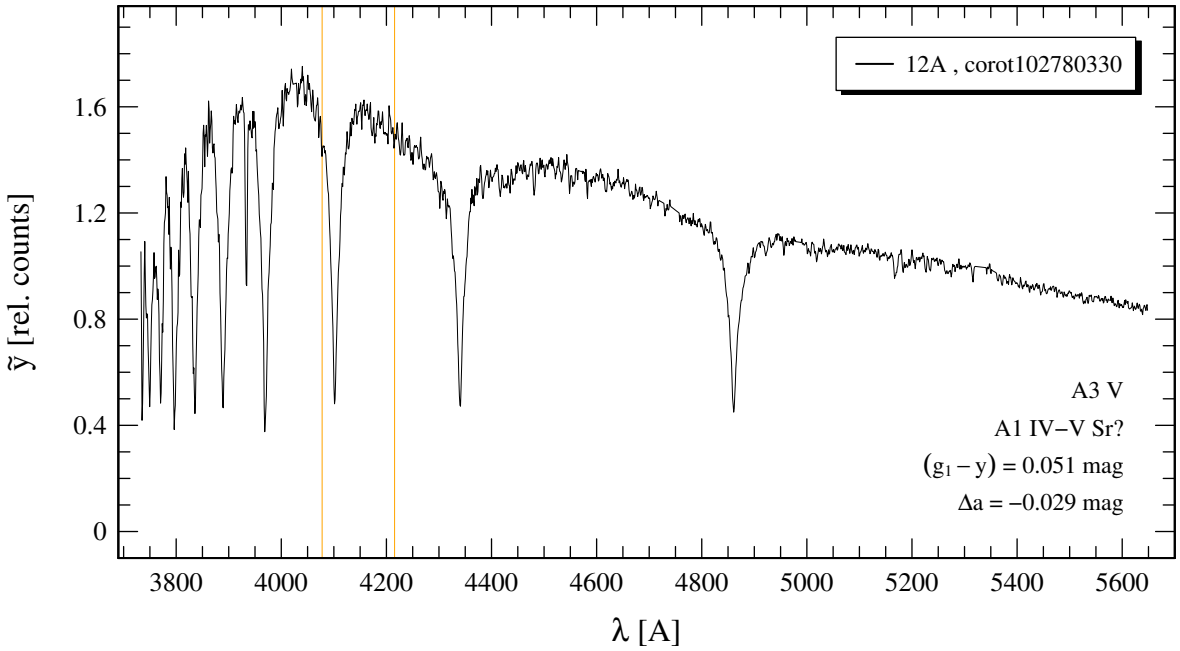

Figure 5.19: The spectrum of corot102780330. According to `MKCLASS` there is a hint of Sr II at 4078 Å and 4216 Å.

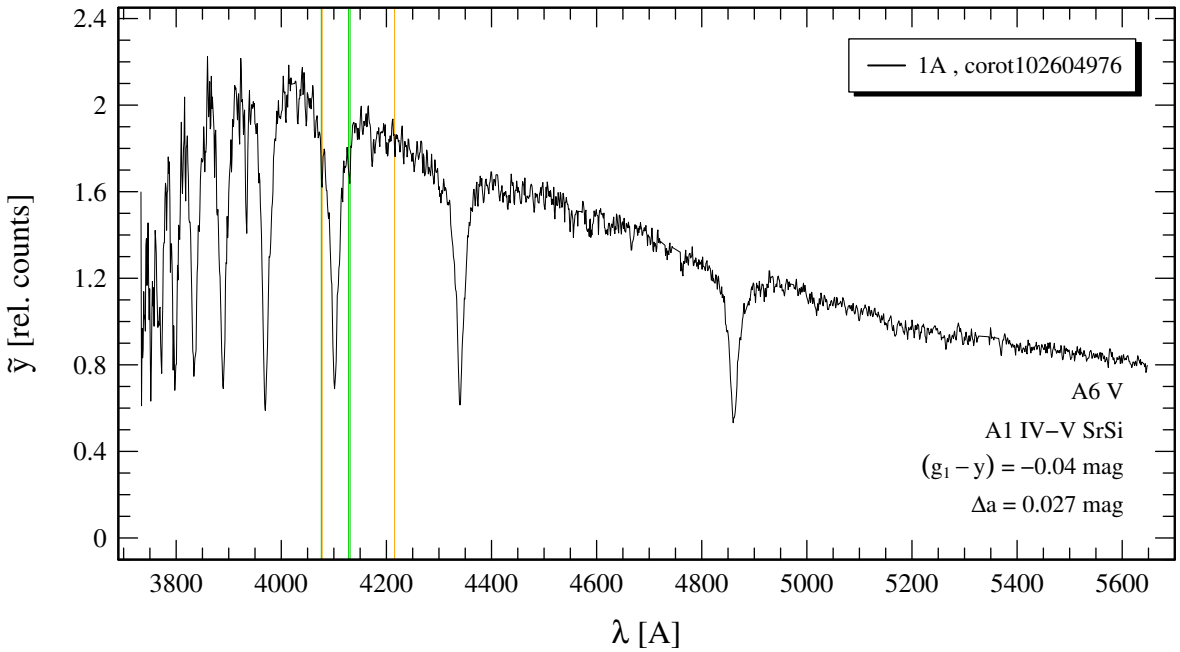

Figure 5.20: The spectrum of corot102604976. One can spot a slight Δa flux depression, Sr II and Si II spectral lines.

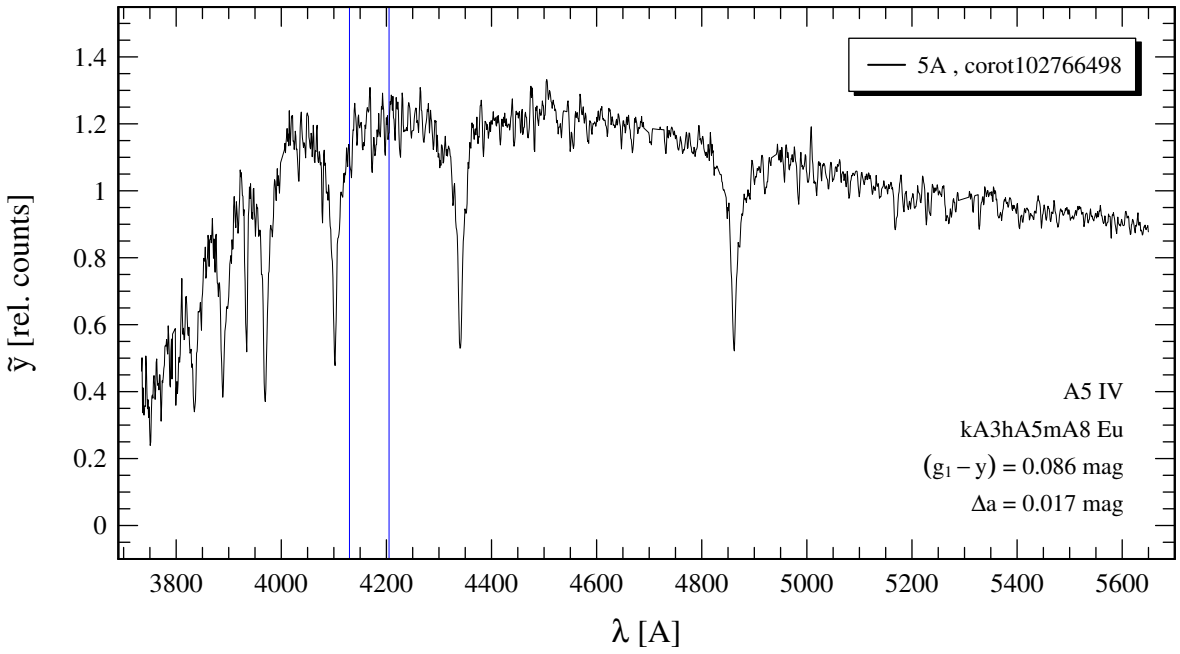

Figure 5.21: The spectrum of corot102766498. MKCLASS classifies this star as Am-type star (CP1) with Eu overabundances and a faint Δa depression appears.

5.2.3 F-type stars

Similar to the F-type stars in the IRa1 sample almost no stars with $|\Delta a| > 2\sigma_{\Delta a}$ are actual CP stars. Only 5 of those stars are CP stars according to MKCLASS, but 3 of them are assigned to the wrong spectral type and another one is heavily polluted by bad counts. The remaining 5th star (15F) is possibly Sr overabundant, according to MKCLASS.

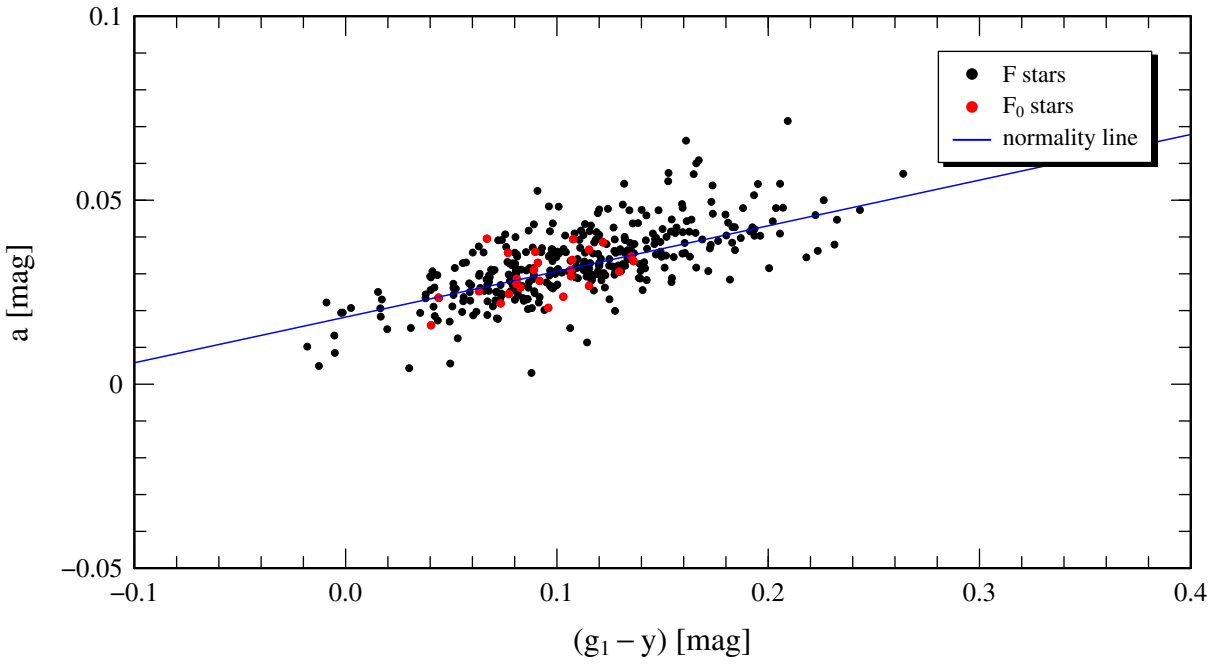

Figure 5.22: Same as in fig. 5.13 for F-type stars.

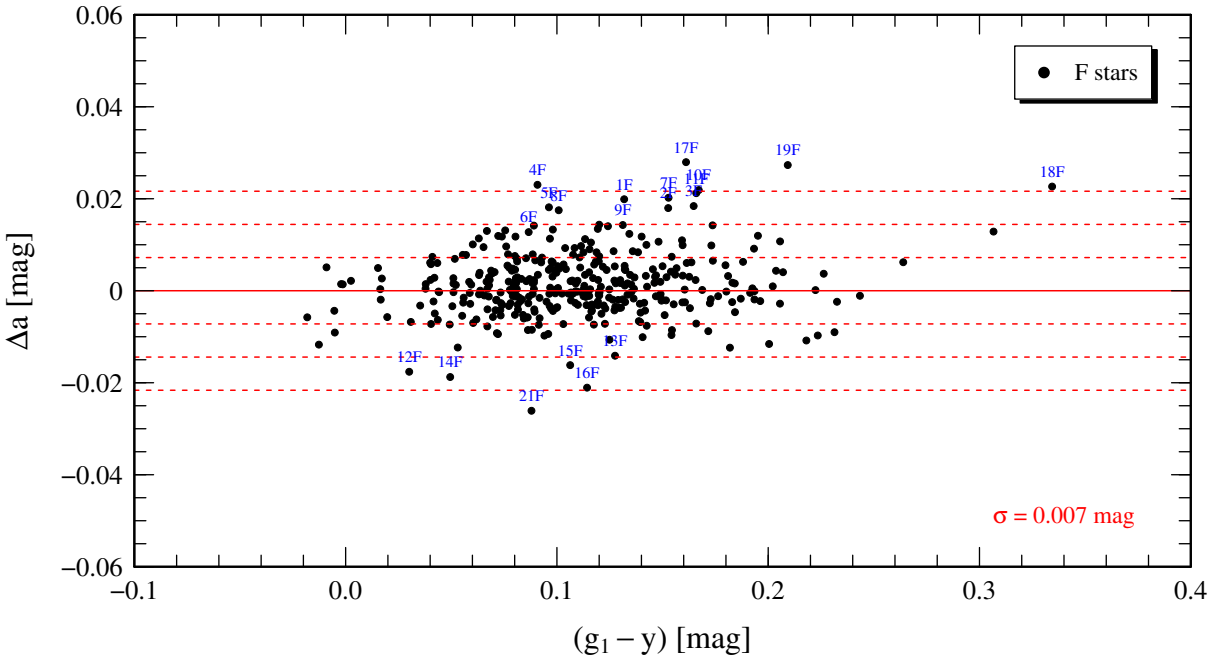

Figure 5.23: Same as in fig. 5.14 for F-type stars.

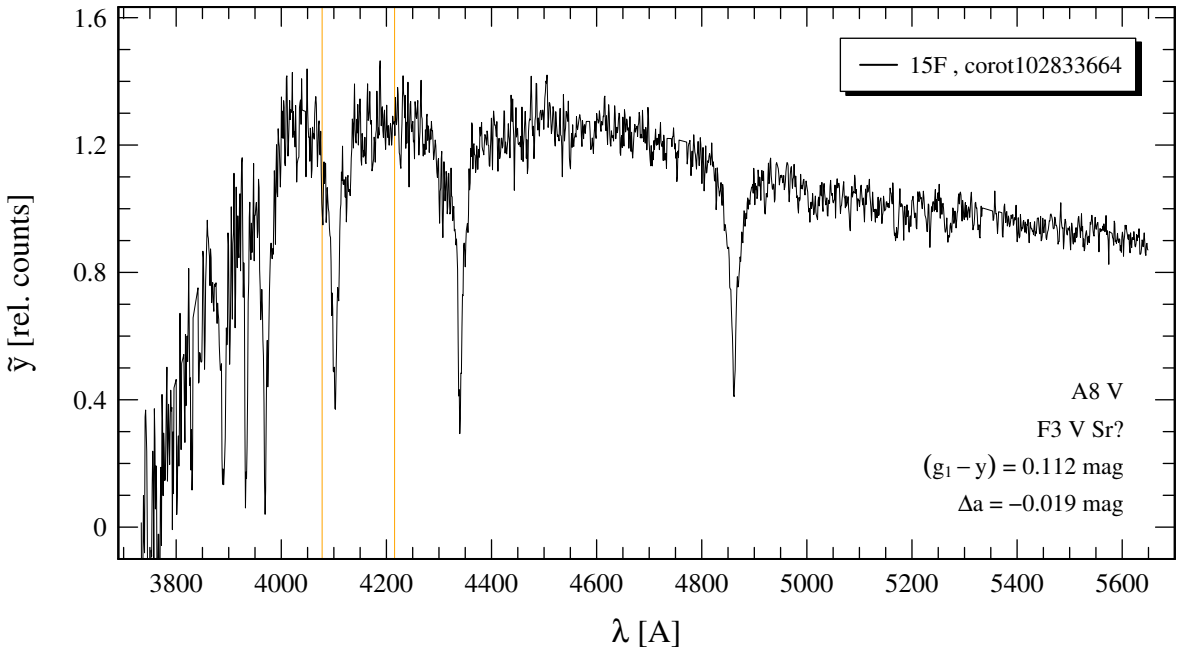

Figure 5.24: The spectrum of corot102833664. This spectrum has a low S/N, a Δa excess and therefore its CP2 membership is highly questionable.

5.3 LRa2 sample

5.3.1 B-type stars

The LRa2 sample holds the largest number of stars and therefore the highest number of CP candidates. In case of B-type stars the Δa technique detects four chemically peculiar stars, three CP2 stars (1B, 2B, 3B) and one λ Bootis star (4B).

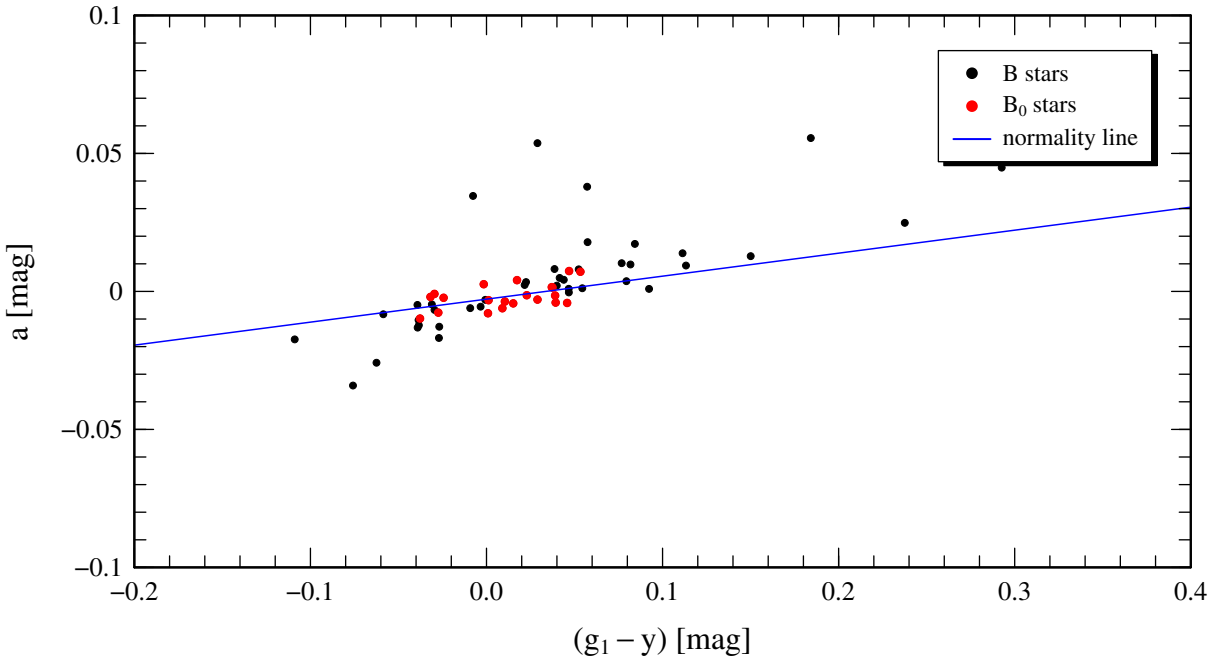

Figure 5.25: The a - color diagram of all LRa2 B-type stars with luminosity class of IV-V or V, plus all λ Bootis stars and spectra with "metal rich" or "metal weak" suffixes in the same spectral range.

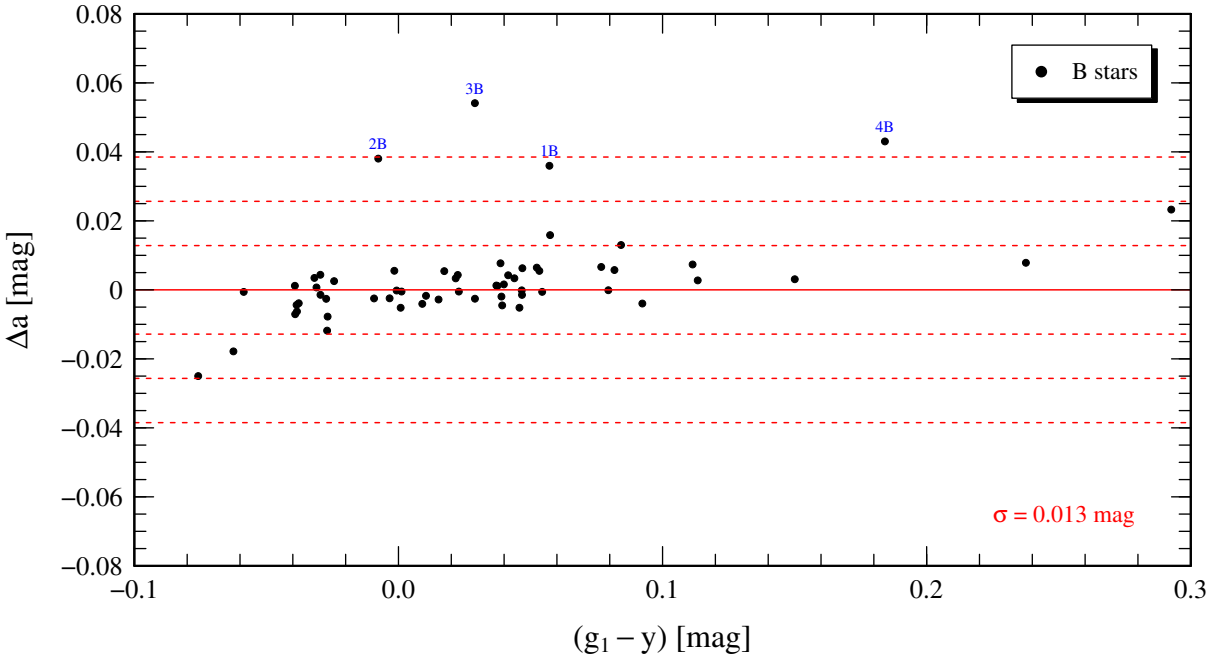

Figure 5.26: The Δa - color diagram of all LRa2 B-type stars with luminosity class of IV-V or V, plus all λ Bootis stars and spectra with "metal rich" or "metal weak" suffixes in the same spectral range. The horizontal lines represent ± 1, ± 2 and $\pm 3\sigma_{\Delta a}$ ranges of the underlying sample of spectra.

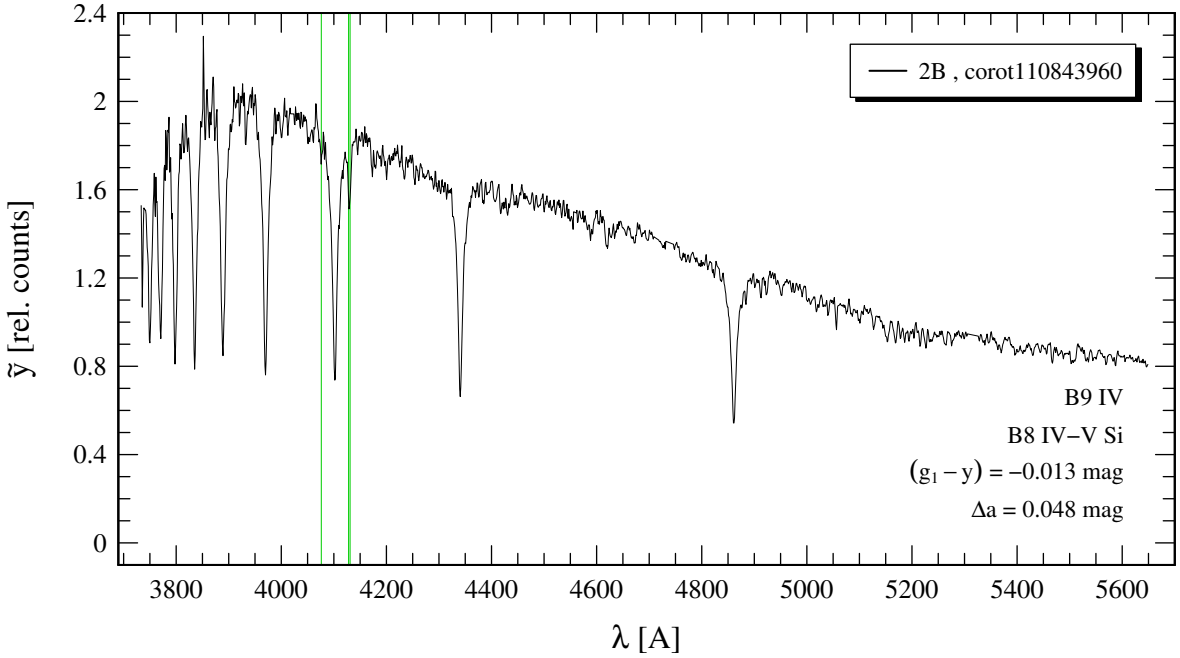

Figure 5.27: The spectrum of corot110843960. Distinct Si II doublet and Δa depression can be seen.

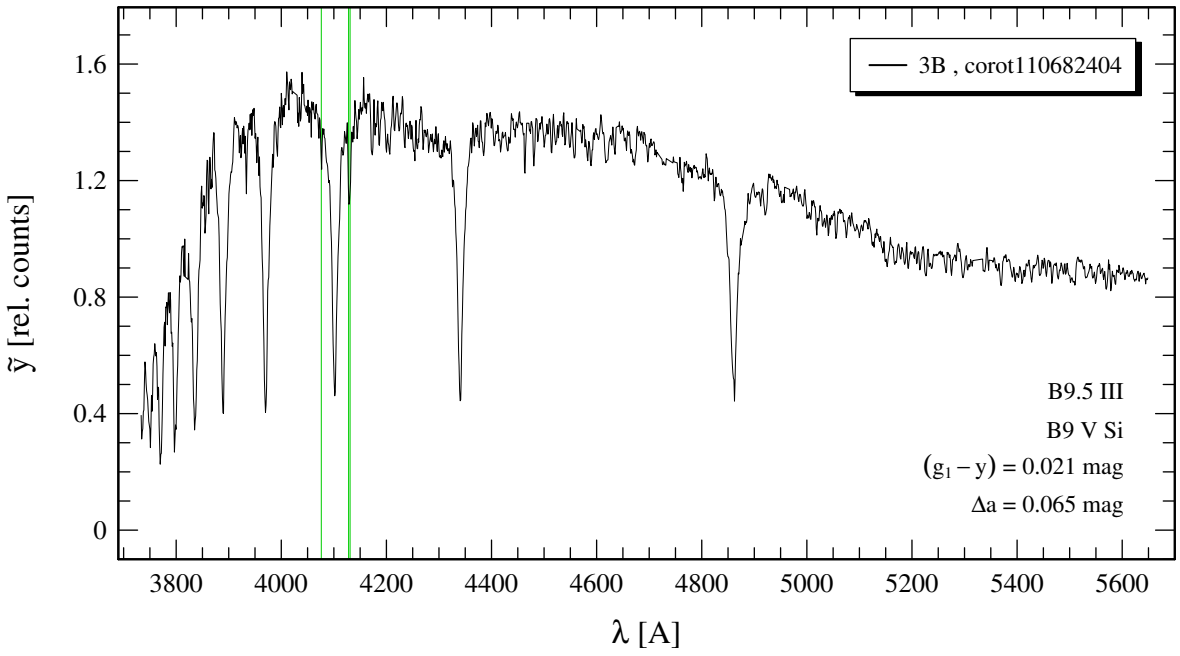

Figure 5.28: The spectrum of corot110682404. In this spectrum a huge Δa depression occurs and it suggests that there might be a Δa and a wave-like pattern acting at the same time, when following the continuum from about 4400 Å towards the red end. The strong Si II line at 4076 Å and the Si II doublet at $4128 - 4131$ Å confirm that this is a CP2 star.

5.3.2 A-type stars

Unlike for the B-type stars, the Δa detections contain a variety of CP and non CP spectra: doubtful and distinct CP stars, spectra with wave-like patterns, non CP stars and spectra affected by the calibration bias.

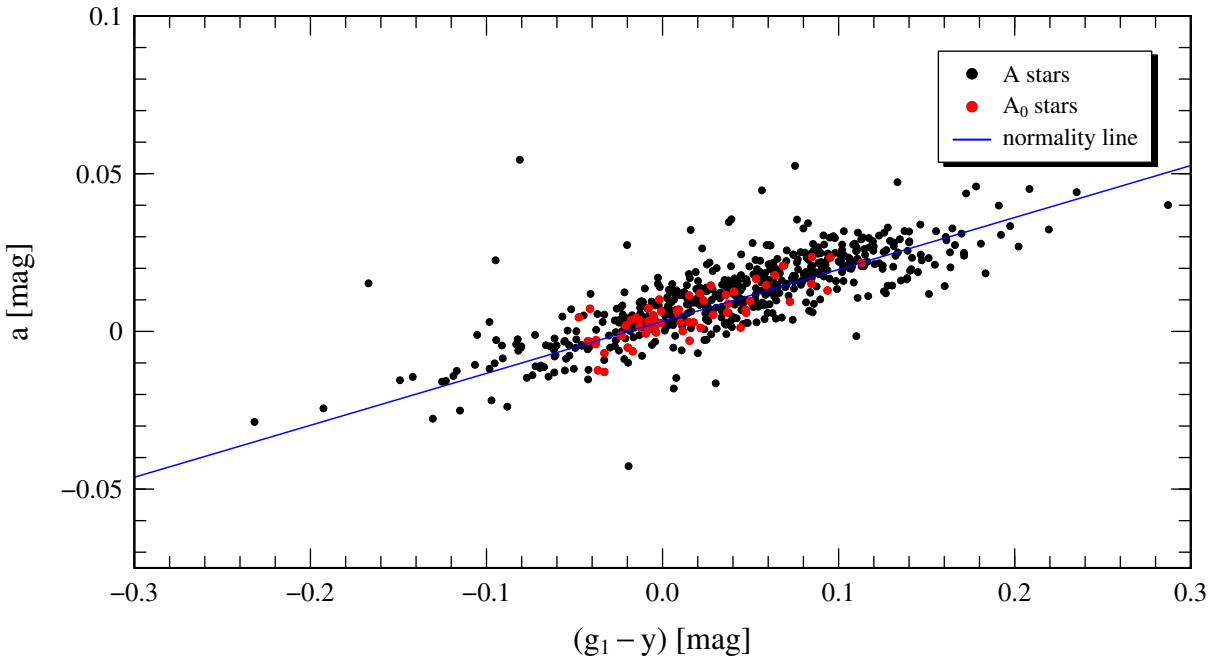

Figure 5.29: Same as in fig. 5.25 for A-type stars.

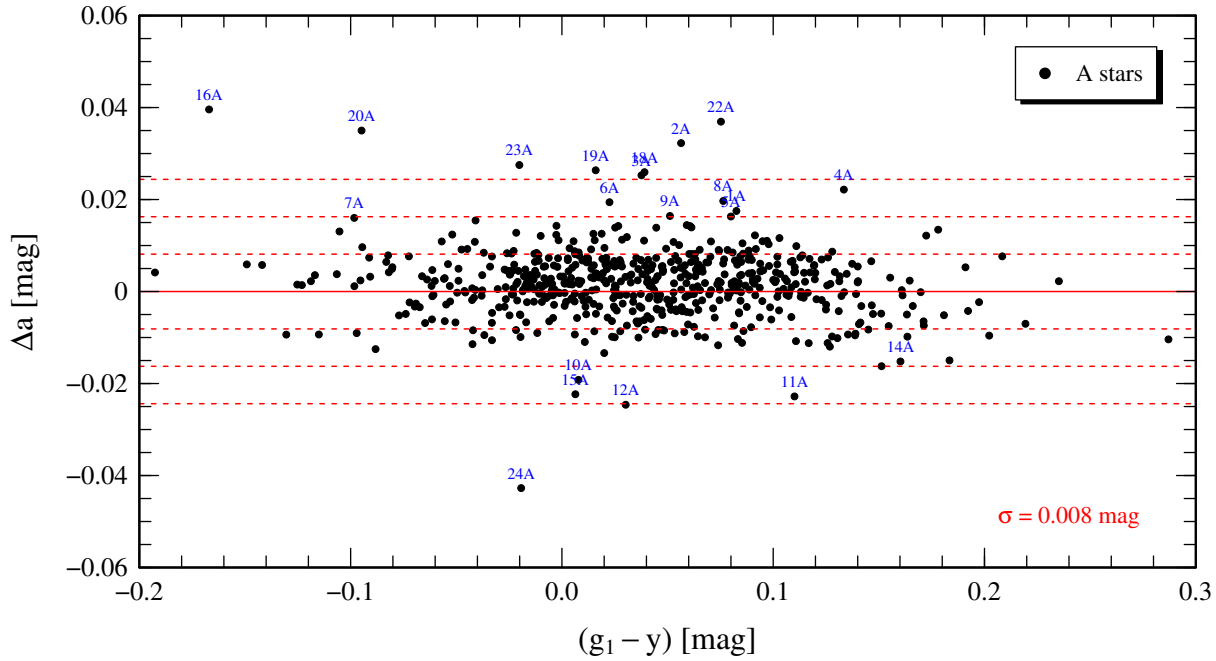

Figure 5.30: Same as in fig. 5.26 for A-type stars.

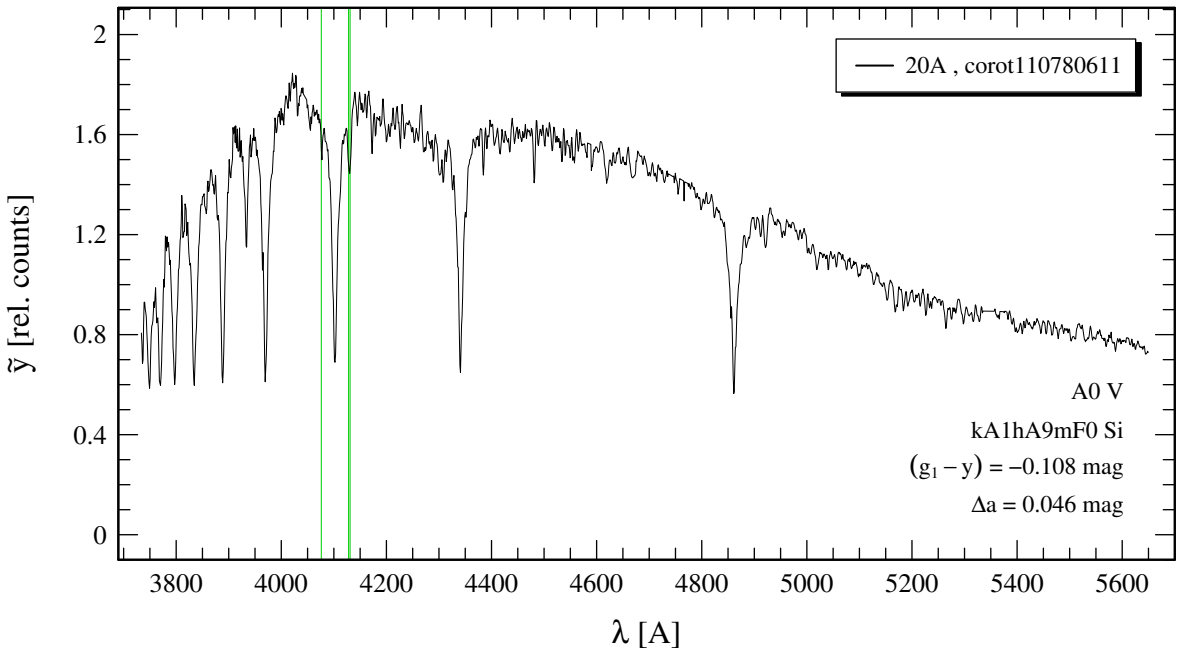

Figure 5.31: The spectrum of corot110780611. This star appears to be an Am-type star (CP1), because the calcium K line suggests an early A-type star, while an evolving G-band can be spotted (early F-type star). Apart from that strong Si absorption lines and a distinct Δa flux depression are visible and suggesting that it is a CP2 star.

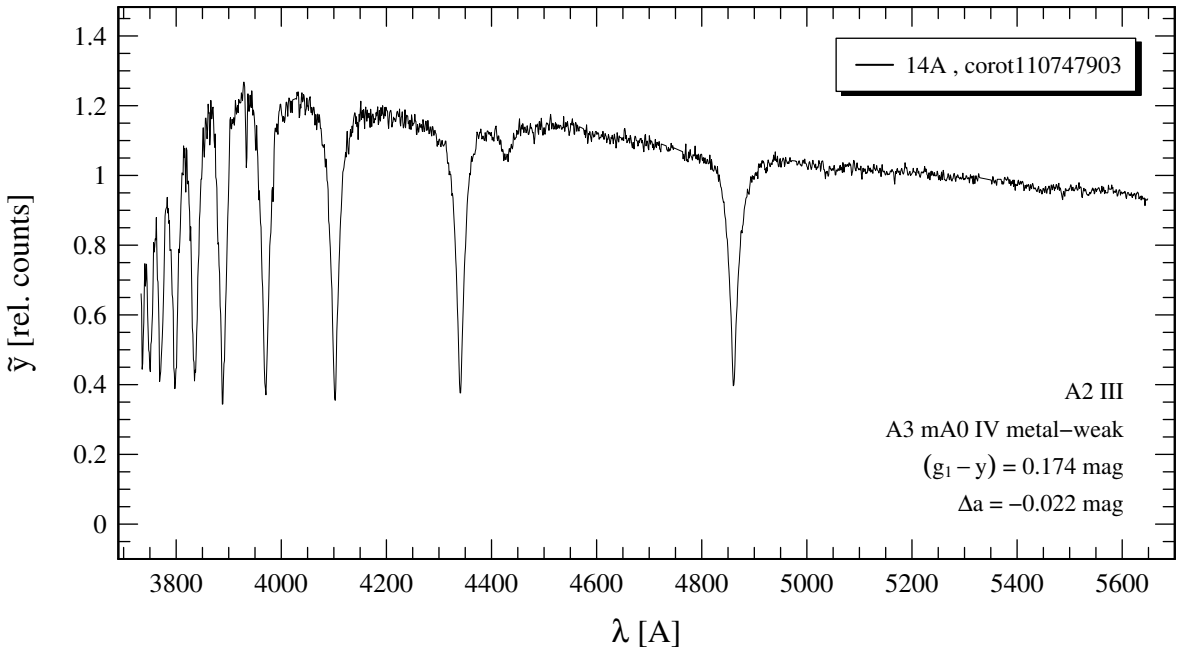

Figure 5.32: The spectrum of corot110747903. This star hardly shows any metals at all, hence the suffix "metal-weak" from MKCLASS.

5.3.3 F-type stars

As we have seen already in the IRa1 and LRa1 samples of F-type stars, most of the Δa detections lack of prominent chemically peculiar spectral features. Apart from a few metal rich stars only 5 CP2 stars are detected via MKCLASS (25F, 58F, 62F, 64F, 84F).

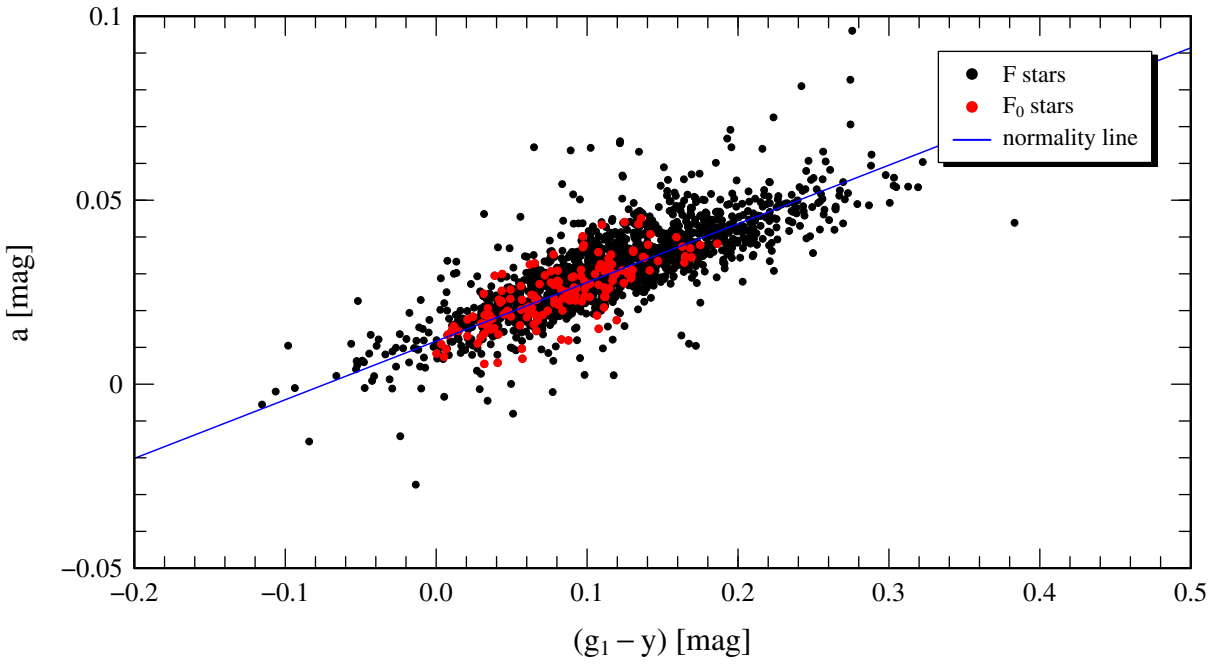

Figure 5.33: Same as in fig. 5.25 for F-type stars.

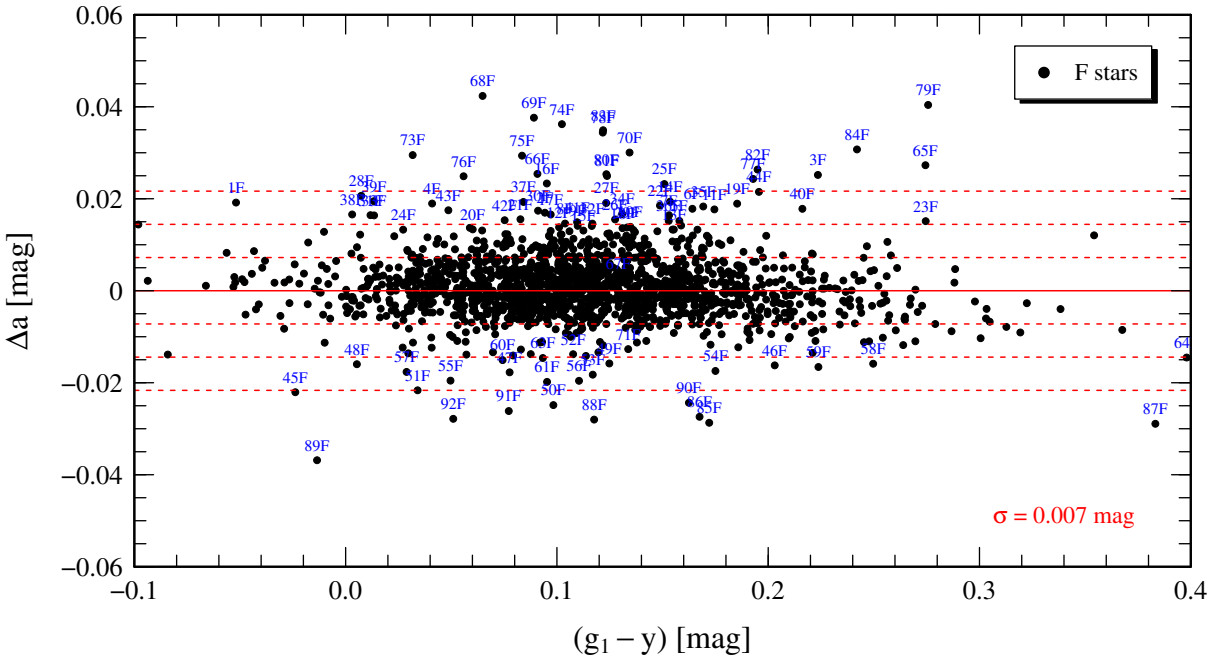

Figure 5.34: Same as in fig. 5.26 for F-type stars.

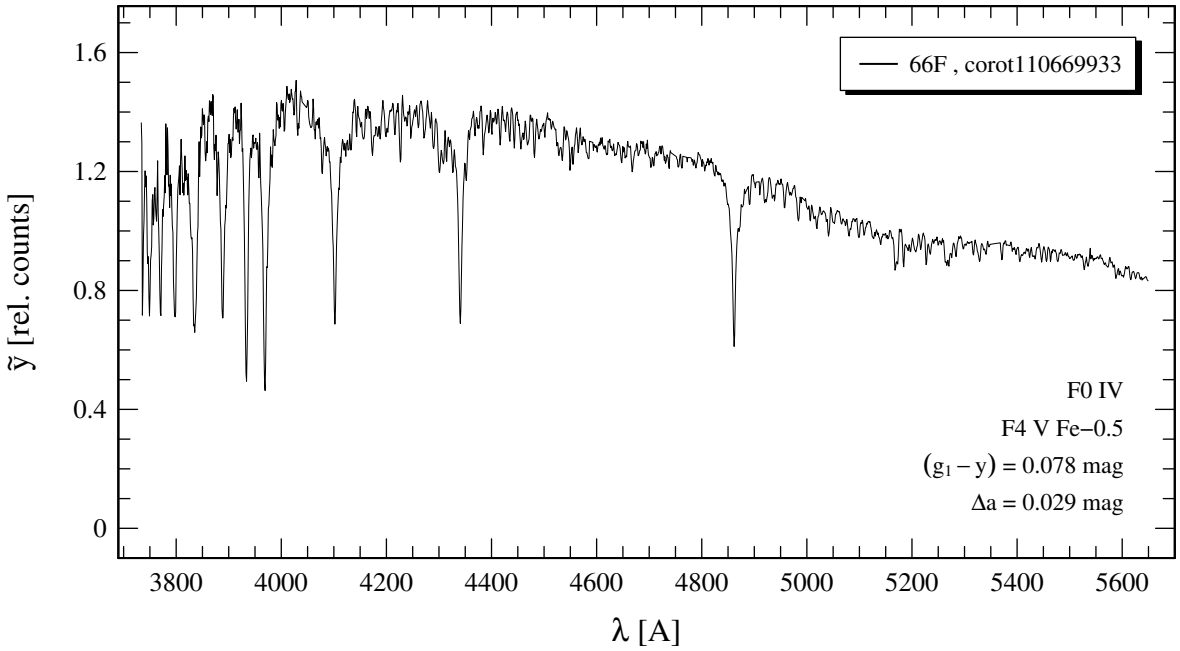

Figure 5.35: The spectrum of corot110669933. A strong Δa depression can be seen, but apart from that no CP2 spectral features are detected. However, no clear wave-like pattern appears.

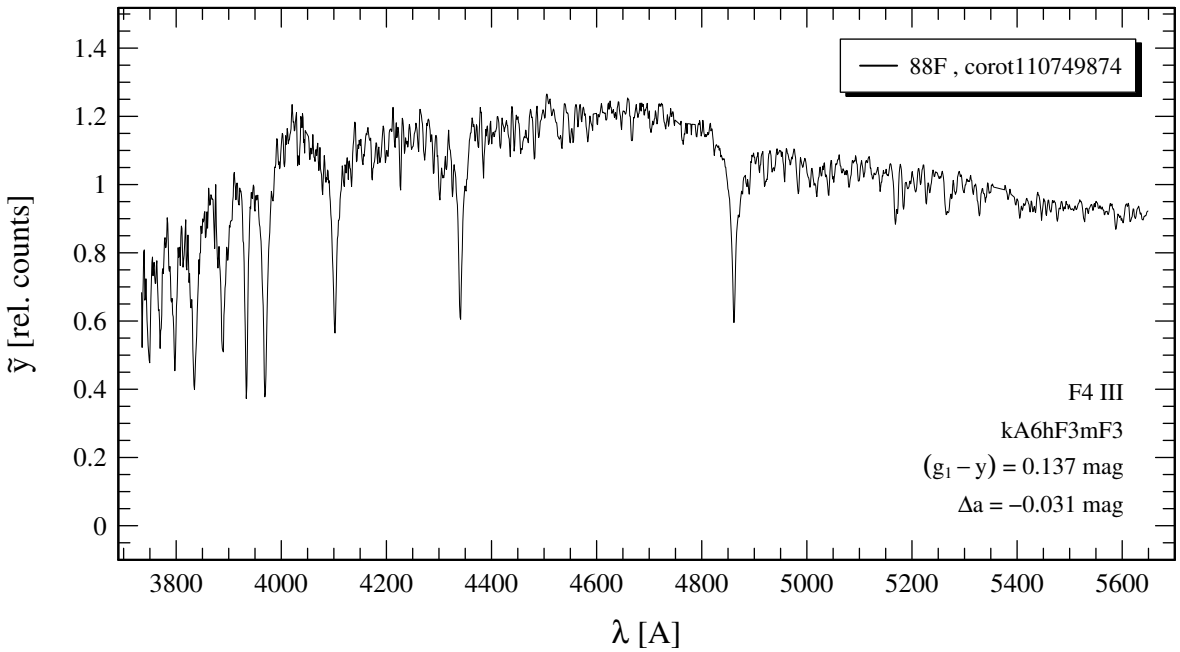

Figure 5.36: The spectrum of corot110749874.

6 Summary

In my Master's Thesis I analyzed 14468 precalibrated spectra obtained from Sebastian et al. (2012) from stars in the CoRoT fields IRa1 (4000 spectra), LRa1 (1659 spectra) and LRa2 (8809 spectra). I recalibrated the IRa1 sample spectra and removed spectra from all three samples, which provided low signals (section 3.3), too many negative counts (section 3.4) or rectangular-like patterns (section 3.5). After that I removed systematically and non systematically occurring bad counts from the remaining spectra by reconstructing the H_β absorption line and interpolating all remaining cut out counts, which reduced the samples' sizes down to 2867 (IRa1), 1525 (LRa1) and 7533 (LRa2) spectra. In the next step I searched for wave-like patterns in the samples by fitting a heuristic function to each spectrum and excluded further spectra which matched this function, cutting down on the samples' sizes to 2671 (IRa1), 1456 (LRa1), 7162 (LRa2). Since many stars have been observed several times by Sebastian et al. (2012), I always picked one out of multiple spectra and removed the others from the samples, which leads to a sample size of 2154 (IRa1), 1297 (LRa1) and 3989 (LRa2) spectra. In the final step of data reduction I computed the radial velocities according to an individual spectral line list with respect to their rest frames for the sample stars (if possible) and shifted their spectral counts distributions to the corresponding rest frames. Afterwards I determined accurate spectral types via `MKCLASS` (Gray & Corbally 2014) for all remaining stars (IRa1: 2144, LRa1: 1289, LRa2: 3974), computed the required magnitudes for the Δa analysis of two filter systems and selected reference spectra for the normality lines. In the last step I calculated Δa and $\Delta a'$ magnitudes for each subsample of B-, A- and F-type stars per CoRoT field and examined the $|\Delta a| > 2\sigma_{\Delta a}$ and $|\Delta a'| > 2\sigma_{\Delta a'}$ outliers from the respective normality lines to verify the CP stars and reject false detections. By doing so I found 15 stars which are most likely chemically peculiar stars and further 158 stars, which are possible chemically peculiar stars or chemically peculiar star candidates, see table 6.1.

field	index	CoRoT id	Sp_1	Sp_2	$\Delta a/\sigma_{\Delta a}$	$\Delta a'/\sigma_{\Delta a'}$	CP type	CP flag
IRa1	1B	102966594	A0 III	B9.5 V Si	4.76	4.74	2	a
IRa1	1A	102769760	A5 IV	A8 V Si	2.26	2.41	2	a
IRa1	2A	102817153	A4 III	A1 IV-V	2.47	2.72	0	c
IRa1	3A	102825399	F4 III	kA2hA7mF0 Eu	2.41	2.48	1	b
IRa1	4A	102831489	A9 IV	kA7hA9mF0	2.06	2.10	1	b
IRa1	5A	102856489	F1 V	kA2hA9mF2 Si	2.12	2.01	1	b
IRa1	6A	102873405	A7 V	A7 Vn	2.57	2.89	2	b
IRa1	7A	102911074	F4 III	kA5hA9mF0	2.04	1.82	1	b
IRa1	8A	102918375	A5 IV	A7 Vn	2.25	1.95	2	b
IRa1	9A	102921874	A5 IV	A7 V Eu	2.27	1.99	2	b
IRa1	10A	102983355	F0 V	A7 Vn	2.50	2.30	2	b
IRa1	11A	102988300	A5 IV	kA5hA9mF3 SiEu	2.23	2.34	1	b
IRa1	12A	102962047	B8 V	A9 mA0 V Lam-Boo	3.36	2.93	5	b
IRa1	13A	102830877	A1 Ib	kA0hA1mA2	-2.83	-2.72	1	b
IRa1	14A	102896075	A6 V	A1 V	6.96	7.80	0	c
IRa1	15A	102898310	F0 IV	A9 V Eu	5.13	4.82	2	a
IRa1	16A	102798478	A5 IV	kA2hA5mA8	-3.83	-3.69	1	b
IRa1	17A	102991179	A8 V	A9 mA4 V Lam-Boo	-4.28	-4.68	5	a
IRa1	1F	102697754	F9 IV	F8 IV-V	2.54	2.08	2	b
IRa1	2F	102700263	F4 V	F6 V	2.51	2.19	2	b
IRa1	3F	102720345	G2 IV	F9 V	2.31	1.81	2	b
IRa1	4F	102758435	F9.5 V	F9 IV-V	2.26	2.01	2	b
IRa1	5F	102774699	F7 IV	F5 V	2.15	2.55	2	b
IRa1	6F	102780079	F8 V	F8 V	2.02	1.93	2	b
IRa1	7F	102798231	F8 V	F9 V	2.45	2.34	2	b
IRa1	8F	102828903	G2 IV	F9 V	2.06	1.54	2	b
IRa1	9F	102829039	F9.5 V	F9 V	2.24	1.88	2	b
IRa1	10F	102831226	F9.5 V	F9 IV-V	2.11	1.89	2	b
IRa1	11F	102844674	G0 V	F9 V	2.15	1.76	2	b
IRa1	12F	102856147	F7 IV	F6 V	2.07	1.39	2	b
IRa1	13F	102863995	F9 IV	F9 V	2.17	2.12	2	b
IRa1	14F	102864379	G0 V	F9 IV-V	2.10	1.66	2	b
IRa1	15F	102913695	F8 V	F9 V	2.03	1.61	2	b
IRa1	16F	102922787	G0 V	F9 V	2.76	2.10	2	b
IRa1	17F	102947113	G2 IV	F9 V	2.00	1.90	2	b
IRa1	18F	102969618	G1 V	F8 IV-V	2.60	2.23	2	b
IRa1	19F	102971306	F9 IV	F8 IV-V	2.15	1.71	2	b
IRa1	20F	102985692	F8 V	F9 IV-V	2.16	1.83	2	b
IRa1	21F	102723949	F7 IV	F7 V	3.18	2.81	2	b
IRa1	22F	102754035	F6 V	F6 V	1.99	2.24	2	b
IRa1	23F	102850896	F7 V	F7 V	3.09	2.59	2	b
IRa1	24F	102874709	G4 V	F9 V	3.46	2.87	2	b
IRa1	25F	102719294	F4 III	F3 V	-2.84	-3.21	1	b
IRa1	26F	102945448	A9 III	F1 V	-2.68	-1.95	1	b
IRa1	27F	102955422	F6 V	F6 V	-2.85	-2.89	1	b
IRa1	28F	102941727	F0 IV	F3 V Fe-0.8	-1.99	-2.03	1	b

IRa1	29F	102791241	F7 IV	F5 V	4.19	3.84	0	c
IRa1	30F	102795487	F5 IV	F5 V	3.96	3.71	0	c
IRa1	31F	102814603	G1 V	F9 IV-V	3.47	3.10	2	b
IRa1	32F	102817430	F8 V	F9 V	4.84	4.88	2	b
IRa1	33F	102869824	F3 V	F5 V Fe-0.5	5.94	5.32	0	c
IRa1	34F	102936636	NA NA	F6 V	7.02	4.94	0	c
IRa1	35F	102958468	F7 IV	F5 V	5.34	4.76	2	b
IRa1	36F	102974775	F9 IV	F8 V	4.56	4.13	0	c
IRa1	37F	102917296	G2 IV	F9 V	-10.78	-14.51	0	c
LRa1	1B	102574584	F8 V	F7 II-III	2.98	2.92	0	c
LRa1	2B	102835814	B9 V	B8 IV-V Si	2.17	2.08	2	a
LRa1	3B	102780040	A4 III	B9 V Si	3.76	3.78	2	a
LRa1	1A	102604976	A6 V	A1 IV-V SrSi	2.98	2.77	2	a
LRa1	2A	102744681	A6 V	A1 IV-V SrSi	2.69	2.88	2	b
LRa1	3A	102683983	A9 IV	A7 Vn	1.88	2.07	0	c
LRa1	4A	102748440	A7 V	A9 V Eu	1.92	2.00	2	b
LRa1	5A	102766498	A5 IV	kA3hA5mA8 Eu	1.89	2.13	1	b
LRa1	6A	102778314	A2 IV	A0 IV-V Sr?Si	1.91	2.26	2	a
LRa1	7A	102838310	B9 V	F1 mA0 V Lam-Boo	-2.46	-2.47	5	b
LRa1	8A	102710394	A0 IV	kB9.5hA1mA2	-1.99	-2.21	0	c
LRa1	9A	102563684	A6 V	kA4hA6mK0	4.22	4.33	1	b
LRa1	10A	102584289	F4 III	kA6hA9mF3	3.69	3.76	1	b
LRa1	11A	102795463	B9 V	kA0hA1mA2 Si	3.36	3.38	2	a
LRa1	12A	102780330	A3 V	A1 IV-V Sr?	-3.23	-3.15	2	b
LRa1	1F	102549040	F9 IV	F9 V	2.71	2.76	2	b
LRa1	2F	102630967	G0 V	F9 IV-V	2.60	2.49	2	b
LRa1	3F	102637311	F9 V	F9 IV-V	2.51	2.55	2	b
LRa1	4F	102724668	F4 III	F2 V	2.61	3.19	2	b
LRa1	5F	102725947	F6 V	F6 V Sr?	2.18	2.52	2	b
LRa1	6F	102739590	F9.5 V	F9 V	2.16	1.77	2	b
LRa1	7F	102768056	F7 V	F8 IV-V	2.50	2.80	2	b
LRa1	8F	102788790	G2 IV	F9 V	2.22	2.43	2	b
LRa1	9F	102829324	F9.5 V	F9 V	2.15	1.98	2	b
LRa1	10F	102829349	F7 IV	F6 V	2.97	3.04	2	b
LRa1	11F	102807343	F9.5 V	F9 V	3.15	2.94	2	b
LRa1	12F	102562563	F8 V	F9 IV-V	-2.17	-2.44	1	b
LRa1	13F	102598954	B9 V	kA0hF0mG0	-2.37	-1.96	0	c
LRa1	14F	102623451	F5 V	F7 IV-V	-2.37	-2.60	1	b
LRa1	15F	102833664	A8 V	F3 V Sr?	-2.15	-2.24	2	b
LRa1	16F	102788427	NA NA	F6 V Fe-0.9	-3.10	-2.92	0	c
LRa1	17F	102612430	G0 V	F9 V	3.89	3.88	2	b
LRa1	18F	102714636	G0 V	F9 V Sr	3.00	3.14	2	b
LRa1	19F	102751765	F7 IV	F5 V	3.92	3.79	2	b
LRa1	20F	102835912	G4 V	kA0hF8mF8 Sr	4.23	3.76	0	c
LRa1	21F	102769848	F1 V	F2 V	-3.49	-3.62	1	b
LRa2	1B	110760937	A0 V	B8 V Si	2.77	2.80	2	a
LRa2	2B	110843960	B9 IV	B8 IV-V Si	3.18	2.96	2	a
LRa2	3B	110682404	B9.5 III	B9 V Si	4.27	4.22	2	a

LRa2	4B	110683580	A0 III	F1 mB7 V Lam-Boo	3.17	3.35	5	a
LRa2	1A	110660141	A5 IV	kA7hA9mF1	2.04	2.15	1	b
LRa2	2A	110744388	F4 III	kA5hA9mF2 Si	2.93	3.97	1	b
LRa2	3A	110751459	A6 V	A6 V	2.64	3.11	0	c
LRa2	4A	110830595	A8 V	F0 mA5 V metal-weak	2.24	2.73	6	b
LRa2	5A	110834647	A7 V	kA2hA5mA8	2.03	2.01	1	b
LRa2	6A	110849901	A6 V	kA2hA7mA6	2.44	2.39	1	b
LRa2	7A	110857541	A5 IV	kA2hA7mF1 SrSi	2.15	1.97	1	b
LRa2	8A	110661913	NA NA	kA1hA4mA8	1.92	2.42	1	b
LRa2	9A	110677136	F4 III	kA2hA6mA9	1.90	2.02	1	b
LRa2	10A	110673703	B6 V	F0 mA2 IV-V metal-weak	-2.43	-2.36	6	b
LRa2	11A	110675609	A8 V	A9 V	-2.68	-2.81	0	c
LRa2	12A	110691459	A6 V	A2 V	-2.74	-3.03	0	c
LRa2	13A	110745154	A2 II	kA2hA5mA8	-2.19	-2.06	1	b
LRa2	14A	110747903	A2 III	A3 mA0 IV metal-weak	-2.21	-1.87	6	b
LRa2	15A	110841774	A8 V	A8 Vn	-1.84	-2.75	0	c
LRa2	16A	102983551	A5 IV	kA1hA8mF0 Eu	4.96	4.87	1	b
LRa2	17A	110658353	A6 V	A5 IV-V	8.54	7.95	0	c
LRa2	18A	110676985	A5 IV	kA5hA7mA9	3.22	3.19	1	b
LRa2	19A	110765453	A0 IV	A1 IV-V	3.52	3.24	0	c
LRa2	20A	110780611	A0 V	kA1hA9mF0 Si	4.67	4.31	2	b
LRa2	21A	110830407	A4 V	kA1hA2mA4	5.22	5.34	0	c
LRa2	22A	110834454	A1 V	A0 V	4.09	4.54	0	c
LRa2	23A	110863788	A6 V	A2 IV-V SrSi	3.26	3.38	2	a
LRa2	24A	110694412	A0 V	A0 V	-5.34	-5.26	0	c
LRa2	1F	102990364	F4 III	F2 V	2.57	2.66	2	b
LRa2	2F	103036346	F7 V	F9 V Fe-0.5	2.31	2.35	2	b
LRa2	3F	110657539	F3 V	F5 V	2.95	3.49	2	b
LRa2	4F	110658250	G4 V	F9 V	2.72	2.63	2	b
LRa2	5F	110667213	G4 V	F8 IV-V	2.27	2.10	2	b
LRa2	6F	110670090	F4 V	F5 V Fe-0.6	2.04	2.47	0	c
LRa2	7F	110672405	F7 V	F8 IV-V	2.53	2.27	2	b
LRa2	8F	110673171	F7 IV	F6 V	2.03	2.02	2	b
LRa2	9F	110678927	F9 IV	F9 V	2.06	1.95	2	b
LRa2	10F	110680245	G0 V	F9 V	2.07	1.92	2	b
LRa2	11F	110682141	G0 IV	F9 IV-V	2.96	2.45	2	b
LRa2	12F	110686281	G0 V	F8 V	2.25	1.91	2	b
LRa2	13F	110688326	G0 V	F9 V	2.08	1.84	2	b
LRa2	14F	110689441	G0 IV	F9 V	2.18	2.68	2	b
LRa2	15F	110690837	G2 IV	F6 V	2.09	1.81	2	b
LRa2	16F	110691402	F8 V	F9 V	2.99	3.23	2	b
LRa2	17F	110691975	F9.5 V	F9 V	2.50	2.29	2	b
LRa2	18F	110743696	G0 V	F9 V	2.01	1.88	2	b
LRa2	19F	110744123	F9 IV	F8 IV-V	2.92	2.62	2	b
LRa2	20F	110744984	G0 V	F9 V	2.15	1.85	2	b
LRa2	21F	110753398	G4 V	F9 IV-V	2.42	2.15	2	b
LRa2	22F	110757396	F3 V	F5 V	2.57	2.57	2	b
LRa2	23F	110760165	F7 IV	kA2hF6mF8	2.44	2.10	2	b

LRa2	24F	110763666	F6 V	F7 V	2.19	1.84	2	b
LRa2	25F	110773049	F5 V	F5 V Sr	2.50	3.21	2	b
LRa2	26F	110777372	G2 IV	F9 V	2.47	2.15	2	b
LRa2	27F	110777507	F7 IV	F3 V	2.26	2.64	2	b
LRa2	28F	110779757	G1 V	F9 IV-V	2.87	2.86	2	b
LRa2	29F	110828202	F9.5 V	F9 IV-V	2.01	1.91	2	b
LRa2	30F	110838988	G0 V	F9 V	2.46	2.41	2	b
LRa2	31F	110842257	F7 V	F8 IV-V	2.12	2.28	2	b
LRa2	32F	110843991	G0 V	F9 IV-V	2.13	2.02	2	b
LRa2	33F	110845420	F9 IV	F8 V	2.38	2.27	2	b
LRa2	34F	110849305	F9 IV	F7 V	2.21	2.31	2	b
LRa2	35F	110850324	F7 V	F7 IV-V	2.95	2.54	2	b
LRa2	36F	110852179	G1 V	F9 V	2.21	2.10	2	b
LRa2	37F	110858518	F7 IV	F6 V	2.93	2.67	2	b
LRa2	38F	110860381	F0 V	F1 V	2.40	2.30	0	c
LRa2	39F	110862438	G0 V	F9 IV-V	2.61	2.70	2	b
LRa2	40F	110864413	G0 V	F8 IV-V	2.56	2.47	2	b
LRa2	41F	110865641	G2 V	F9 IV-V	2.00	2.06	2	b
LRa2	42F	110681955	F4 III	F2 V	1.68	2.12	2	b
LRa2	43F	110751598	F1 V	F1 V	1.86	2.42	2	b
LRa2	44F	110755046	G0 V	F9 IV-V	3.18	2.98	2	b
LRa2	45F	102891109	F0 V	F2 V	-2.84	-3.06	1	b
LRa2	46F	110661209	F7 IV	F2 V	-2.41	-2.25	1	b
LRa2	47F	110663736	F4 III	F3 V	-2.10	-2.46	1	b
LRa2	48F	110668182	NA NA	kA0hF5mK2	-2.47	-2.21	0	c
LRa2	49F	110668654	F3 V	F5 V	-2.20	-2.19	1	b
LRa2	50F	110675161	F3 V	F5 V	-2.96	-3.44	1	b
LRa2	51F	110684354	F4 III	F4 IV-V Fe-0.7	-2.61	-3.00	1	b
LRa2	52F	110693837	F4 III	F3 V	-2.07	-1.91	1	b
LRa2	53F	110757622	F4 V	F5 V Fe-0.6	-2.10	-2.53	1	b
LRa2	54F	110758150	F5 IV	F4 V	-2.43	-2.42	1	b
LRa2	55F	110764421	F0 V	F1 V	-2.44	-2.71	1	b
LRa2	56F	110775963	F7 IV	F3 V	-2.30	-2.71	1	b
LRa2	57F	110777793	F4 III	F5 V Fe-0.6	-2.37	-2.45	1	b
LRa2	58F	110830830	F7 IV	F4 V Sr	-2.22	-2.20	1	b
LRa2	59F	110834599	F7 IV	F6 V	-2.43	-2.30	1	b
LRa2	60F	110851908	F0 IV	F3 V Fe-0.7	-2.13	-2.09	1	b
LRa2	61F	1110849884	NA NA	F0 V	-2.81	-2.74	1	b
LRa2	62F	110684641	F2 III	F5 V Sr?	-1.78	-2.20	1	b
LRa2	63F	110765783	F4 III	F3 V	-1.89	-2.03	1	b
LRa2	64F	110834263	F1 V	F2 V Sr	-1.81	-2.01	1	b
LRa2	65F	110668307	F7 IV	F6 V	3.74	3.78	2	b
LRa2	66F	110669933	F0 IV	F4 V Fe-0.5	3.17	3.52	2	b
LRa2	67F	110673408	G4 V	F9 V	3.55	0.36	0	c
LRa2	68F	110686860	F8 V	F5 V	5.10	5.87	0	c
LRa2	69F	110694880	F9 IV	F9 V	4.18	5.21	0	c
LRa2	70F	110746552	G0 V	F9 IV-V	4.54	4.16	2	b
LRa2	71F	110749548	G4 V	F9 V	4.23	-1.76	0	c

LRa2	72F	110757421	G2 IV	F9 IV-V CN1	5.22	4.97	2	b
LRa2	73F	110760723	F5 IV	F5 V	3.94	4.09	2	b
LRa2	74F	110762535	F8 V	F5 V Fe-0.6	4.68	5.02	0	c
LRa2	75F	110765178	F7 IV	F5 V	4.24	4.07	0	c
LRa2	76F	110767488	G5 III	F9 V	3.80	3.45	2	b
LRa2	77F	110773610	F7 V	F6 V	3.25	3.37	0	c
LRa2	78F	110777272	G1 V	F9 IV-V	3.04	4.76	2	b
LRa2	79F	110779395	F4 III	kA1hF3mF7	5.66	5.60	2	b
LRa2	80F	110838402	F6 V	F6 V Fe-0.6	3.74	3.51	0	c
LRa2	81F	110853609	G4 V	F7 V	3.57	3.46	2	b
LRa2	82F	110857960	G4 V	F9 V	3.43	3.65	2	b
LRa2	83F	110858375	G0 V	F9 IV-V	4.71	4.84	2	b
LRa2	84F	110861330	F7 IV	F5 V Sr?	3.55	4.26	2	b
LRa2	85F	102995134	F7 IV	F6 V	-3.88	-3.98	1	b
LRa2	86F	110659009	B5 III	kA0hF5mK2	-3.96	-3.80	0	c
LRa2	87F	110677158	NA NA	F8 V Fe-0.6 Sr	-3.48	-4.01	1	b
LRa2	88F	110749874	F4 III	kA6hF3mF3	-3.40	-3.88	1	b
LRa2	89F	110763718	B9 IV	F0 mB8 V Lam-Boo	-5.00	-5.11	5	a
LRa2	90F	110773254	B5 V	kB9hF4mK2	-3.53	-3.38	0	c
LRa2	91F	110851270	F3 V	F5 V	-3.43	-3.62	0	c
LRa2	92F	110866082	F6 V	F6 V	-3.83	-3.86	0	c

Table 6.1: The detailed results of Δa and $\Delta a'$ values in all three CoRoT fields. The deviation from the normality line is expressed as fraction of $\sigma_{\Delta a}$ and $\sigma_{\Delta a'}$. Spectral types are taken from Sebastian et al. (2012) and obtained from `MKCLASS`. The CP types range from "1" to "6" with "1" to "4" corresponding to CP1 to CP4 stars, "5" are λ Bootis stars and "6" are other "metal weak" stars, assigned by `MKCLASS`. The CP flag is heuristically determined and its symbolism is: "a" for "very likely", "b" for "possible", "c" for "very unlikely". Please note that "c" also includes not only spectra, where no chemical peculiarities have been found, but also spectra with obviously incorrect spectral types, wave-like patterns, too many bad counts or low signal. Therefore "c" flagged spectra could still be chemically peculiar, but based on the underlying spectra it is hard or even impossible to tell.

7 Discussion

It has been shown by Paunzen et al. (2005) that the efficiency of detecting magnetic CP stars with Δa photometry in a sample can go up to 95%. Based on the results of this Master's Thesis I can not obtain such high efficiency. The reasons for that do not lie in the nature of Δa photometry or synthetic photometry, but in the quality of the spectra and assumed simplifications, which mainly are

1. color bias,
2. neglecting interstellar reddening,
3. wave-like patterns,
4. no check for variabilities of the stars.

(1) It has already been pointed out that there appears to be a "color bias" (sec. 5) leading to a vast spread in colors, but not in a magnitudes, hence shifting the spectra in the a versus color plot only along the color axis, but not in a direction. For instance an intermediate B and CP2 star with a clear Δa flux depression stands out in the B-type star sample, but if its color is much redder due to the "color bias" and corresponds to the color of a F-type star, it will be compared to a higher a_0 magnitude derived from the normality line, so its actual Δa is rather small and therefore considered as "normal". To counter this effect I lower the Δa detection criterion to $|\Delta a| > 2\sigma_{\Delta a}$, but this adds many possibly non peculiar stars to the Δa detection list and lowers the efficiency as a consequence.
(2) Similar to the reasons given above, interstellar reddening shifts a star along the color axis, but only towards red colors. Again, a definite CP star can be mistaken as normal star, if interstellar extinction is considerable.
(3) Since not all spectra with wave-like patterns have been detected and removed from the samples, they either exceed the $|\Delta a| > 2\sigma_{\Delta a}$ threshold by far, if the pattern amplitudes are high (which is the case for few spectra) or they increase $\sigma_{\Delta a}$ since there are many more spectra with low pattern amplitudes, hence increasing the threshold for detecting CP2 stars.
(4) Variable stars and chemically peculiar stars provide similar Δa effects, so one has to

check the sample stars for known variabilities, which I have not done in my Master's Thesis.

Despite these difficulties the 15 stars that have been verified manually are inevitably chemically peculiar stars, but both they and the possible CP stars have to be confirmed spectroscopically still.

8 Acknowledgements

I would like to thank Ernst Paunzen for his valuable support and patience during my work on my Master's Thesis. Also thanks to Richard Gray and Christopher Corbally for clearing up the use of `MKCLASS` to me.
This Master's Thesis has been supported by WTZ (Wissenschaftlich-Technische Zusammenarbeit) by funding accomodation and attendance fee to the extend of 8 days in Brno, Czech Republic from 2012 to 2016.
This Master's Thesis uses `R` Version 3.3.1, `RStudio` Version 0.99.903, `MKCLASS` Version 1.07 and `fits viewer` Version 5.4.

9 References

9.1 Publications

Bessell, M. S. 2005, ARA&A, 43, 293

Conti, P. 1970, PASP, 82, 781

Fossati, L., Bagnulo, S., Landstreet, J., et al. 2008, A&A 483, 891

Gray, R. O., Corbally, C. J. 2014, AJ, 147, 80

Kochukhov, O. 2017, A&A, 597, A58

Kochukhov, O. 2014, "Magnetic fields in A-type stars besides Ap stars" in "Putting A-type stars into Context: Evolution, Environment, and Related Stars"

Maitzen, H. M. 1976, A&A, 51, 223

Paunzen, E., 2004, in IAU Symposium, Vol. 224, The A-Star Puzzle, Zverko, J., Ziznovsky,
J., Adelman, S. J., Weiss, W. W., eds., pp. 443-450

Paunzen, E., Stütz, Ch., Maitzen, H. M. 2005, A&A, 441, 1111

Paunzen, E., Maitzen, H. M. 2017, lecture "Vertiefungsmodul: Pekuliare Sterne (NPI)", University of Vienna

Pence, W. D., Chiappetti, L., Page, C. G., et al. 2010, A&A, 524, A42

Pöhnl, H., Paunzen, E., Maitzen, H. M. 2005, A&A, 441, 1111

Preston, G. W. 1974, ARA&A, 12, 257

Roman, N. G., Morgan, M. M., Eggen, O. J. 1948, ApJ, 107, 107

Sebastian, D., Guenther, E. W., Schaffenroth, V., et al. 2012, A&A, 541, A34

Smalley, B., Southworth, J., Pintado, O. I., et al. 2014, A&A, 564, A69

Smalley, B., Antoci, V., Holdsworth, D. L., et al. 2017, MNRAS, 465, 2662

Smith, K. C. 1996, Ap&SS, 237, 77

Stibbs, D. W. N. 1950, MNRAS, 110, 395

Urpin, V. 2015, Astron. Nachr., 336, 266

Valdes, F., Gupta, R., Rose, J. A., Singh, H. P., Bell, D. J. 2004, ApJS, 152, 251

9.2 Websites

NASA/GSFC Astrophysics Data Facility
NASA Goddard Space Flight Center, Greenbelt MD 20771
http://fits.gsfc.nasa.gov/fits_standard.html
(visited: 2016, June 15, last update: Friday, 29-Jul-2016 14:13:12 EDT)

NIST Atomic Spectra Database (ver. 5.3)
National Institute of Standards and Technology, Gaithersburg, MD.
http://physics.nist.gov/PhysRefData/ASD/lines_form.html
(visited: 2017, May 25, last update: 2015)

9.3 Software & programming

9.3.1 `R` & `R`-packages

R Core Team (2013).
R: A language and environment for statistical computing.
R Foundation for Statistical Computing, Vienna, Austria.
http://www.R-project.org/.

Packages:

- Andrew Harris (2016).
  FITSio: FITS (Flexible Image Transport System) Utilities.
  R package version 2.1-0.
  https://CRAN.R-project.org/package=FITSio

9.3.2 `RStudio`

RStudio Team (2015).
RStudio: Integrated Development for R.
RStudio, Inc., Boston, MA URL
http://www.rstudio.com/.

9.3.3 `MKCLASS`

Richard Gray, Christopher Corbally
`MKCLASS`: An expert stellar spectral classification system
http://www.appstate.edu/~grayro/mkclass/

9.3.4 `fits viewer`

Fv: The Interactive FITS File Editor (2015)
https://heasarc.gsfc.nasa.gov/docs/software/ftools/fv/

A Appendix

A.1 Abstract in English

In this Master's Thesis I search for chemically peculiar stars of the upper main sequence in 14468 precalibrated, low resolution spectra $(R = 1300)$ from stars in the CoRoT fields IRa1 (4000 spectra), LRa1 (1659 spectra) and LRa2 (8809 spectra) obtained by Sebastian et al. (2012). Low quality spectra have been removed from the three samples of spectra because they contain too low signals, too many negative counts or rectangular-like patterns. I removed bad counts that occurred systematically and non systematically from the remaining spectra and interpolated all cut out counts. Spectra where wave-like patterns are present have been detected by fitting a heuristic function to the spectral counts distribution and those spectra have been removed from the samples. The samples contained multiple spectra per observed star, so only one spectrum per star has been kept in the samples. A basic algorithm to compute the radial velocities and the rest frame wavelengths of the spectra has been applied. The spectral types are derived via `MKCLASS` (Gray & Corbally 2014) or taken from Sebastian et al. (2012). Photometric magnitudes of the spectra are computed for two similar photometric systems (Δa, $\Delta a'$), a Δa and a $\Delta a'$ analysis is carried out for each of the three samples. The statistical outliers of each sample with $|\Delta a| > 2\sigma_{\Delta a}$ or $|\Delta a'| > 2\sigma_{\Delta a'}$ are investigated manually to confirm the chemical peculiarities or reject false detections. By doing so, 15 stars which are most likely chemically peculiar stars and further 158 stars which are possible chemically peculiar stars have been found.

A.2 Abstract in German

In dieser Masterarbeit suche ich in 14468 vorkalibrierten Sternspektren von geringer Auflösung ($R = 1300$), die aus den CoRoT Feldern IRa1 (4000 Spektren), LRa1 (1659 Spektren) und LRa2 (8809 Spektren) stammen und von Sebastian et al. (2012) aufgenommen worden sind, nach chemisch pekuliaren Sternen der oberen Hauptreihe. Spektren von schlechter Qualität wurden aus den drei Datensätzen herausgenommen, weil sie zu geringe Signale, zu viele negative Einträge oder rechtecksartige Muster beinhalteten. Ich entfernte systematisch wie nicht systematisch auftretende defekte Einträge aus den übrig gebliebenen Spektren und führte an all diesen ausgeschnittenen Positionen eine Interpolation durch. Spektren, wo wellenartige Muster präsent sind, wurden durch einen Fit einer heuristischen Funktion an die spektrale counts Verteilung detektiert und aus dem Datensatz entfernt. Die Datensätze beinhalten mehrere Spektren pro beobachtetem Stern, weshalb jeweils nur je ein Spektrum pro Stern in den Datensätzen verbleibt. Die Radialgeschwindigkeiten und Ruhenwellenlängen der Spektren wurden mittels eines einfachen Algorithmus berechnet. Die Spektraltypen der Spektren wurden durch `MKCLASS` (Gray & Corbally 2014) bestimmt oder von Sebastian et al. (2012) übernommen. Es werden die photometrischen Magnituden aller Spektren für zwei ähnliche photometrische Systeme (Δa, $\Delta a'$) berechnet und eine Δa und $\Delta a'$ Analyse durchgeführt. Statistische Ausreißer jedes Datensatzes mit $|\Delta a| > 2\sigma_{\Delta a}$ oder $|\Delta a'| > 2\sigma_{\Delta a'}$ werden manuell überprüft um chemische Pekuliaritäten zu bestätigen oder falsche Detektionen zu verwerfen. Auf diese Art und Weise wurden 15 Sterne, welche sehr wahrscheinlich chemisch pekuliare Sterne sind und weitere 158 Sterne, die mögliche chemisch pekuliare Sterne sind, gefunden.